COMPARATIVE SKIN TRANSCRIPTOME PROFILING OF CHANGTHANGI AND MUZZAFARNAGRI SHEEP

THESIS SUBMITTED TO THE

ICAR-NATIONAL DAIRY RESEARCH INSTITUTE, KARNAL

(DEEMED UNIVERSITY)

IN PARTIAL FULFILMENT OF THE REQUIREMENTSFOR THE AWARD OF THE DEGREE OF

MASTERS OF SCIENCE

IN

ANIMAL BIOTECHNOLOGY

BY

MAHANTHI VASU

B.Sc. B.Z.C

ANIMAL BIOTECHNOLOGY CENTRE

ICAR-NATIONAL DAIRY RESEARCH INSTITUTE

(DEEMED UNIVERSITY)

KARNAL-132001 (HARYANA), INDIA

2023

Regd. No. 20-M-BT-14

COMPARATIVE SKIN TRANSCRIPTOME PROFILING OF CHANGTHANGI AND MUZZAFARNAGRI SHEEP

BY

MAHANTHI VASU

THESIS SUBMITED TO THE

ICAR-NATIONAL DAIRY RESEARCH INSTITUTE, KARNAL

(DEEMED UNIVERSITY)

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE AWARD OF THE DEGREE OF

MASTERS OF SCIENCE

IN

ANIMAL BIOTECHNOLOGY

Approved by

8/9/2023

EXTERNAL EXAMINER

8/9/23

Dr. SONIKA AHLAWAT

MAJOR ADVISOR

MEMBERS, ADVISORY COMMITTEE

1. **Dr. SACHINANDAN DE**
   Principal Scientist
   Animal Biotechnology Center
   ICAR-NDRI, Karnal

   8/9/2023

2. **Dr. REENA ARORA**
   Principal Scientist
   Animal Biotechnology Division
   ICAR-NBAGR, Karnal

   13/9/23

3. **Dr. REKHA SHARMA**
   Principal Scientist
   Animal Genetic Resources Division
   ICAR-NBAGR, Karnal

   13/9/23

4. **Dr. SUMAN KAPILA**
   Principal Scientist
   Animal Biochemistry Division
   ICAR-NDRI, Karnal

   13/9/23

राष्ट्रीय पशु आनुवंशिक संसाधन ब्यूरो
NATIONAL BUREAU OF ANIMAL GENETIC RESOURCES
(I.C.A.R.)
Makrampur Campus, G.T. Road Bye Pass, Near Basant Vihar,
P.O. Box. No. 129. KARNAL - 132 001 (Haryana)

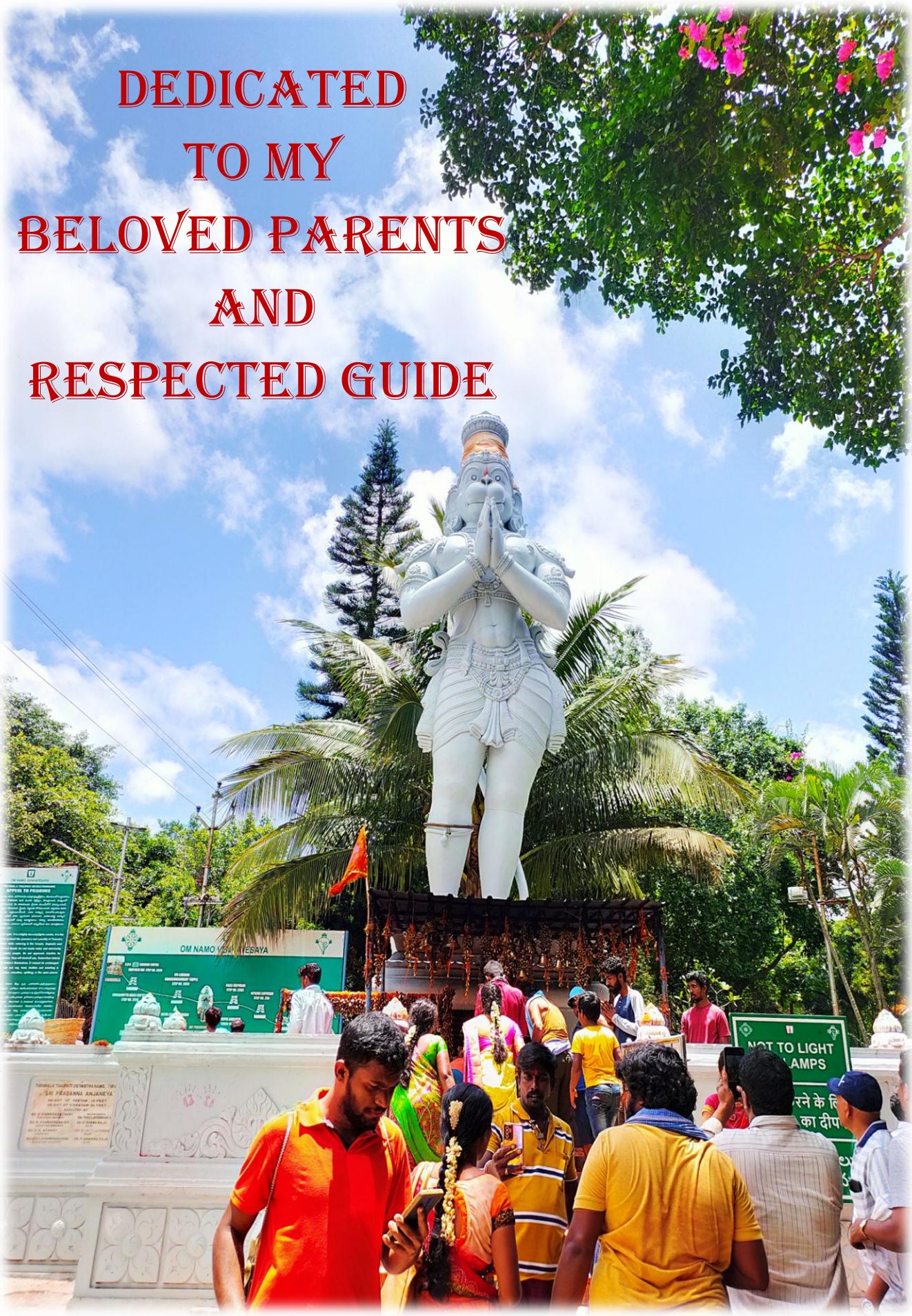
Dr. SONIKA AHLAWAT
Senior Scientist
(Animal Biotechnology Division)
ICAR-NBAGR, Karnal-132001, Haryana, India

CERTIFICATE

This is to certify that the thesis entitled "**COMPARATIVE SKIN TRANSCRIPTOME PROFILING OF CHANGTHANGI AND MUZZAFARNAGRI SHEEP**" submitted by Mr. **MAHANTHI VASU** towards the partial fulfilment of the requirement for the award of the degree of **MASTERS OF SCIENCE IN ANIMAL BIOTECHNOLOGY** of the **ICAR-NATIONAL DAIRY RESEARCH INSTITUTE (DEEMED UNIVERSITY)**, Karnal (Haryana), India, is a bonafide research work carried out by her under my supervision and guidance and no part of the thesis has been submitted for any other degree or diploma.

Dated: **28th July, 2023**

Dr. SONIKA AHLAWAT
MAJOR ADVISOR


DEDICATED
TO MY
BELOVED PARENTS
AND
RESPECTED GUIDE

ACKNOWLEDGEMENTS

Success is impossible without the blessings of almighty God and the elders. This is perhaps the most enjoyable and thoughtful moment to acknowledge those who made significant contribution in completing my Master's degree. Finally, the day has come tolook around and think about the people who were source of support and inspiration for me.Words in my lexicon fail to express my profound sense of gratitude to my mentor ***Dr. Sonika Ahlawat****, Senior Scientist, Animal Biotechnology Division, ICAR-NBAGR, Karnal,for her inspiring guidance, outstanding cooperation and constant encouragement during the entire course of study. Her constant support, critical appreciation and care gave a glitterof confidence to my work. Her intense hard work and countless efforts to provide every possible facility in the lab made this strenuous task achievable. I am grateful to my advisory committee members,* ***Dr. Reena Arora****, Principal Scientist, ICAR-NBAGR and* ***Dr. Rekha Sharma****, Principal Scientist, ICAR-NBAGR, for their appreciation, encouragement and precious suggestions during my scientific endeavor.I owe my sincere thanks to* ***Dr. Sachinandan De****, Principal Scientist, Animal Biotechnology Center, ICAR-NDRI, and* ***Dr. Suman Kapila,*** *Principal Scientist, Animal Biochemistry Division, ICAR-NDRI, for their critical evaluation along with prolific ideas and time to time suggestions that helped me in the successful execution of my work.*

My sincere thanks to ***Dr. Dheer Singh,*** *Director, ICAR-NDRI, Karnal and****Dr. B.P. Mishra,*** *Director, ICAR-NBAGR, Karnal for providing all the necessary facilitiesto carry out the research. Financial assistance from ICAR in the form of institute fellowshipis gratefully acknowledged. Also, a heartfelt thanks to* ***Dr. Vikas Choudhary*** *and* ***Mr. Rakesh****, from District Disease Diagnostic Laboratory, Karnal who have constantly helped me in getting the samples to carry out my research in a smooth manner. Further, co-operation, support and help of my lab-mates,* ***Ashish Kumar, Mandeep Kaur, Pooja Chhabra, Upasana*** *and* ***Sushma Pandey*** *during my research work isthankfully acknowledged. I would like to thank* ***Neeraj bhaiya, Rishi bhaiya*** *and* ***Rohit bhaiya*** *for providing day to day support during my lab work. A heartfelt gratitude to my wonderful seniors* ***Amarjeet, Rashmeet*** *for always supporting, caring and encouraging me in odd times. Where emotions are involved, words cease to mean; words don't seem adequate to express my deepest feeling for the love, care, affection, concern and support rendered by my parents* ***Mahanthi Satyam, Narayanamma,*** *my grandparents* ***Mahanthi lakshmu*** *and* ***Suramma*** *and my beloved brother* ***Lucky*** *as well as my dear friends,* ***Shubham, Gagan****. I am deeply indebted to them for always being there for me. Finally, i wish to express my deep gratitude to* ***Dr. Tunagula.Srikaran Uddesh*** *(M.B.B.S, M.D (Internal Medicine)) for financial support to studies.*

For their immutable inspirationand strength to undertake this master's study and successful completion of the same, I feel thankful to them. I am proud to have them as my parents and there are no words but only feelings to honorably pay my regards to them for their ceaseless perspiration, encouragement, moral support and for instilling in me focused attitude towards my goal since childhood.

Again, I want to show my sincere and heartfelt gratitude and appreciation to a wonderful guide, ***Dr. Sonika Ahlawat*** *for providing me with the homely environment, guidance and counseling I needed to succeed in the M.Sc. program. She has been a great mentor in mapping my M.Sc. journey, advising on a research topic, connecting me with theresources I need, being available and responding to my emails and questions. My supervisorisa doyen in my area of research interest and her wealth of experience has been the main driver of my program.*

Finally, I am extremely thankful to The Almighty God for his grace and blessings, which provided me the courage to sail through this journey. Thank you, God!

Date: **28th July, 2023** **MAHANTHI VASU**

ABSTRACT

This study investigated the skin transcriptome of Changthangi and Muzzafarnagri sheep through RNA sequencing to identify differentially expressed genes and pathways contributing to fiber quality, coat color and adaptation in different environments. The Changthangi animals included in this study were from the high altitude region of Leh, Ladakh and had grey skin and black coat color, while Muzzafarnagri sheep from the plains of Uttar Pradesh had white skin and coat. Global gene expression data was generated for four biological replicates of each breed. The reads were aligned with the sheep reference genome resulting in mapping percentage of 98.27% to 99.37% in different samples. A total of 149 genes were up-regulated and 2139 were down-regulated in Changthangi sheep as compared to Muzzafarnagri sheep (padj ≤ 0.05 and Log2 fold change ≥ 1.5). Gene Ontology analysis of the up-regulated genes showed enrichment of terms related to melanin biosynthesis and developmental pigmentation in Changthangi sheep. The enriched KEGG pathways were tyrosine metabolism and metabolic pathways. The melanogenesis related genes with higher expression in Changthangi sheep were *TYR, TYRP1, DCT, SLC45A2, PMEL, MLANA,* and *OCA2*. In addition, four genes, namely *KRTAP6, KRTAP7, KRTAP13,* and *KRTAP21* associated with fiber quality showed more abundant expression in the fine wool producing sheep from Ladakh. The results of RNA sequencing were validated by real time PCR of 10 genes governing fiber quality and coat color using *ACTB* and *PPIB* as reference genes. These reference genes were selected after carrying out a comprehensive assessment of 18 potential reference genes (*ACTB, BACH1, B2M, GAPDH, HMBS, HPRT1, PGK1, PPIA, PPIB, RPLP0, RPL19, RPS9, RPS15, RPS28, SDHA, TBP, UXT,* and *YWHAZ)* in 10 ovine tissues (muscle, skin, kidney, liver, intestine, rumen, lung, testis, heart, and spleen) from 5 sheep to identify stably expressed genes across different tissues. The ranking of genes was determined using four algorithms, namely, geNorm, NormFinder, BestKeeper, and Delta Ct (ΔCt). Similar trend in ranking was obtained using these approaches. Finally, comprehensive ranking of the potential RGs was achieved using RefFinder, which integrates the output obtained from different algorithms. This study recommends the use of *ACTB, PPIB, BACH1* and *B2M* as internal control genes in qPCR experiments involving different sheep tissues. In conclusion, candidate genes associated with fiber quality, coat color and adaptation to high altitude hypoxic conditions were identified in Changthangi sheep.

सारांश

वर्तमान शोध में फाइबर की गुणवत्ता, कोट के रंग और विभिन्न वातावरणों के अनुकूलन में योगदान देने वाले विभेदित जीन और मार्गों की पहचान करने के लिए आरएनए अनुक्रमण के माध्यम से चांगथांगी और मुज़फ़्फ़रनगरी भेड़ की त्वचा प्रतिलेख की जांच की गई थी। इस अध्ययन में शामिल किए गए चांगथांगी पशु लेह, लद्दाख के ऊंचाई वाले क्षेत्र से थे और उनकी त्वचा ग्रे और कोट का रंग काला था, जबकि उत्तर प्रदेश के मैदानी इलाकों की मुजफ्फरनगरी भेड़ की त्वचा और कोट सफेद थे। प्रत्येक नस्ल की चार जैविक प्रतिकृतियों के लिए वैश्विक जीन अभिव्यक्ति डेटा तैयार किया गया था। रीड्स को भेड़ संदर्भ जीनोम के साथ संरेखित किया गया जिसके परिणामस्वरूप विभिन्न नमूनों में 98.27% से 99.37% तक मैपिंग प्रतिशत प्राप्त हुए थे । मुजफ्फरनगरी भेड़ ($P \leq 0.05$ और Log2 fold change ≥ 1.5) की तुलना में चांगथांगी भेड़ में कुल 149 जिन की अभिव्यक्ति अधिक एवं 2139 जिन की अभिव्यक्ति काम थी I अप-विनियमित जीन के जीन ओन्टोलॉजी विश्लेषण ने चांगथांगी भेड़ में मेलेनिन जैवसंश्लेषण और विकासात्मक रंजकता से संबंधित शब्दों का संवर्धन दिखाया। समृद्ध के.ई.जी.जी. मार्ग टायरोसिन उपापचय और चयापचयी मार्ग थे। चांगथांगी भेड़ में उच्च अभिव्यक्ति वाले मेलानोजेनेसिस से संबंधित जीन *TYR, TYRP1, DCT, SLC45A2, PMEL, MLANA,* और *OCA2* थे। इसके अलावा, फाइबर गुणवत्ता से जुड़े चार जीन, अर्थात् *KRTAP6, KRTAP7, KRTAP13* और *KRTAP21* ने लद्दाख की महीन ऊन पैदा करने वाली भेड़ों में अधिक प्रचुर अभिव्यक्ति दिखाई थी। आरएनए अनुक्रमण के परिणामों को संदर्भ जीन के रूप में ACTB और PPIB का उपयोग करके फाइबर गुणवत्ता और कोट रंग को नियंत्रित करने वाले 10 जीन के रियल टाइम पी.सी.आर. द्वारा मान्य किया गया था। विभिन्न ऊतकों में स्थिर रूप से व्यक्त जीन की पहचान करने के लिए 5 भेड़ों के 10 ऊतकों (मांसपेशियों, त्वचा, वृक्क, यकृत, आंत, रुमेन, फेफड़े, वृषण, हृदय, और प्लीहा) में 18 संभावित संदर्भ जीन (*ACTB, BACH1, B2M, GAPDH, HMBS, HPRT1, PGK1, PPIA, PPIB, RPLP0, RPL19, RPS9, RPS15, RPS28, SDHA, TBP, UXT, and YWHAZ)* का व्यापक मूल्यांकन करने के बाद इन संदर्भ जीन का चयन किया गया था। जीन की रैंकिंग चार एल्गोरिदम, अर्थात् जीनॉर्म, नॉर्मफाइंडर, बेस्टकीपर और डेल्टा सीटी (ΔCt) का उपयोग करके निर्धारित किया की गई I इन एल्गोरिदम का उपयोग करके रैंकिंग में समान रुझान प्राप्त किया गया। अंत में, RefFinder का उपयोग करके उपयुक्त जीन की व्यापक रैंकिंग हासिल की गई, जो विभिन्न एल्गोरिदम से प्राप्त उत्पादन को एकीकृत करता है। यह अध्ययन विभिन्न भेड़ के ऊतकों से जुड़े क्यू.पी.सी.आर. प्रयोगों में आंतरिक नियंत्रण जीन के रूप में *ACTB*, *PPIB*, *BACH1* और *B2M* के उपयोग की सिफारिश करता है। निष्कर्ष में, चांगथांगी भेड़ में फाइबर की गुणवत्ता, कोट के रंग और उच्च ऊंचाई वाले हाइपोक्सिक स्थितियों के अनुकूलन से जुड़े कैंडिडेट जीन की पहचान की गई।

CONTENTS

LIST OF FIGURES

LISTOFTABLES

CHAPTER-1

Introduction

INTRODUCTION

Sheep are an important element of the agricultural economy because of their multifaceted uses for wool, meat, skin, and manure, especially in India's arid, semi-arid, and hilly regions. India possesses the second-largest sheep population in the world (74.26 million) and its population has increased by 14.13% between the last two censuses (Livestock census, 2019). The sheep genetic resources of India are represented by 44 defined breeds and many lesser-known populations that have adapted to diverse agro-climatic regions through natural selection (Sharma *et al.,* 2020). They are primarily reared for carpet wool, apparel wool, and meat. Due to low resource need, sheep farming is particularly significant for small and marginal farmers in the harshest areas and contributes in their livelihood security (Rout *et al.,* 2018). Rapid advancements in unravelling of the sheep genome have opened vistas for underpinning the genetic mechanisms controlling various phenotypic traits associated with production and adaptation (Kalds *et al.,* 2022). Among various economically important products from ovines, fibre/wool is an important agricultural commodity that contributes immensely to the global economy. Wool follicles across the skin surface produce wool through a complex process involving many proteins (Galbraith, 2010). The quality, quantity and color of wool/fleece are regulated by both heredity and environment. The green movement and customer demand for natural products have led to an increased interest in the natural coat colours of sheep (Yang *et al.,* 2016). Gaining a comprehensive grasp of the processes governing wool fineness and color can help to enhance the economic advantages of the sheep industry (Tian *et al.,* 2022). Since the beginning of the 20th century, attempts have been made to comprehend the genetic mechanism of sheep coat color and consequently, multiple alleles at several loci have been shown to govern the trait (Wu *et al.,* 2021). Similarly, many scientific investigations have identified the genes and molecular mechanisms associated with fine and coarse wool phenotypes in sheep (Zhang *et al.,* 2017; Shi *et al.,* 2022).

The adaptive capability of sheep across diverse climates ranging from high altitude hypoxic conditions to heat stressed environments is unparalleled and has facilitated their existence across widespread geographical regions (Wei *et al.,* 2015). Adaptation to climatic extremes in sheep has been documented to have physiological, behavioral, and morphological basis. However, the genetic landscape of the animals has the most profound influence in acclimatization to varying agro-climatic zones (Berihulay *et al.,* 2019). The skin of animal plays an important role in adaptive evolution because it directly interacts with the external

environment (Huang *et al.,* 2016). Animal adaptability can be described as the degree to which an organism, population, or species can continue to become adjusted to a variety of environments by physiological or genetic processes (Prayaga and Henshall, 2005; Barker, 2009). Skin is the largest thermoregulatory organ and has vital roles such as protection from adverse environments, microbial invasion, and mechanical damage (Ahmad and Mukhtar, 2004). Native populations of highland Tibetans exposed to high ultraviolet (UV) radiation at high altitudes have been documented to be under significant selective pressure resulting in adaptations in the form of skin pigmentation as a response (Yang *et al.,* 2022). The pigmentation of the skin and hair is caused by melanocytes which are critical in providing protection against the harmful effects of ultraviolet light (Lin and Fisher, 2007).

A clear understanding of the genetic architecture of phenotypic traits associated with production and adaptation offers the opportunity to improve production efficiency and diversity in breeding programs. Despite many efforts across the globe to decipher the molecular basis of economic traits in sheep, there is a dearth of scientific knowledge regarding the key genes and molecular mechanisms governing wool traits and adaptation to different habitats from the Indian subcontinent. Therefore, this study sought to compare the skin transcriptome profiles of two Indian sheep breeds varying in phenotype, utility and geographical location. Among the ovine genetic resources of India, the Changthangi breed is found in the Changthang region of the Leh district of Jammu and Kashmir. Referred to as Changluk locally by the Changpa nomadic clan, it is raised for a variety of reasons, including mutton, wool, pelt, and manure. Changthangi sheep is acknowledged for its adaptation to high altitude hypoxic conditions of Leh (Khan *et al.,* 2022). Unlike other Indian sheep breeds, Changthangi sheep possesses double coat with down fibers that provide protection against the inhospitably cold climate of Leh. The extremely fine undercoat presents an opportunity to manufacture fine textured luxurious fibers from this sheep breed (Malik *et al.,* 2021). On the contrary, Muzzafarnagri sheep, which is indigenous to the western Uttar Pradesh produces coarse carpet-quality fleece. It is one of the biggest and heaviest sheep breeds in India which is primarily kept for its meat and is adapted to arid and semi-arid environments (Mandal *et al.,* 2000).

The high-throughput RNA sequencing platforms have provided new opportunities to delineate the gene expression profiles at genome-wide scale and have delivered valuable insights into the regulatory mechanisms governing the economically relevant phenotypic traits in sheep (Kaur *et al.,* 2020; Arora *et al.,* 2021; Michailidou *et al.,* 2021). The aim of this study

was to compare the skin transcriptomes of Changthangi and Muzzafarnagri sheep to identify differentially expressed genes and molecular pathways contributing to variation in fiber quality, coat color and adaptation to different agro-ecological regions.

Deep sequencing approaches have unravelled the global gene expression profiles and identified molecular pathways governing the economically relevant phenotypic traits in sheep (Kaur *et al.,* 2020; Arora *et al.,* 2021; Michailidou *et al.,* 2021). However, the high throughput expression data from RNA sequencing and/or microarray-based experiments needs to be validated by the real-time reverse transcription PCR method (qPCR) (Schmittgen and Livak, 2008; Kumar *et al.,* 2021; Kaur *et al.,* 2023). By virtue of being an extremely sensitive, simple, cost-effective, and efficient technique, qPCR is a routinely used molecular method for assessing the expression profiles of target genes (Sonowal *et al.,* 2022). However, many variables such as amount of template, RNA integrity, and enzymatic efficiencies influence the precision of qPCR. Therefore, normalization of expression data of target genes using stable reference genes is a critical step to ensure the accuracy of relative quantification using qPCR (Kaur *et al.,* 2023).The reference gene is subjected to the same experimental procedure as the target gene during qPCR quantification in order to normalise the results and reduce potential errors.

Housekeeping genes are often used as internal controls because they are essential for maintaining cellular homeostasis and are generally accepted to have stable spatial or temporal expression across tissues. Under varying experimental conditions, the expression of reference genes must remain consistent between cells of diverse tissues (Thellin *et al.,* 1999). Yet, several investigations have shown that a variety of circumstances have a major impact on their expression kinetics (Yang *et al.,* 2014). Therefore, selection of an appropriate reference gene or a panel of normalizer genes is the most important step for accurate interpretation of qPCR results (Bustin, 2002; Kaur *et al.,* 2023). Actins, tubulins, 18S and 28S ribosomal RNA, glyceraldehyde-3-phosphate dehydrogenase, albumins, ubiquitin, and hypoxanthine-guanine phosphoribosyl transferase are the most often used reference genes (Thellin *et al.,* 1999). Research has revealed that the transcription levels of frequently used housekeeping genes such as β-actin and *GAPDH* is not always constant (Hamalainen *et al.,* 2001). It has become evident that no reference gene can be taken to be acceptable for all experimental designs after multiple instances of variation in the expression of the professed housekeeping genes (Schmittgen and Zakrajsek, 2000; Huggett *et al.,* 2005). As a single gene cannot serve as a reference gene for diverse biological systems or processes, it is necessary to establish a number of reference

genes (GarciaVallejo *et al.,* 2004; Robinson *et al.,* 2007). Many statistical techniques, including geNorm, BestKeeper, NormFinder, comparative ΔCt method have been developed to evaluate the stability of various housekeeping genes (Vandesompele *et al.,* 2002; Andersen *et al.,* 2004; Pfaffl *et al.,* 2004; Silver *et al.,* 2006). The freely available online tool, RefFinder combines the geometric means of these four most widely used stability algorithms resulting in a comprehensive ranking (http://leonx ie.esy.es/RefFinder/). Successful attempts have been made to identify the most acceptable reference genes for ovine blood (Peletto *et al.,* 2011), neurological tissue (Lyahyai *et al.,* 2009), pulmonary tissue (Passmore *et al.,* 2009), spleen, lung, ileum (Garcia-Crespo *et al.,* 2005) for the estimation of gene transcription. This study also attempted to identify most stable reference genes that may be used to normalize qRT-PCR results for experiments involving 10 ovine tissues (muscle, skin, kidney, liver, intestine, rumen, lung, testis, heart, spleen). Eighteen candidate reference genes were evaluated for their expression stability, including Beta-actin (*ACTB*), BTB and CNC homology1 (*BACH1*), Beta-microglobulin (*B2M*), Glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*), Hydroxymethylbilane synthase (*HMBS*), Hypoxanthine-guanine phosphoribosyl transferase (*HPRT1*), Phosphoglycerate kinase 1 (*PGK1*), Peptidylprolyl isomerase A (*PPIA*), Peptidylprolyl isomerase B (*PPIB*), Ribosomal protein lateral stalk subunit P0 (*RPLP0*), Ribosomal Protein L19 (*RPL19*), Ribosomal protein S9 (*RPS9*), Ribosomal protein S15 (*RPS15*), Ribosomal protein S28 (*RPS28*), Succinate dehydrogenase complex flavoprotein subunit A (*SDHA*), TATA binding protein *(TBP*), Ubiquitously expressed prefoldin like chaperone (*UXT*), and Tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein zeta (*YWHAZ*).

Thus, the present study was carried out with the following objectives:

Objective1: To compare the skin transcriptome profiles of Changthangi and Muzzafarnagri sheep.

Objective2: To validate the differentially expressed genes by quantitative real time PCR.

CHAPTER–2

Review of Literature

REVIEW OF LITERATURE

2.1 Breed

Changthangi breed is generally called Changluk, meaning sheep of northern plateau. These are large-framed, well-built animals with excellent fleece that has an incredibly long staple (Malik *et al.,* 2021). On the other hand, Muzzafarnagri sheep is one of the tallest and heaviest mutton breeds in the country. The breed gets its name from the Muzaffarnagar area of western Uttar Pradesh, where it was first domesticated (Mandal *et al.*, 2000) (Figure 2.1).

Changthangi sheep **Muzzafarnagri sheep**

Figure 2.1: Animals of Changthangi and Muzzafarnagri breeds

2.2 Distribution

Changthangi sheep can be found in the Changthang region of the Leh district of Jammu and Kashmir and is reared by the Changpa nomadic clan (Malik *et al.,* 2021). Muzzafarnagri sheep are mostly found in Western Uttar Pradesh's Muzaffarnagar district and neighboring areas. However, some areas of the states of Delhi and Haryana also include a sizable number of animals (Mandal *et al.,* 2000) (Figure 2.2).

Figure 2.2: Native tracts of Changthangi and Muzzafarnagri sheep.

2.3 Importance

The Changthangi breed is raised for mutton, wool, manure, and dung energy. Pashmina goats and Changthangi sheep contribute to the economic stability of Changpas. The breed has two fleece coats; the outer course fiber and fine inner coat, which protects the breed during harsh winters when temperature drops to -40 degree Celsius. The wool is used for weaving of the woolen products (Malik, *et al.,* 2021). The Muzzafarnagri sheep is mostly used to make mutton but also produces a good amount of wool. It is recognized for its quicker development rate and strong adaptability to semi-arid environments. However, the wool it produces is coarse type (https://www.sheepfarm.in/muzaffarnagar-sheep).

2.4 Transcriptome

The collection of all RNA transcripts, both coding and non-coding, found in a single cell or in a population of cells is known as the transcriptome (Figure 2.3). The term "transcriptome," which is a combination of the term's "transcript" and "genome," refers to the process of producing transcripts during the biological transcription process. The field of genetic sequencing technology has grown rapidly during the last few decades (Shendure *et al*., 2017). As a result of increased throughput, higher accuracies and lower costs, there has been an exponential growth in genomic sequence databases (Lathe *et al.,* 2008). Numerous

studies have been conducted in various organisms on transcriptomics, which provide significant insights on the structure, expression, and control of genes (Jain *et al.*, 2006). Because of the underlying sequencing technology's fast progress, the transcriptomics investigations have improved significantly (Wang *et al.*, 2009). cDNA libraries that were released in the 1980s marked the beginning of the transcriptome annotation process. After that, the development of high-throughput technologies gave rise to quicker and more effective means of getting information about the transcriptome. The first transcriptome sequencing was done by Walter Fiers and his colleagues in the year 1970s, and they successfully sequenced the 3,569 nucleotide-long transcriptome of the bacteriophage MS2 (Fiers *et al.*, 1976). Transcriptomics is now a viable choice for several research projects involving primary agricultural animals, including sheep, goats, poultry, pigs, bovines and fish which offers information about uncovered prospective candidate genes linked to meat quality, fiber quality, milk, adaptation, lactation, reproductive effectiveness, and disease response (Parreira and de Sousa Araujo, 2018).

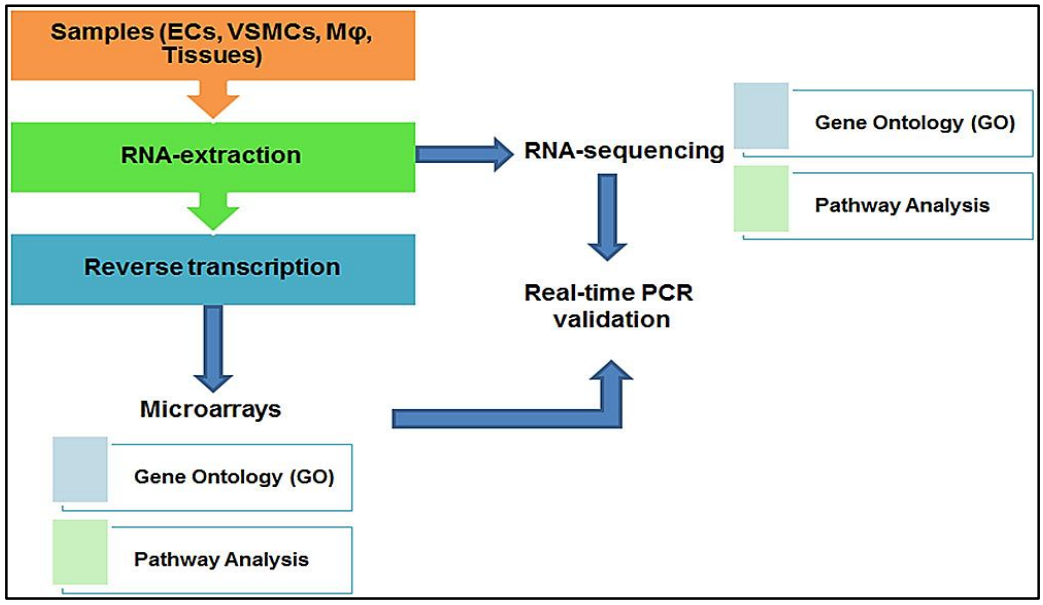


Figure 2.3: Transcriptome analysis

Expression sequencing tags (ESTs), serial analysis of gene expression (SAGE), cap analysis of gene expression (CAGE), massive parallel signature sequencing (MPSS), long SAGE, DNA microarray are employed to analyse the transcriptome (Livak and Schmittgen, 2001). Microarray technology utilizes the principle of molecular hybridization to mount hundreds of known partial sequence DNA probes on a solid support slide or nylon membrane

using an automated instrument, Arrayer. Through a single hybridization, a sizable number of genes may be statistically identified. Another emerging technique, RNA-Seq, has been gaining popularity over the last few years. RNA-Seq is one of the most important applications of next-generation sequencing technology, and one of the most important tools for transcriptome study. It overcomes the drawbacks of microarray and brings deeper insight for transcriptome research. NGS, or high-throughput sequencing, is exemplified by the 454 technologies from Roche, the Solexa technology from Illumina, and the SOLiD technology from ABI.

Technology		**Principle**	**Signal**	**Throughput**	**Resolution**	**Sequencing Read**	**Initial amount of nucleic acids**
DNA Sequencing	First Generation Sequencing	Dideoxy chain-termination method (Sanger Sequencing)	Digital signal	Low	Single bp	Long	High
	Next Generation Sequencing	High Throughput Sequencing	Digital signal	High	Single bp	Short	Low
Microarray		Hybridization	Fluorescence	High	Several 100 bp	Long	High
RNA Sequencing		High throughput Sequencing	Digital signal	High	Single bp	Short	Low

2.5 Techniques for transcriptomics studies

2.5.1 Expressed Sequence Tags (ESTs)

Expressed Sequence Tags (ESTs) are short lengths of DNA sequences (200 to 800 nucleotides) obtained by sequencing either one or both ends of an expressed gene. ESTs are generated through individual sequencing procedures conducted on randomly selected clones from a cDNA library (Parkinson and Blaxter 2009). By utilizing ESTs, researchers can perform structural, functional genomic analysis, comparative genetics studies, and gene comparison across species including those with significant genetic differences (Hatey *et al.*,1998). The 5' ESTs were generated by sequencing the initial few bases of the cDNA, which are relatively consistent, whereas the 3' ESTs were obtained by sequencing the terminal

segment of the cDNA molecule, encompassing the non-coding or untranslated regions (UTR). Human ESTs have contributed significantly to the identification of novel gene family members, localization of polymorphisms, and the characterization of tissue or species specific gene expression (Gill *et al.*, 2000). The GenBank section known as dbEST provides access to sequence information and additional details regarding single pass cDNA sequences from various species, specifically focusing on data generated through single pass reads.

2.5.2 Serial Analysis of Gene Expression (SAGE)

Serial analysis of gene expression (SAGE) is a valuable technique that provides quantitative assessment of the number of genes expressed within a specific cell type (Tuteja and Tuteja, 2004). The SAGE method involves the sequential joining of long DNA molecules with distinct sequence tags, typically ranging from 9 to 10 base pairs in length. SAGE is an advanced technology that not only enables the generation of a comprehensive gene expression profile for a specific cell or tissue type but also facilitates the comparison of expression profiles between two cells maintained under different conditions. This comparison helps in the identification of a specific set of genes that are associated with different conditions (Yamamoto *et al.*, 2001). SAGE studies have been used to evaluate the developmental production attributes in sheep, goat, cattle, parasites, plants, and yeast (Peters *et al.*, 1999; Anisimov, 2008).

2.5.3 Cap Analysis of Gene Expression (CAGE)

Cap analysis gene expression (CAGE) is an alternative variant of the SAGE method that provides an information of the messenger RNA population specifically at the 5' end in a given biological sample. By utilizing single base-pair resolution, the CAGE approach enables the identification and monitoring of transcription start site activity throughout the genome. This technique facilitates the detection of active promoter and enhancer regions, enables gene network expression analysis, and holds potential for identifying crucial components in immunotherapy (Shiraki *et al.*, 2003; Morioka *et al.*, 2020).

2.5.4 DNA Microarray

DNA microarrays consist of numerous small DNA patches affixed to a surface, facilitating the analysis of gene expression levels for hundreds of genes within a single sample. It can help researchers in gaining fresh understanding of diseases, comparative studies, pharmaceutical development, and environmental studies. In order to assess the relative amounts of nucleic acid

species within a solution, the targets make bonds with the probes on the DNA array through hybridization. A solution containing a variety of labelled nucleic acids is probed using DNA arrays. By expanding the application of this method to a vast number of DNA spots, an array can be employed to quantify numerous distinct nucleic acid sequences present in the solution (Bumgarner, 2013). DNA microarrays have been utilized for genotyping, identification of disease-relevant genes, mutation analysis, SNP screening, chromosomal abnormality detection, and post translational modification analysis (Yoo *et al.,* 2009).

2.5.5 RNA Sequencing

RNA-Seq was first developed in mid 2000s with the advent of next-generation sequencing technology. The first manuscripts that used RNA-Seq even without using the term includes those of prostate cancer cell lines (2006), *Medicago truncatula* (2006), maize (2007), and *Arabidopsis thaliana* (2007), while the term "RNA-Seq" itself was first mentioned in 2008. Early analyses of gene expression depended on low-throughput techniques that can only measure a single transcript, such as northern blots and quantitative polymerase chain reaction (qPCR). Transcriptomics, as it is more often known, is the process of measuring gene expression across the whole genome (Kukurba *et al.,* 2015). RNA sequencing is the process of sequencing the entire mRNA transcripts. RNA sequencing is a next-generation, high throughput RNA sequencing (Figure 2.4) and quantification method used for studying the gene expression. It captures both well-known and novel characteristics. It allows researchers to find biomarkers in the widest range of transcripts possible and provides a thorough understanding of phenotypes of interest (Li *et al.,* 2015). Since the introduction of high-throughput sequencing in 2005, it has become possible to perform in-depth research to unravel the genome and transcriptome and to comprehend biological activities at the molecular level.

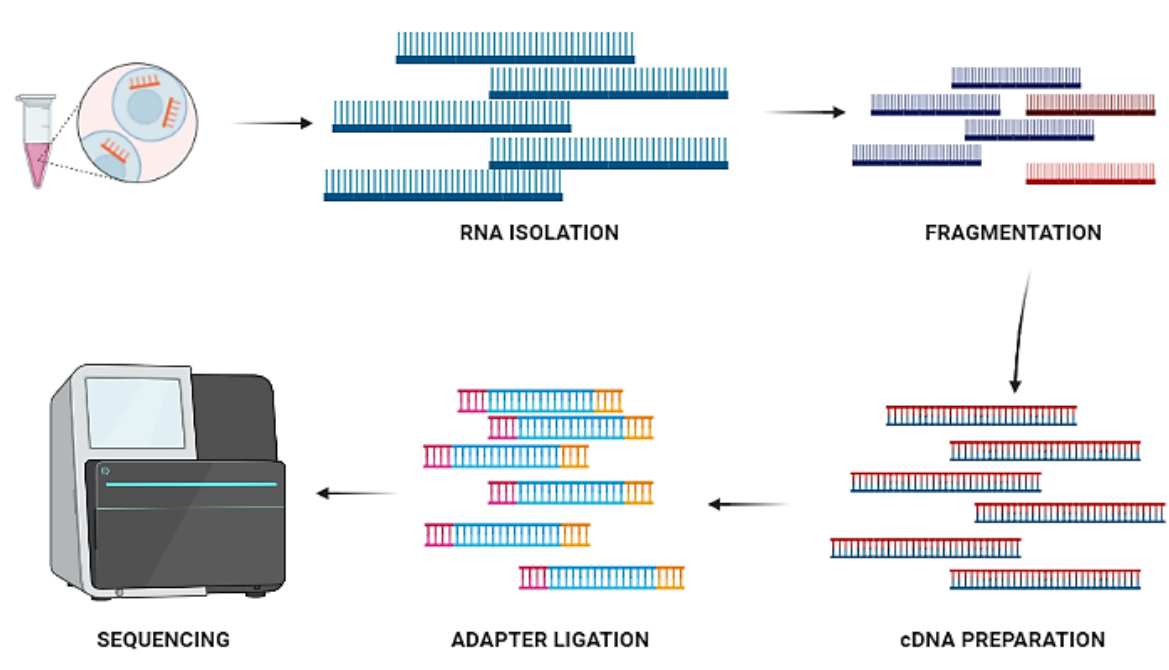


Figure 2.4:Schematic overview of RNA-Sequencing

2.6 Next-generation sequencing platforms

In the year 1953, Watson and Crick utilized the crystallographic findings of Rosalind Franklin and Maurice Wilkins to identify the three-dimensional structure which gave a theoretical support to DNA replication and protein processing (Watson and Crick, 1953). However, it took a considerable amount of time for scientists to achieve the ability to sequence the DNA. The techniques employed to determine the sequence of protein chains did not readily translate to the study of nucleic acids, as DNA molecules were significantly longer and composed of fewer, more uniform components (Hutchison, 2007). Sanger and Maxam-Gilbert are credited with establishing the field of DNA sequencing which was termed Sanger Sequencing (First-generation sequencing). The phiX174 genome, which has a length of 5374 base pairs (Sanger and Coulson, 1975), and the bacteriophage genome, which has a length of 48501 base pairs (Sanger *et al.*, 1980) were the first genomes to be sequenced using the Sanger method. From last three decades, Sanger sequencing has been extensively utilized and continues to be employed for single or low-throughput DNA sequencing, but it is difficult to accelerate the analysis while sequencing of complex genomes, such as the genomes of animal and plant species, and sequencing is still very expensive and time-consuming (Kchouk *et al.*, 2017). In

2005, a second generation of sequencers emerged, addressing the limitations of the first generation. The generation of many millions of short reads in parallel, acceleration of the sequencing process over first-generation sequencing, low cost of sequencing, and direct detection of the sequencing output without the need for electrophoresis are the key characteristics of second-generation sequencing technology. Second-generation sequencing technology revolutionized the field by enabling the generation of millions of short reads. simultaneously, accelerating the sequencing processes, reducing sequencing costs, and allowing for direct detection of sequencing output without the need for electrophoresis.

2.6.1 RNA sequencing technology platforms

Platform	Company	Read length (bases)	Run time	Volume per Run	Cost	Template preparation	Sequencing chemistry
Sanger	Life sciences	800 bp	2 h	1 read	$2400 million bases	Bacterial cloning	Dideoxynucleosides terminator
Roche 454 pyrosequencing	454 life sciences	700 bp	<24 h	0.7 gb	$10 per million bases	Emulsion PCR	Sequencing by synthesis, pyrosequencing
Illumina HiSeq	illumina	100 bp	3-10 days	120-1500 gb	$0.02-0.07 per million bases	Bridge PCR	Reversible terminator sequencing
Illumina MiSeq	Illumina	100 bp	1-2 days	0.3-15 gb	$0.13 per million bases	Bridge PCR	Reversible terminator sequencing
Ion torrent	454 life sciences	200-500 bp	4-5 h	600 Mb; 11 Mb	$300 to 750 per run	Emulsion PCR	Sequencing by synthesis
SMRT	Pacific biosciences	>900 bp	1-2 h	0.5-1 gb	$2 per million bases	No need	Sequencing by synthesis
Helicos sequencing	Helicos biosciences	25-60 bp	8 days	21-35 gb	$0.01 per million bases	No need	Hybridization and synthesis
Nanopore sequencing	Oxford nanopore technologies	Up to 98 kb	48/72 h	Up to 30 gb	<$ 1 per million bases	No need	nanopore

2.7 Transcriptome studies in farm animals

Transcriptomics is now a viable choice for several research projects involving primary agricultural animals, including sheep, goats, poultry, pigs, bovines and fish which offers an information about uncovered prospective candidate genes linked to meat quality, fiber quality, milk, adaptation, lactation, reproductive effectiveness, and disease response (Parreira and de Sousa Araujo, 2018). The best natural fibre in the world, pashmina, comes from Changthangi goat secondary hair follicles. By employing transcriptome profiling, Ahlawat *et al.,* 2020 explored the metabolic and gene networks associated with pashmina production by comparing

skin samples from Changthangi and Barbari goats, thereby unraveling the molecular drivers of the natural fibre, Pashmina.

Meat production plays a vital role in the economic aspect of animal husbandry and serves as a significant benchmark in sheep breeding. A recent study by Kumar *et al*., 2021 investigated the expression of genes in the biological replicates of the longissimus thoracis muscles from Barbari and Changthangi goats. Distinct variations were observed in the expression of genes related to lipid metabolism (*FASN, SCD, THRSP, DGAT2,* and *FABP3*) between the breeds. Furthermore, through gene co-expression network analysis, highly connected genes associated with the triacylglycerol biosynthesis pathway were found in the Barbari goat. This study provided an insight into the differential expression of genes in longissimus thoracis muscles between Barbari and Changthangi goats that are adapted to and reared in different agro-climatic regions. Another study by Arora *et al.*, 2019 compared the muscle transcriptome profile of the popular Indian meat sheep breed Bandur and local sheep. The highly linked genes (*CNOT2, CNOT6, HSPB1, HSPA6, MAP3K14,* and *PPARD*) discovered by network analysis were identified as crucial regulators of skeletal muscle fatty acid metabolism, cellular stress, and energy metabolism.

India holds the distinction of producing the highest per capita milk output globally, largely attributed to the significant contributions made by cattle and buffalo. Ahlawat *et al.,* 2021 utilized RNA sequencing to explore the transcriptome profile of milk somatic cells derived from Sahiwal cattle and Murrah buffaloes. Genes involved in the biological functions such as lactation, immunological response, cellular oxidant detoxification, and hormonal response were found to have the highest transcript abundance in both species. Interestingly, the study revealed that buffaloes possess more robust immune systems compared to cattle, which could potentially contribute to their decreased susceptibility to mastitis. Using RNA sequencing, the milk transcriptome from three lactation phases of Murrah buffalo was examined to investigate how genes and pathways interact. In the early stage, genes related to milk production and secretion were activated and their expression steadily decreased as lactation progressed (Arora *et al.,* 2019).

High ambient temperatures and humidity cause heat stress, which has an impact on health and productivity of animals. With the increasing impact of global warming, native cattle breeds renowned for their heat tolerance are gaining prominence over crossbred varieties. However, there is still limited comprehension regarding the underlying systems biology responsible for this advantageous trait in native cattle. Malik *et al.,* (2021) conducted

a transcriptome analysis to compare Tharparkar (a native cattle breed known for heat tolerance) and Crossbred cattle under heat stress conditions. The key molecules and genes that exhibited interaction in a manner that could potentially counteract the detrimental effects of heat stress. These findings provided a valuable insight into the molecular mechanisms underlying the heat tolerance of Tharparkar cattle and highlighted potential targets for further research aimed at improving heat resilience in cattle populations.

2.8 Sheep skin transcriptomic studies

Shi *et al.*, (2022) identified the key genes and signaling pathways related to wool density in Hetian sheep by the transcriptome analysis of the skins with different wool densities. A total of 1,452 differentially expressed genes were found between high wool density and low wool density sheep. Many pathways related to hair growth such as MAPK and TGF-β/BMP signaling were represented by the DEGs. Many key genes such as *MAP3K7, MAP2K2, FST, TNF, PTPN11, KIT, INHBA,* and *BMPR1A,* were observed to be involved in growth and density of the wool.

A recent study by Bai *et al.*, (2021) investigated the expression of genes in the skin of Xinji fine- wool sheep at different stages of growth (3 and 12 months old) as well as Tan sheep (12- month- old) to correlate the transcript profiles with the growth of hair follicle. Between the two breeds and across development stages, the differentially expressed genes were related to gene ontology terms such as multicellular organism development, skin development, animal organ development, and system development. Many genes linked with hair follicle development such as *SRF, LAMA5, HOXC13, SMO, CDSN, DHCR24, OVOL1, KDF1,* and *NGFR* were identified as candidate genes. Members of homeobox and zf- C2H2 transcription factor families showed higher expression in fine- wool Xinji sheep in comparison to the Tan sheep. These genes are associated with the production of fine and dense wool.

Qi *et al.*, (2021) compared the skin transcriptome of sheep with high fibre diameter (Coarse wool) and low fibre diameter (Fine wool) using RNA-Seq technology. Many genes related to skin–hair cycle, namely *Wnt10b, KRTAP7-1, Wnt2b, β-catenin, KRT14,* and *FGF5* were observed in the skin expression profiles of both groups. The gene ontology analysis of the differentially expressed genes showed enrichment of KEGG pathways involving lipoprotein and lipid metabolism, suggesting that the fibre diameter is influenced by lipid metabolism. The RNA-Seq results were validated by quantitative real-time PCR (qPCR).

Integrated analysis of transcriptome and methylome in the skin tissue of Merino sheep across different embryonic stages and postnatal stages was carried out by Zhao *et al.,* (2021). Distinct patterns of expression of protein-coding genes and other types of RNAs (lncRNAs, circRNAs, and miRNAs) and their correlation with DNA methylation were evident across different hair follicle developmental stages. Gene co-expression and regulatory network analyses delineated key genes governing hair follicle morphogenesis. Many transcriptional factors (*eHOXC13, LEF1, RBPJ, VDR, KLF4, RARA,* and *STAT3*) associated with stage-specific involvement in hair follicle development were also predicted. Integration of the results of transcriptome and methylome analysis with genome-wide association studies (GWAS) in Merino sheep identified many genes related to wool-related traits in ovine.

Transcriptome profiling analysis revealed key genes associated with different coat color in sheep skin in a study by Yao *et al.,* (2019). Six differentially expressed genes (*DCT, TYR, TYRP1, PMEL, SLC45A2,* and *MLANA*) were shown to be involved in the regulation of coat color in the transcriptome sequencing investigation of the four groups of sheep with different coat color phenotypes (cyan grey, black, light brown, and white). These genes were linked to developmental pigmentation, melanosome trafficking, melanocyte differentiation, and the biosynthesis of melanin. The identified genes were considered as candidate genes for future research to understand their functional role in the control of sheep coat color.

The transcriptome data from the skin of sheep with coarse wool (Small Tail Han) and fine wool (Super Merino) was analyzed to identify wool and fiber characteristics (Zhang *et al.,* 2017). According to the study's findings, sheep with fine wool had greater expression levels of genes related to lipid metabolism, keratin-associated proteins, keratins, and markers for wool follicle stem cells. Many genes such as *KRT7927, KRTAP1-1, KRT3626, KRTAP6-1, KRTAP4-9L, KRTAP9-2, KRTAP6-2L,* and *KRTAP9-9L* showed higher expression in Super Merino sheep and these genes belong to the family of the keratin (*KRT*) and keratin associated proteins (*KRTAP*). On the other hand, greater expression of genes coding for actin cytoskeleton structural proteins was observed in the Small Tail Han sheep.

Sheep skin transcriptome from black and white coat color sheep was compared and a total of 2,235 genes differentially expressed genes were identified. Among the up-regulated genes in the black sheep, *TYR, MATP, DCT,* and *TYRP1* belong to the structural components of melanosome (Fan *et al.,* 2013). Yao et al., 2019 carried out skin transcriptome sequencing in sheep with different color phenotypes (Black, light brown, cyan grey, and white) and observed

six genes (*PMEL*, *SLC45A2*, *MLANA*, *DCT*, *TYR*, and *TYRP1*) with exhibit differential expression patterns. These genes were found to play crucial roles in the regulation of coat color, developmental pigmentation, melanosome transport, melanocyte differentiation, and the biosynthesis of melanin.

RNA sequencing was employed to investigate the skin structure, color variations, and gene expressions in two native Chinese sheep breeds, namely the Small-tail Han and Minxian black fur sheep. Several candidate genes that controlled the production of melanin, including *DDC*, *MC1R*, *COA2*, *FZD2*, *TYR*, *TYRP1* and *DCT* were shown to have a distinct expression pattern of Minxian black fur sheep. These genes may be required for melanin synthesis via the signaling pathways of melanogenesis and tyrosine metabolism (Shi *et al*., 2022).

CHAPTER–3

Materials & Methods

MATERIALS AND METHODS

3.1 Ethics statement

The Institutional Animal Ethics Committee (IAEC) of the ICAR-National Bureau of Animal Genetics Resources (NBAGR), Karnal, approved the tissue sampling of sheep in accordance with the appropriate rules and regulations (F.No. NBAGR/IAEC/2017, dated 21.01.2017).

3.2 Animal selection and sampling

In this study, skin samples were collected from two predominant Indian sheep breeds, Changthangi (N=4) and Muzzafarnagri (N=4), aged between 1 to 1.5years from their breeding tracts in Leh and Uttar Pradesh, respectively. Leh belongs to the Northern temperate agro-ecological zone with extremely cold and hypoxic environment. Muzzafarnagri sheep from Uttar Pradesh represents the North western arid and semi-arid zone of India (Sharma *et al.,* 2020). The Changthangi sheep included in this study had grey skin and black coat color, while, Muzaffarnagri sheep had white skin and coat (Figure 3.1). Sample collection for both the breeds was carried out by a trained veterinarian. About 200 to 500 mg of skin samples were immersed in 2.5 ml of RNA later, transported to laboratory within 24 hours and preserved at -80°C to ensure their integrity for future evaluation. These samples were used for generating RNA sequencing data for the skin of Changthangi and Muzzafarnagri sheep. To identify suitable reference genes for qPCR studies in different sheep tissues, ten tissues, including muscle, skin, kidney, liver, intestine, rumen, lung, testis, heart, and spleen were collected from 5 sheep by a trained veterinarian. A small amount of the tissue (3-5 g) was promptly immersed in 2 ml of RNA*later* solution (Thermo Fisher Scientific, United States) and transported to the laboratory within 24 hours. The RNA*later* solution was subsequently drained out and the samples were kept at -80°C for future use.

3.3 Tissue homogenization and RNA isolation

Approximately 80 - 100 mg of tissue was cut into thin slices and immersed in Trizol reagent at 4 °C (Sigma-Aldrich Co., United States). Using tissue rupture, tissues were homogenised and stored at -80^0 C. Total RNA was isolated from tissue samples using Trizol reagent (Sigma-Aldrich) according to the manufacturer's protocol. Subsequently, the extracted RNA was purified using RNase-free DNase enzyme treatment and RNeasy Mini Kit columns (Qiagen, Germany).

In the similar manner, RNA isolation was also carried out for 10 tissues each from 5 sheep.

3.4 Evaluation of RNA quality and quantity

To assess the purity and concentration of the RNA, NanoDrop ND-1000 spectrophotometer and Experion Bioanalyzer from Thermo Scientific (USA) were utilized. Furthermore, the integrity of each RNA sample was verified by visualizing the presence of distinct 18S and 28S ribosomal bands on a 1% agarose gel. RNA samples having RIN (RNA Integrity Number) values greater than or equal to 8.0 and OD 260/280 ratio greater than 2.0 were considered suitable for futher experimentation.

3.5 cDNA synthesis

First strand cDNA was synthesized using the RevertAid First Strand cDNA Synthesis Kit (Termo Fisher Scientifc, CA, USA). The cDNA was synthesized using purified RNA (500 ng), DNase/RNase-free water, and oligo(dT) primers. After incubating at 65°C for 5 minutes, the mixture was chilled for 3 minutes. After that, 8 µl of master mix consisting of 4 µl of 5X reaction buffer, 1 µl of M-MuLV reverse transcriptase enzyme, 2 µl of 10 mM dNTP mix, and 1µl of Ribolock RNase Inhibitor enzyme was added. The mixture was then incubated for 60 minutes at 42^0 C, followed by 10 minutes at 70^0 C. Before proceeding for qPCR, each cDNA sample was diluted 1:9 (v:v) in DNase/RNase-free water.

3.6 RNA sequencing and data analysis

The RNA sequencing libraries were generated using Illumina-compatible NEBNext Ultra Directional RNA Library Prep Kit (New England Biolabs, MA, USA). Following library preparation, the NovaSeq6000 Sequencing System (Illumina) was employed to generate 150 bp paired-end reads. The quality of raw sequences obtained from RNA sequencing run was checked using the FastQC software (v 0.11.5) (Andrews, 2010). The trimming and filtering of the raw reads was performed using FastXToolKit. CLC Genomics Workbench 6.5.1 (CLC Bio, Aarhus, Denmark) was used for the data analysis. The high-quality reads or clean sequences were aligned with the available reference assembly of Oar v3.1 genomic.fna by means of the 'Map Reads to Reference Tool' embedded in the CLC Genomics Workbench. The aligned sequences were normalized as reads per kilobase per million transcripts (RPKM) and transcripts or reads having RPKM values less than 0.01 were not included in further analysis. The CLC transcriptomics analysis tool was used for differential gene expression analysis between Changthangi and

Muzzafarnagri sheep skin. Transcripts or genes that showed differential expression with a log2 fold change of more than 1.5 and a 'p' adjusted (padj) value less than 0.05 were used for further analysis.

3.7 Small RNA-seq data analysis

The COMPSRA (Comprehensive Platform for Small RNA-seq data Analysis) (https://github.com/cougarlj/COMPSRA) was utilized to evaluate the analysis of quality and reliability of RNA-seq data, including assessing alignment statistics, identifying potential issues such as adapter contamination or low-quality reads, and comparing multiple alignment results. To eliminate poor quality reads, a quality control filter was applied based on the specified parameters set in the command line (-rh 20 –rt 20 –rr 20). This trimming step plays a crucial role in improving the accuracy and reliability of the downstream analysis, as it ensures that only high-quality and biologically meaningful sequences are considered. COMPSRA utilizes STAR as its default RNA sequence aligner to align the clean reads to the reference genome (Dobin *et al.,* 2013). High-quality reads from the QC modules were initially aligned to the human genome versions hg19/hg38. Subsequently, the aligned reads were quantified and annotated within the Annotation Module. COMPSRA uses different small RNA databases such as circBase for circular RNA (Glazar *et al.,* 2014); miRBase for miRNA (Kozomara and Griffiths-Jones, 2011); piRNABank, piRBase and piRNACluster (Sai Lakshmi and Agrawal, 2008; Zhang *et al.,* 2014; Rosenkranz, 2016) for piRNA; gtRNAdb for tRNA (Chan and Lowe, 2016); GENCODE for snRNA and snoRNA (Harrow *et al.,* 2012) to extract different kinds of non-coding RNA.

3.8 Prediction of target genes for potential miRNAs

It was done using miRNet which is a web-based tool (http://www.mirnet.ca/) that provides comprehensive information about potential miRNA target genes, by integrating data from eleven different miRNA databases (EpimiR HMDD, miRanda, miR2Disease, miRecords, miRTarBase, PhenomiR, PharmacomiR, SM2miR, starBase, and TarBase).

3.9 Functional annotation of the differentially expressed genes and pathway analysis

The differentially expressed transcripts/genes were analyzed to ascertain their biological functions and to enable discovery of enriched pathways using DAVID software (Huang *et al.,* 2009). The functions of the differentially expressed genes were interpreted in terms of their involvement in biological, molecular and cellular activities. Kyoto Encyclopedia of Genes and

Genomes (KEGG) (Kanehisa *et al.,* 2016) and Reactome (https://reactome.org) (Jassal *et al.,* 2020) were used to discover the enriched biological pathways in this study.

3.10 Network analysis

Gene-protein co-expression networks which give an overview of the connectivity of genes and proteins for deducing significant biological inferences were also constructed. Cytoscape was used for construction of the co-expression network together with cytoHubba app to discover the key genes with highest degree of connectivity (Shannon *et al.,* 2003; Chin *et al.,* 2014). The protein–protein interactions (PPI) of the DEGs were studied using STRING with default parameters (Szklarczyk *et al.,* 2017). Network construction was done for credible interactions with high confidence PPI score (0.7).

3.11 Candidate reference gene selection and primer designing for qPCR

To validate the differentially expressed genes identified by RNA Sequencing in Changthangi and Muzzafarnagri sheep, a need was felt to identify suitable reference genes for qPCR analysis of the target genes. Therefore, in this study, 18 commonly used reference genes from a variety of functional categories were selected, namely, *ACTB*, *BACH1*, *B2M*, *GAPDH*, *HMBS*, *HPRT1*, *PGK1*, *PPIA*, *PPIB*, *RPLP0*, *RPL19*, *RPS9*, *RPS15*, *RPS28*, *SDHA*, *TBP*, *UXT*, and *YWHAZ*. These candidate reference genes were chosen based on previous studies involving diverse tissues from various species (Nygard *et al.,* 2007; Vorachek *et al.,* 2013; Zhang *et al.,* 2013; Jiang *et al.,* 2015; Puech *et al.,* 2015; Mahakapuge *et al.,* 2016; Schulze *et al.,* 2017; Sahu *et al.,* 2018; Wang *et al.,* 2020; Sonowal *et al.,* 2022). Primers were either taken from published studies or designed using the Primer3 software (Untergasser *et al.,* 2012). BLAST analysis was carried out to confirm the specificity of primers for each gene. To prevent DNA amplification, exon-exon junction spanning primers were designed. The target specificity of the primers was evaluated using end-point PCR and the size of their amplicons was determined using electrophoresis on a 1.5% agarose gel. Table 3.1 summarizes the information about the potential reference genes, their function, accession number, primer sequence, annealing temperature, and amplicon length for each gene.

3.12 Real time quantitative PCR (qPCR)

qPCR was used to evaluate the expression of each candidate reference gene using the PowerUp™ SYBR™ Green Master Mix (Thermo Fisher Scientific, Lithuania). The PowerUp™ SYBR™ Green Master Mix, MicroAmp® Optical Adhesive Films (Applied Biosystems), and

MicroAmp® Optical 96-Well Reaction Plates (Applied Biosystems) were used in the reactions, which were carried out in the QuantStudio™ 5 Real-Time PCR System (Applied Biosystems). The 10 µL reaction mixture consisted of 2 µL cDNA, 8 µL master mix comprising of 5 µL of 2X PowerUp™ SYBR™ Green (Thermo Fisher Scientific, Lithuania), 2.4 µL of DNase/RNase free water, and 0.3 µL each of forward and reverse primers. The cycling protocol consisted of 40 cycles of denaturation at 95°C for 15s, annealing at specific temperature for 30s, and extension at 72°C for 30s. Non-template control (Negative control) of each gene was included in the assay to detect unwanted nucleic acids. The specificity of each primer was determined by the melting curve analysis. The standard curves of each primer pair were constructed using ten-fold serial dilutions of cDNA samples. Quantification cycle (Ct) values and logarithms of the cDNA concentrations were used to determine the amplification efficiency, correlation coefficients, and slope of each primer pair.

3.13 Candidate reference gene expression stability analysis

The average Ct (Quantification cycle) values of all 50 tissues (10 tissues each from 5 sheep) were analysed for the 18 potential reference genes using the web-based tool called RefFinder. The comparison and ranking of the experimental candidate genes were achieved using NormFinder, GeNorm, BestKeeper, and comparative ΔCt algorithms. The reference genes were ranked in accordance with the stability values (SV) by the NormFinder algorithm based on their relative quantities. Lower the SV, higher is the expression stability of the gene. The geNorm algorithm investigates the reference gene expression stability (M) individually and then estimates its pair-wise variance with other genes. Here again, the gene with the lowest M value is regarded to be the most appropriate reference gene. The threshold M value for any gene to be considered as a reference gene is <1.5.The stability of the reference genes was also assessed using the BestKeeper method by examining the coefficient of variance (CV), p-values (p), standard deviation (SD), and correlation coefficient (r) between the genes for each selected gene. For a gene to be considered a reference gene, its SD and CV should be low and the correlation coefficient should be high. SD value >1 is indicative of high difference in expression of the gene among samples, thus, higher instability. The ΔCt method calculates the SD of the genes based on their relative expression levels and then determines the most stable candidate

Table 3.1: The selected reference genes, primer sequences, annealing temperature, and amplicon size.

Gene	Primer sequence (5'→3')	Annealing temperature (^{0}C)	Amplicon size (bp)	Reference
ACTB	F: CTCTTCCAGCCTTCCTTCCT R: TAGAGGTCCTTGCGGATGTC	60	101	Kaur *et al.*, 2023
BACH1	F: GCTGCAAAACGAAAAGGAGAGCT R: TCAGGTTCTGCTTGGTCTCGC	60	73	Kaur *et al.*, 2023
B2M	F: GGCCTGCTGTCGCTGTCT R: TTCTGGCGGGTGTCTTGAGT	60	78	Kaur *et al.*, 2023
GAPDH	F: GGTCGGAGTGAACGGATTTG R: TGGCAACGATGTCCACTTTG	58	83	Kaur *et al.*, 2023
HMBS	F: CTCGCATACAGACGGACAGT R: CAGGTCCTTCAGCGAGTGA	58	209	This study
HPRT1	F: TTCCTATGACTGTGGATTTTATCAGACT R: CAATTACTTTTATGTCGCCTGTTGAC	58	74	French *et al.*, 2006
PGK1	F: GCCTTCCGAGCTTCACTTTC R: GGCCTTGGCAAAGTAGTTCA	58	156	This study
PPIA	F: CATACAGGTCCTGGCATCTTGTCC R: CACGTGCTTGCCATCCAACC	62	108	Kaur *et al.*, 2023
PPIB	F: ACACCAACGGCTCCCAGT R: AGGCTTGTCCCGACCATC	60	140	Kaur *et al.*, 2023
RPLP0	F: CAACCCTGAAGTGCTTGACAT R: AGGCAGATGGATCAGCCA	58	227	Kaur *et al.*, 2023

RPL19	F: AGCCTGTGACTGTCCATTCC R: ACGTTACCTTCTCGGGCATT	58	126	This study
RPS9	F: CCTCGACCAAGAGCTGAAG R: CCTCCAGACCTCACGTTTGT	58	54	This study
RPS15	F: CGTAGACCTCGACCAGCTAC R: TTGTAGACGCCAACCATGCT	58	237	Kaur *et al.*, 2023
RPS28	F: TGCGCGTAGAATTCATGGAC R: GACAACCCAGACATCCAGCA	58	155	Kaur *et al.*, 2023
SDHA	F: GAGTTCGTGCAGTTCCACCC R: CTCTCACCCTGGCTGTTGAT	58	104	This study
TBP	F: GCACCACCGCACTGATATTC R: ACTCCGGCTCATAACTACTGA	58	237	This study
UXT	F: CGCTACGAGGCTTTCATCTC R: CCACTGTGTCAACGAAGAAGT	58	193	Kaur *et al.*, 2023
YWHAZ	F: GTCATCTTGGAGGGTCGTCT R: CTTGTGAAGCGTTGGGGATC	58	167	Kaur *et al.*, 2023

3.14 Validation of differentially expressed genes in Changthangi and Muzzafarnagri sheep

Ten genes *viz.*, *DCT, TYR, TYRP1, OCA2, SLC45A2, PMEL, KAP6, KAP7, KAP13,* and *KRT19* were randomly selected to validate their differential expression in Changthangi and Muzaffarnagri sheep (Table 3.3). Primers for these genes were designed using Primer3 software version 4.1.0 (Untergasser *et al.,* 2012). The first strand cDNA was synthesized from 500 ng of RNA using the RevertAid First Strand cDNA Synthesis Kit (Termo Fisher Scientifc, CA, USA) following manufacturer's instruction. The gene expression was measured by quantitative PCR using PowerUp™ SYBR™ Green Master Mix (Thermo Fisher Scientific, Lithuania). The qPCR was performed in a 10 µl reaction volume comprising of 2 µl diluted cDNA, 5 µl SYBR green, 0.2 µl each of 10 µM forward and reverse primers and 2.6 µl DNase/RNase free water in a QuantStudio™ 5 Real-Time PCR System (Applied Biosystems, Thermo Fisher) in the standard mode. Each qPCR reaction was performed in duplicate in order to assess the intra-assay variation and to check the quality. The amplification for qPCR reaction was carried out with the following conditions: 2 min at 50°C, 5 min at 95°C, 40 cycles of 15s at 95°C (denaturation), 30 s (annealing) at respective annealing temperature of primers and 30 s at 72°C (extension). For the qPCR, the mean cycle threshold (Ct) values of the genes were normalized to geometric mean of the reference genes ACTB and PPIB and the $2^{-\Delta\Delta CT}$ method was used for data analysis (Livak and Schmittgen, 2001

Table 3.2: The randomly selected genes, their function, accession number, primer sequences, annealing temperature, and amplicon size for validation of RNA sequencing data.

Gene	Sequences (5'→ 3')	Annealing temperature	Length (bp)
TYR	F: GCACAACCGGGAATCCTACA	58	169
	R: CCAGATTCGACGTGCTTGTT		
TYRP1	F: ACAGCTTCCGAAACACAGTG	58	204
	R: CAGCATTGTATCGCCTCAGC		
DCT	F: ACTTCGCCGGCTATAACTGT	58	237
	R: CAGTTGGCGATCTGTGGTTG		
OCA2	F: CTTGGGAGGACAGCTCAACT	58	161
	R: TTTTAGCCACCGTACACAGC		
PMEL	F: CCCTGCCGGATGATACCTTA	58	162
	R: GCTCAAATGCGTCTCCTTCG		
SLC45A2	F: GAGTTCTGCTATGGGGTGGA	58	189
	R: CGTGAGGATGTAGGGTCTCC		
KRTAP6	F: GAGAACAACCTCAACAACCAAC	58	257
	R: CTTCCGTGGCATCCTCAATAGT		
KRTAP7	F: GGACCATGGCTCCAGATGACTAC	58	343
	R: ACAAGCAAAACCCCTTCCTACTCA		
KRTAP13	F: CTCCACGTTCTTCAGTCCCT	58	118
	R: GAAACCACCGGATCCAAAGC		
KRT19	F: CGCGACTACAGCCACTACTT	58	154
	R: CAAAGCTTGCTCCGTCTCAA		

CHAPTER–4

Results

RESULTS

4.1 RNA sequencing data

RNA sequencing generated a wide range of raw reads in different samples of Changthangi and Muzzafarnagri breeds, ranging from 36.8 million to 1.04 billion, and processed reads, ranging from 36.5 million to 1.04 billion. The alignment of these reads with the Oar v3.1 genomic.fna reference (GCA 000298735.1) resulted in 98.27% to 99.37% of the reads successfully mapping, indicating that the quality of the RNA-seq data is suitable for further analysis. The absence of sequencing biases in the dataset was confirmed by observing a consistent proportion of mapped reads across the samples (Table 4.1).

Table 4.1:Read statistics of 8 libraries from Changthangi and Muzzafarnagri breeds

Sample	Raw reads	Processed reads	% of high-quality data	% aligned reads
Changthangi1	3,68,27,512	3,65,94,148	99.36	99.37
Changthangi2	6,82,34,368	6,79,62,154	99.60	99.26
Changthangi3	5,61,09,072	5,58,66,399	99.56	99.3
Changthangi4	5,97,17,590	5,94,36,296	99.52	99.27
Muzzafarnagri1	8,05,88,926	8,00,60,940	99.34	98.27
Muzzafarnagri2	10,47,20,948	10,40,86,745	99.39	98.92
Muzzafarnagri3	6,14,25,062	6,10,29,745	99.35	99.02
Muzzafarnagri4	6,07,12,740	6,05,37,484	99.71	99.01

4.2 Gene expression profile

The top 25 most abundantly expressed genes in both groups were *ACTG1, CALD1, COL1A2, COL3A1, DCN, DDX17, EIF4G2, FSTL1, HNRNPA2B1, KRT5, PABPC1, PPP1CB, PTMA, RPL5, RPL7, RPL9, RPS24, RPS4X, VIM,* and *ZC3H10*. A Gene Ontology enrichment analysis was conducted on these top 25 highly expressed genes and these were found to be associated with translation, collagen fibril organization, ribosome, ribonucleoprotein complex, RNA binding etc. The distribution of the top 25 differentially

expressed genes among biological processes, cellular components, and molecular functioning is illustrated in Figure 4.1. A total of eight annotation clusters with enrichment scores >1.5 and $p_{adj} \leq 0.05$ were observed, including ribonucleoproteins, mRNA processing, proteoglycans in cancer, acetylation, extracellular matrix, helicase activity, and acetylation.

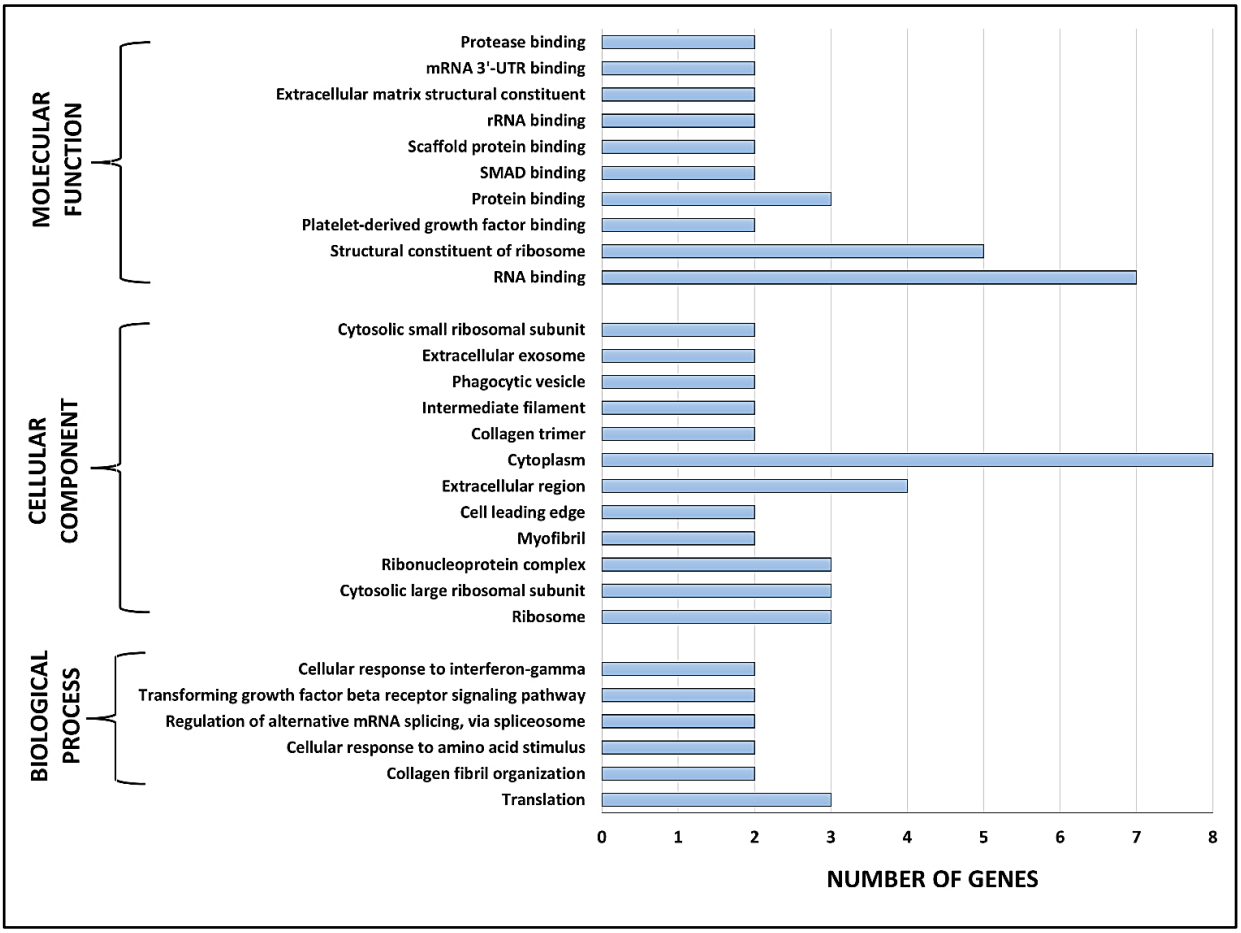


Figure 4.1: Functional classification of the profusely expressed genes in Changthangi and Muzzafarnagri sheep

Among the differentially expressed genes, a total of 149 genes showed significantly higher expression and 2139 genes exhibited significantly lower expression levels (padj ≤ 0.05 & Log2 fold change ≥ 1.5) in Changthangi sheep as compared to Muzzafarnagri sheep.

4.3 Functional analysis of the up-regulated genes in Changthangi sheep

According to GO annotations, DEGs were involved in various biological processes (BP), such as transcription regulation from the promoter RNA polymerase II, developmental pigmentation, melanin biosynthetic process, and response to blue light. Meanwhile, melanosome, melanosome membrane, and endoplasmic reticulum membrane were the main GO terms for DEGs in the cellular component (CC) category. The primary molecular functions (MF) were RNA polymerase II core promoter proximal region sequence-specific DNA binding, RNA polymerase II transcription factor activity, monophenol monooxygenase activity, oxidoreductase activity (Figure 4.2). In terms of fold enrichment, most significant GO terms were response to blue light, melanosome membrane, and monophenol monooxygenase activity for BP, CC and MF, respectively. Genes associated with melanin biosynthesis were *PMEL, TYRP1*, and *TYR*, with developmental pigmentation were *SLC45A2, DCT*, and with response to blue light were *DCT* and *TYR*. The melanosome membrane was represented by OCA2, DCT, TYRP1, and TYR, melanosome by PMEL, DCT, MLANA, TYRP1, and TYR and monophenol monooxygenase activity by TYRP1 and TYR.

4.4 Gene ontology analysis of the down-regulated genes in Changthangi sheep

The significant gene ontology terms for down-regulated genes included 181 terms for biological processes, 98 terms for cellular components, and 104 terms for molecular functions ($P_{adj} \leq 0.05$). Signal transduction, cell differentiation, and intracellular signal transduction were enriched in the BP category, cytoplasm, cytosol, and nucleoplasm in the CC category and ATP binding, metal ion binding, and identical protein binding in the MF category (Figure 4.3). Some of the genes involved in cell differentiation include *THRB, THRA, HNF4G, CXCL17, SPATA20, TMEM100, FGF6, BOLL,* and *HNF4A.* The intracellular signal transduction-related genes such as *DGKD, MAST2, ARAF, HSPB1, PREX1, AKT2, RPS6KA2, JAK3, TNS3,TNS2* etc. are involved in transmitting signals within cells to regulate various cellular processes. Signal transduction-associated genes include *GPI, TOR2A, HHIP, LRRC39, PTPRM, ARRB1, LSP1, LRRC2, ARHGAP4, ARHGAP44,* and *CCND3* etc. These genes play a role in the transmission of signals between cells or within cells.

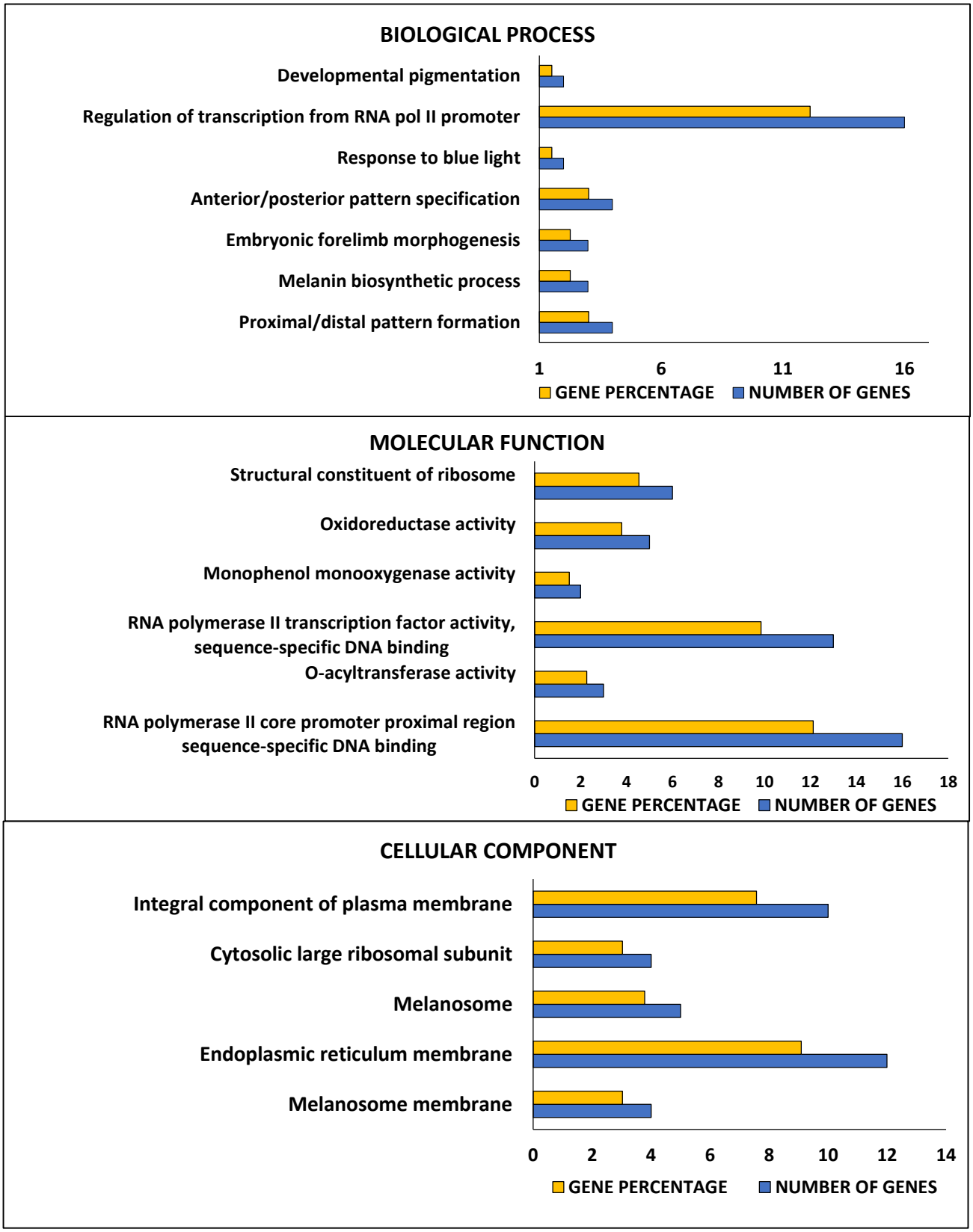


Figure 4.2: Gene ontology terms for up-regulated genes in Changthangi sheep

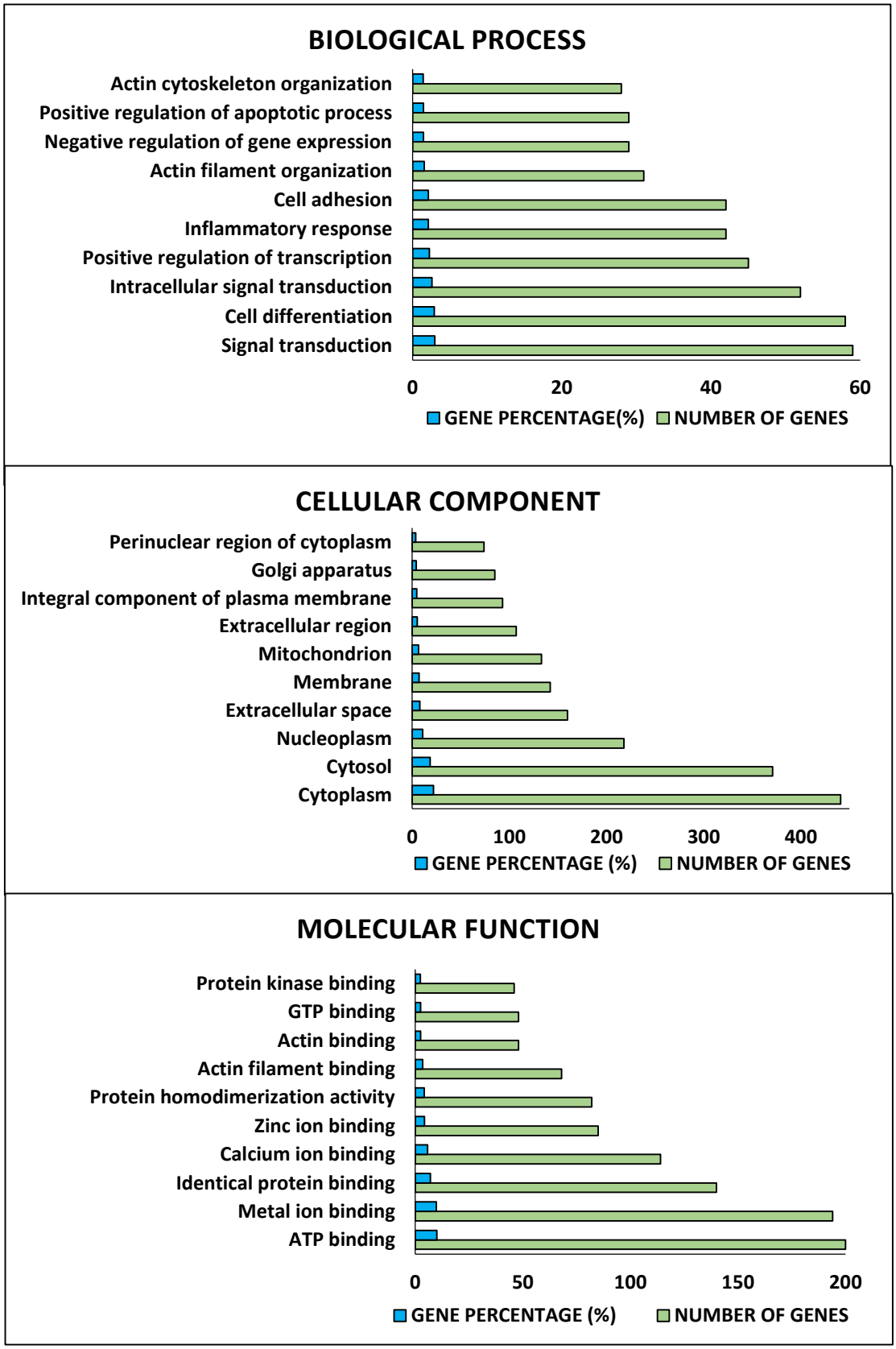


Figure 4.3: Gene ontology terms for down regulated genes in Changthangi sheep

4.5 Pathway analysis for the DEGs

On the basis of KEGG pathway enrichment analysis, up-regulated genes in Changthangi sheep were associated with a number of pathways including tyrosine metabolism, metabolic pathways, and retinol metabolism. Metabolic pathways, pathways of neurodegeneration, pathways in cancer, and MAPK signalling pathways were enriched for the down-regulated genes in Changthangi breed (Figure 4.4). Genes related to tyrosine metabolism included *DCT, TYRP1,* and *TYR* and genes associated with retinol metabolism were *LRAT*, *CYP3A24*, and *RPE65*. Genes representing metabolic pathways included *ARG1*, *GDA, TYRP1, MBOAT2, CYP3A24, ELOVL7, TYR, HACD2, INMT, RPE65, GGCT, CHIT1, COX7A2L, GPAM, DCT, SC5D, LRAT, ENPP3, ATP6V0D2,* and *CA13*. For the down-regulated dataset, *PTPRR, FLT1, ARAF, HSPB1, ARRB1, ECSIT, CRKL, FGF6, AKT2, RPS6KA2,* and *AKT1* etc. are associated with the MAPK signalling pathway which plays a crucial role in the transmitting signals within the cell and is involved in regulating cellular responses through the MAPK signalling cascade. Genes such as *UXS1, ACSM5, ENO1, ENO3, GBA2, FLAD1, MLYCD, GSTK1, ARSA, MECR, SDS, OXSM,* and *DHDH* etc. are connected to metabolic pathways which contribute to various metabolic processes and are involved in the synthesis and breakdown of molecules within cells.

Since many of the up-regulated genes in Changthangi sheep were related to pigmentation and melanin biosynthetic process, the melanogenesis KEGG pathway and the melanin biosynthesis Reactome pathway are in Figure 4.5 and Figure 4.6, respectively. Melanogenesis refers to the synthesis and distribution of melanin, which is facilitated by melanocytes located within the basal cells of the epidermis. Following melanin formation within melanocyte melanosomes, the pigments are subsequently stored in both the basal layer of epidermal cells and dermal macrophages, which transform into melanophores (Maranduca et al., 2019). Melanocytes perform vital functions in mammals, such as controlling inherent pigmentation in the skin, hair, and eyes, contributing to embryonic development, and providing protection against ionizing radiation (Hearing *et al.,* 2000). According to Reactome pathway analysis, *TYR*, *OCA2*, and *SLC45A2* genes are involved in the oxidation of tyrosine to form dopaquinone and leucodopachrome, which is important for the synthesis of pheomelanin. In the presence of *DCT*, dopachrome is converted into *DHICA* and *DHI*. Furthermore, in the presence of *TYRP1, DHICA* can undergo oxidation to form *IQCA* and polymerized *DHI* and *DHICA* are involved in the production of eumelanin (d'Ischia *et al.,* 2013).

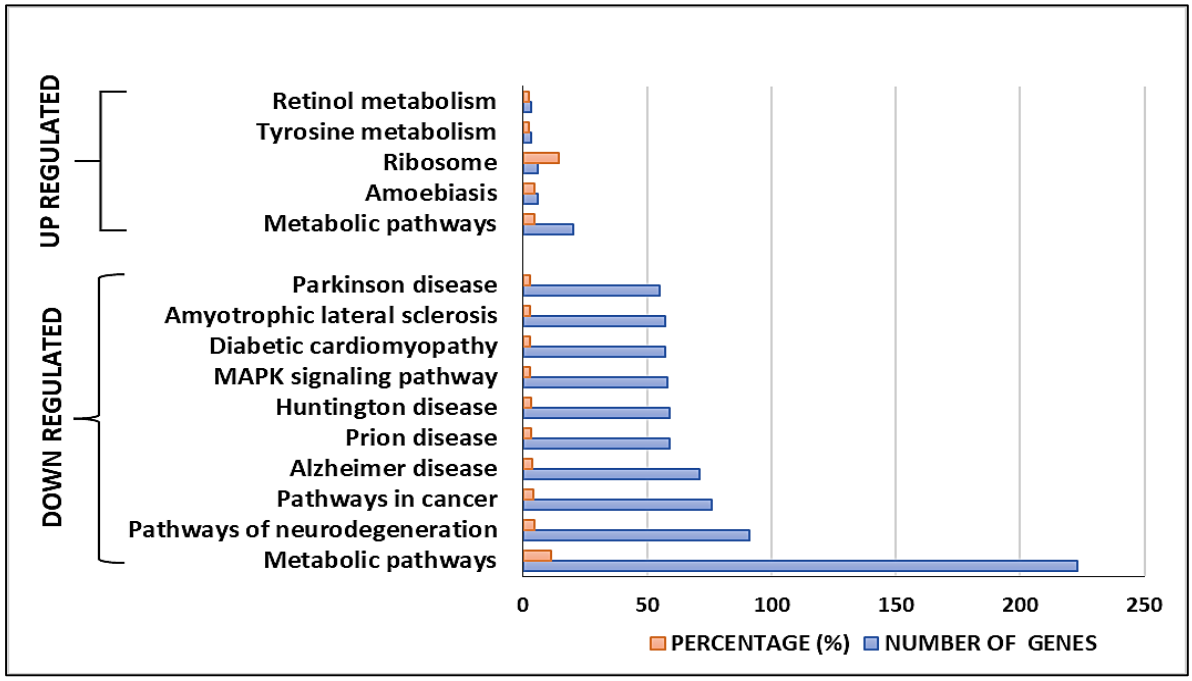


Figure 4.4: KEGG pathway enrichment analysis for both up-regulated and down regulated DEGs in Changthangi sheep

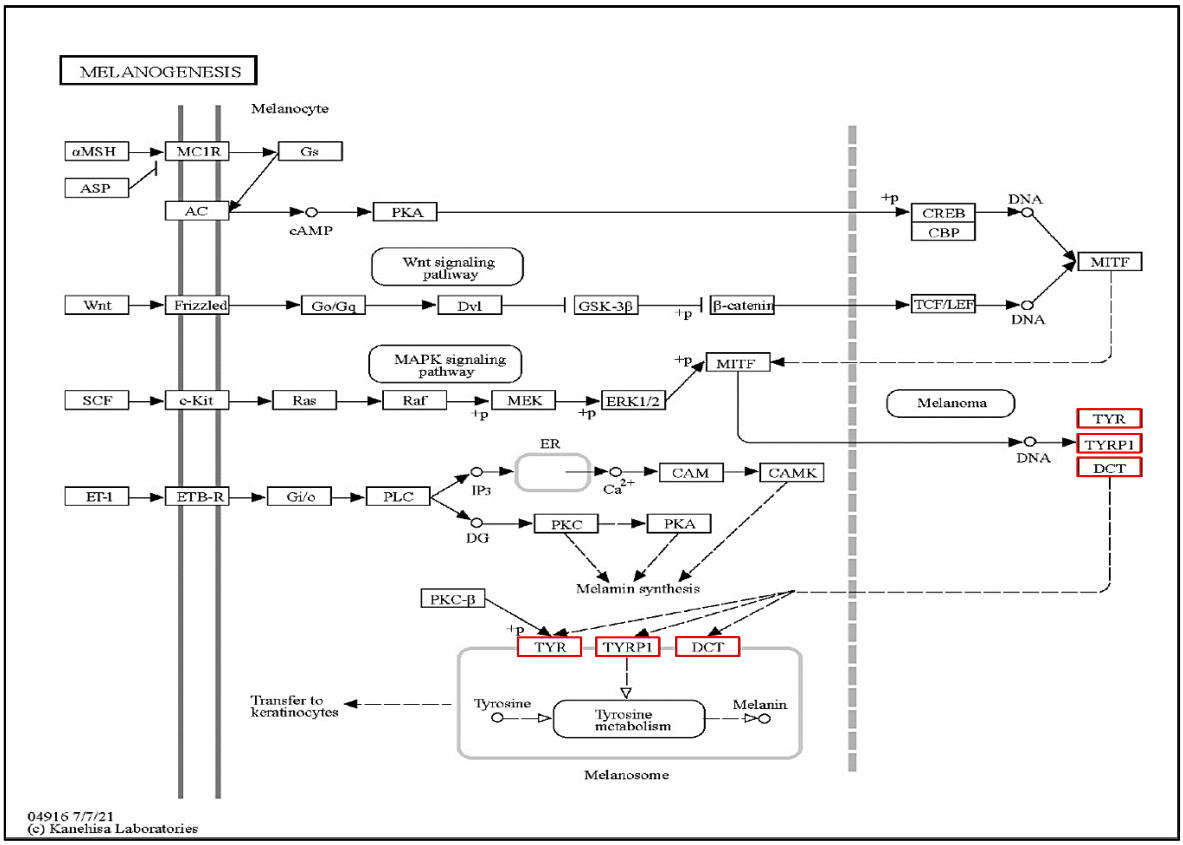


Figure 4.5: Role of up-regulated genes in Changthangi sheep in the melanogenesis pathway. Genes with red boundary showed higher expression in the black coat Changthangi sheep.

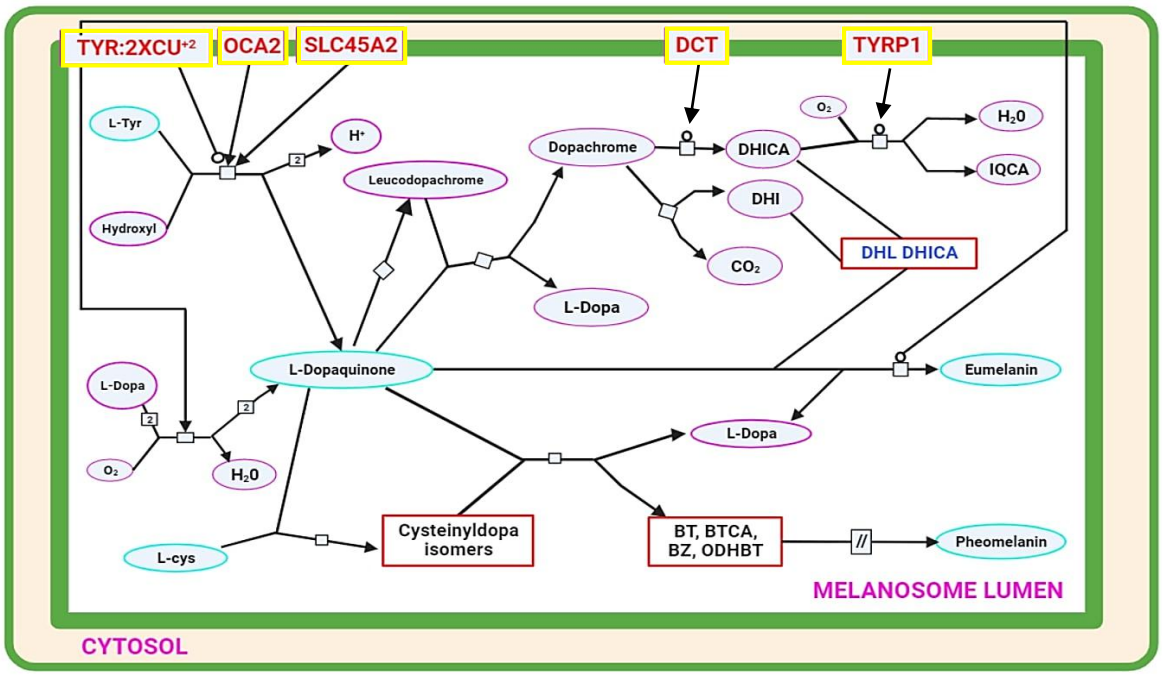


Figure 4.6: Melanin biosynthesis pathway revealing enrichment of up-regulated genes in Changthangi sheep skin (Adapted from Reactome pathway database)

4.6 Expression profiles of candidate genes for pigmentation and fiber quality

We observed significantly higher expression of 7 pigmentation-related genes, namely *TYR*, *TYRP1*, *DCT*, *SLC45A2*, *PMEL*, *MLANA*, and *OCA2* and 4 fiber quality genes, including *KRTAP6*, *KRTAP7*, *KRTAP13*, and *KRTAP21* in Changthangi sheep in comparison to the Muzzafarnagri breed. The difference in the expression of these genes has been depicted in Table 4.2.

Table 4.2: Expression levels of DEGs involved in Changthangi sheep adaptation to skin pigment

Gene	Gene name	Log_2 fold change	P value
TYRP1	Tyrosinase-related protein 1	11.38	0.00
MLANA	Melanoma antigen recognized by T-cells 1	6.96	9.21485E-15
TYR	Tyrosinase	6.89	9.79661E-13
SLC45A2	Solute carrier family 45-member 2	6.26	2.95969E-07
PMEL	Pre-melanosome protein	5.94	3.4723E-10
DCT	Dopachrome tautomerase	4.50	8.40902E-06
OCA2	OCA2 melanosomal transmembrane protein	2.93	7.68589E-05
KRTAP6	Keratin associated protein 6	8.20	0.00
KRTAP7	Keratin associated protein 7	3.90	0.000228667
KRTAP13	Keratin associated protein 13	4.52	0.001052076
KRTAP21	Keratin associated protein 21	4.41	0.000477418

4.7 Network construction

The highly expressed DEGs were subjected to CPDP pathway analysis, to observe interaction between the genes. A co-expression network was constructed for the up-regulated genes in Changthangi sheep which showed the top nodes with maximum interactions to be the genes for various ribosomal proteins and HOX genes (Figure 4.7). Genes such as *RPS25, RPL30, RPS15a, RPL10, RPS3A, RPL23A, RPL26,* and *EIF2A* are associated with the initiation and regulation of protein synthesis, a fundamental process in which cellular machinery translates mRNA into functional proteins. Interaction of a keratin associated protein; KRTAP6-1 was also evident in the co-expression network.

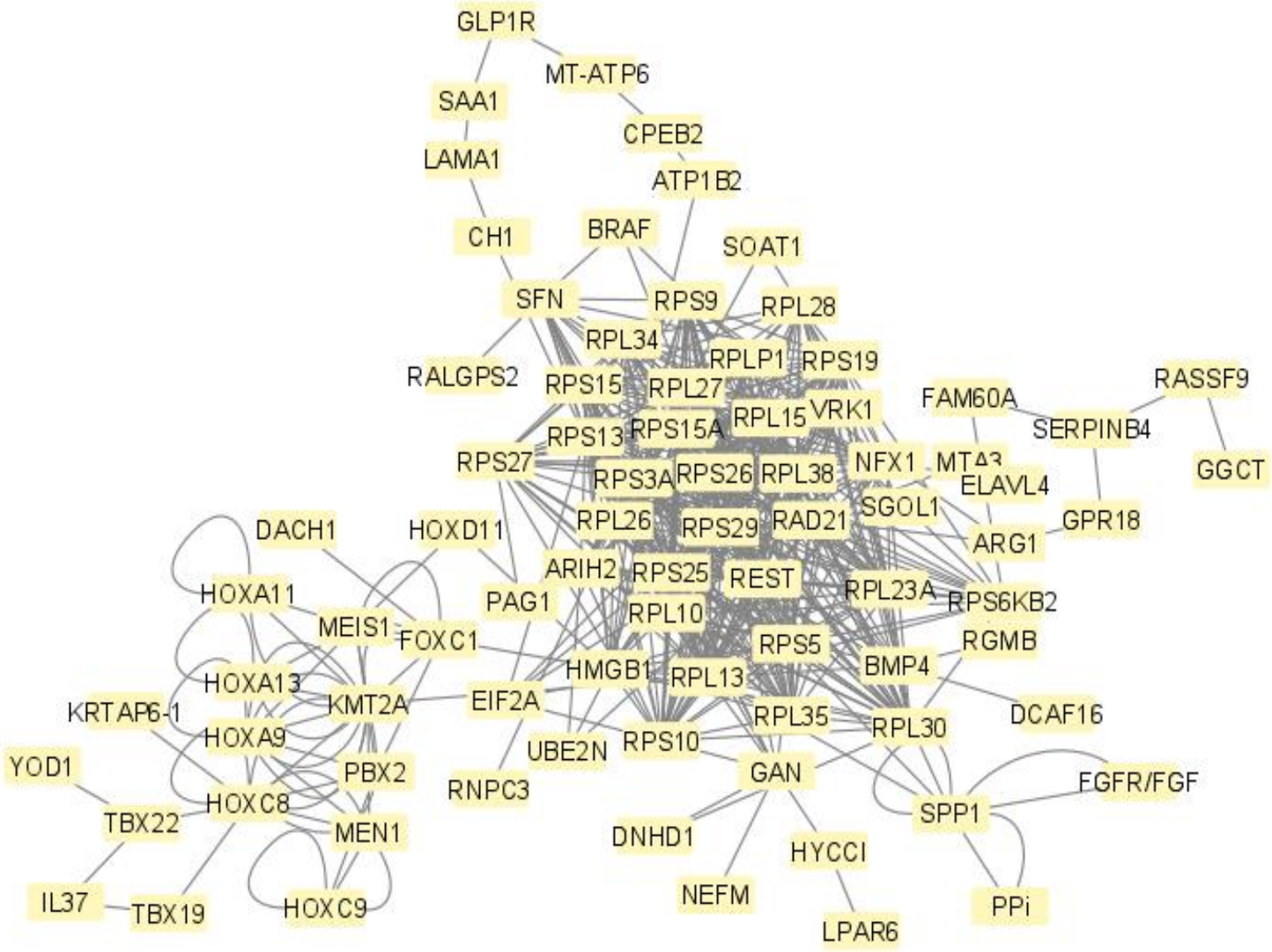


Figure4.7: Highly connected up-regulated genes in Changthangi sheep

The protein-protein interactions of the up-regulated dataset in the Changthangi sheep were also visualized using the STRING database which again revealed conspicuous interactions between ribosomal proteins and proteins associated with pigmentation and melanosome (*PMEL*, *SLC24A5*, *MLANA*, *OCA2*, *TYR*, *SLC45A2*, *TYRP1*, and *DCT*) (Figure 4.8). Retinol metabolism was found to be influenced by genes *LRAT* and *RPE65*, indicating their involvement in the processing and utilization of retinol.

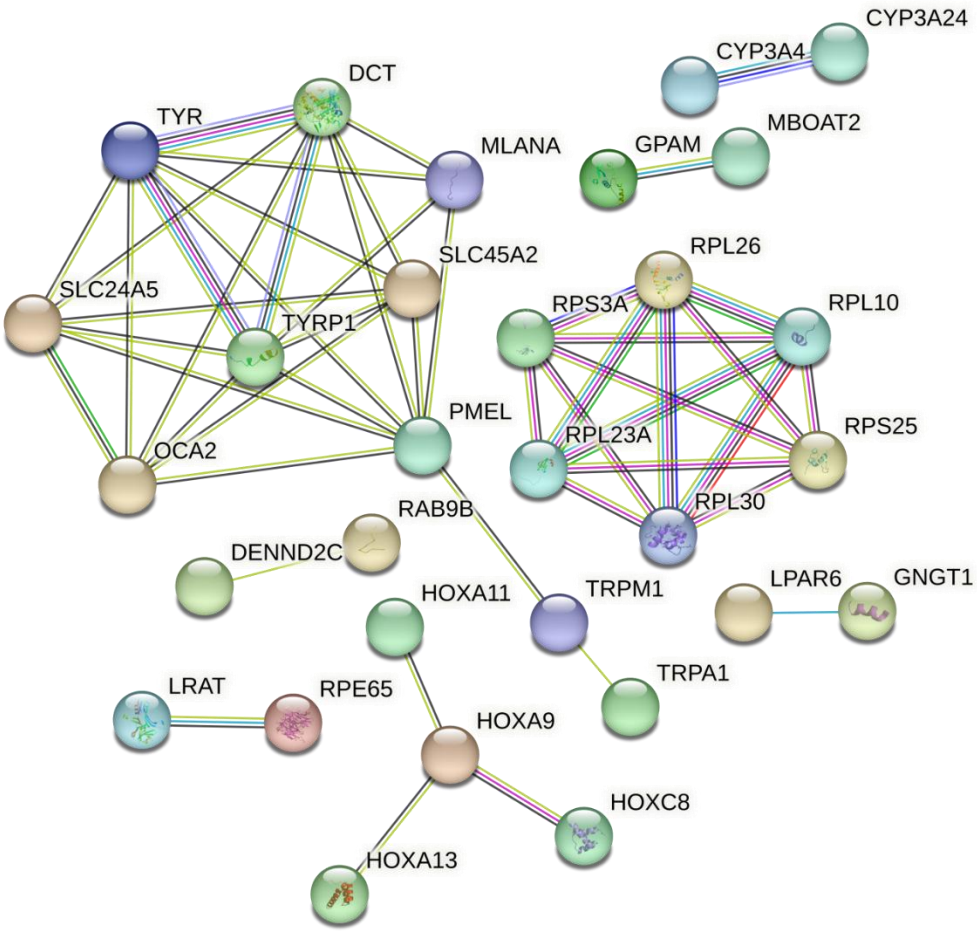


Figure 4.8: Protein-protein interactions of 149 up-regulated genes in Changthangi sheep determined by STRING (https://string-db.org)

Among the 2139 DEGs, a subset of the top 149 genes with highest Log2 fold change was chosen for network construction in order to identify highly connected genes from the down-regulated dataset in Changthangi sheep. The down-regulated genes were found to be primarily associated with pathways involved in cytoskeleton organization, cell development, regulation of transmembrane transport, and muscle growth and contraction. The interaction network divulged that the genes with maximum degree of interactions were *ACTN1*, *ACTN2*, *ACTN3*, *ACTN4*, *TPM1*, *TTN*, *AGRN*, *DYSF*, *MYBPC1*, and *MYBPC2* (Figure 4.9).

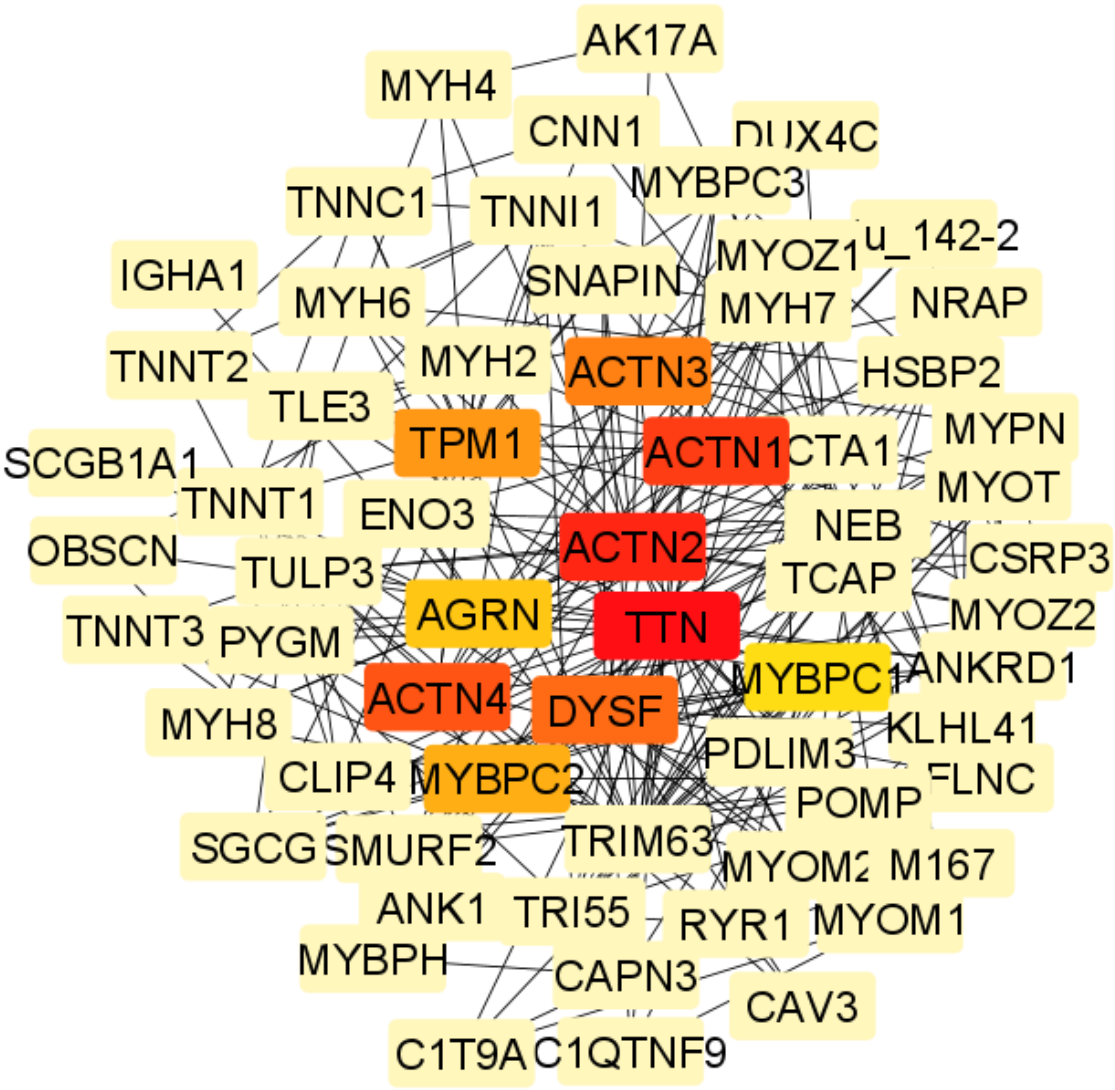


Figure 4.9: Interaction network for the genes down-regulated in Changthangi sheep

4.8 Candidate ncRNAs identification

The most abundant ncRNA molecules were identified in the skin of Changthangi and Muzzafarnagri sheep by using COMPSRA pipeline. In Changthangi, 34,440 circRNAs, 22 miRNAs, 200 piRNAs, 59 snRNAs, 19 snoRNAs, and 27 tRNAs were identified, whereas in Muzzafarnagri, 41,413 circRNAs, 23 miRNAs, 184 pi RNAs, 55 snRNAs, 7 snoRNAs, and 5 tRNAs were identified. The highest number of circRNAs were found in our dataset, followed by the piRNAs, snRNAs, miRNAs, snoRNAs, and tRNAs. A total of 22,581 ncRNAs were commonly found in both the animals (22,417 circRNAs, 11 miRNAs, 114 pi

RNAs, and 39 snRNAs). However, 12,174 ncRNAs were uniquely observed in Changthangi (12,023 circRNAs, 11 miRNAs, 74 pi RNAs, 20 snRNAs, 19 snoRNAs, and 27 tRNAs), and 19,122 ncRNAs were uniquely observed in Muzaffarnagri sheep (18,996 circRNAs, 12 miRNAs, 86 pi RNAs, 16 snRNAs, 7 snoRNAs, and 5 tRNAs) (Figure 4.10). Further analysis was conducted on the miRNAs that were observed both in Changthangi and Muzaffarnagri populations, as well as miRNAs that were unique to Changthangi and Muzaffarnagri (Table 4.3).

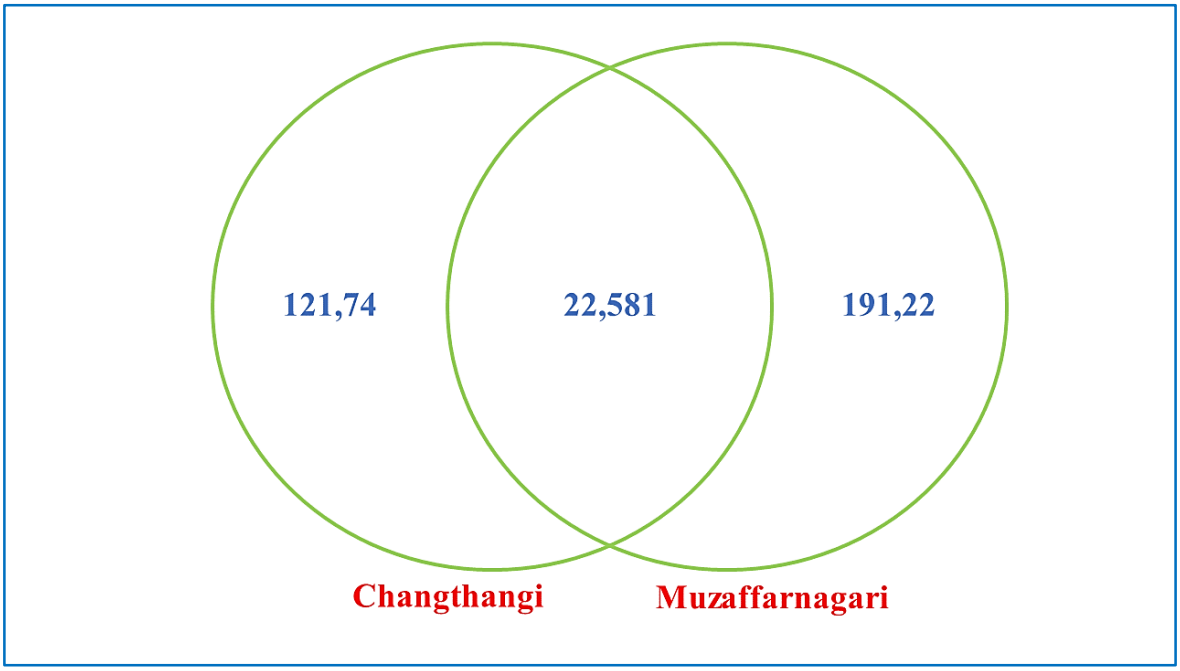


Figure 4.10: 22,581 ncRNAs were commonly found in both the animals, 12,174 ncRNAs were uniquely observed in Changthangi, and 19,122 ncRNAs were uniquely observed in Muzzafarnagri sheep

Table 4.3: miRNAs that were observed in Changthangi and Muzzafarnagri populations, as well as miRNAs that were unique to Changthangi and Muzzafarnagri

Changthangi and Muzaffarnagari	Changthangi	Muzaffarnagari
let-7f-2-3p	mir-4481	miR-15a-3p
let-7f-5p	mir-4532	miR-15a-5p

miR-1244	miR-4680-3p	miR-199a-3p
miR-1282	mir-5047	miR-199a-5p
miR-3120-3p	miR-196b-3p	miR-29b-2-5p
miR-3120-5p	miR-196b-5p	miR-29b-3p
miR-3661	miR-3064-5p	miR-3064-3p
miR-374c-3p	miR-3074-3p	miR-324-3p
miR-374c-5p	miR-3074-5p	miR-324-5p
miR-4444	miR-421	miR-4529-3p
miR-4680-5p	mir-6879	mir-598
		miR-7705

4.9 Target prediction and analysis of candidate miRNAs

The potential miRNAs were predicted by miRNet for the common as well as unique miRNAs observed in the two breeds. The analysis revealed a total of 1104 predicted target genes for the miRNAs that were commonly observed in both breeds, while 1219 and 2023 predicted target genes were identified for the unique miRNAs of Changthangi and Muzzafarnagri, respectively. In order to enhance the visualization of the data, miRNA-mRNA networks were constructed by using the miRNet database, as illustrated in Figure 4.11.

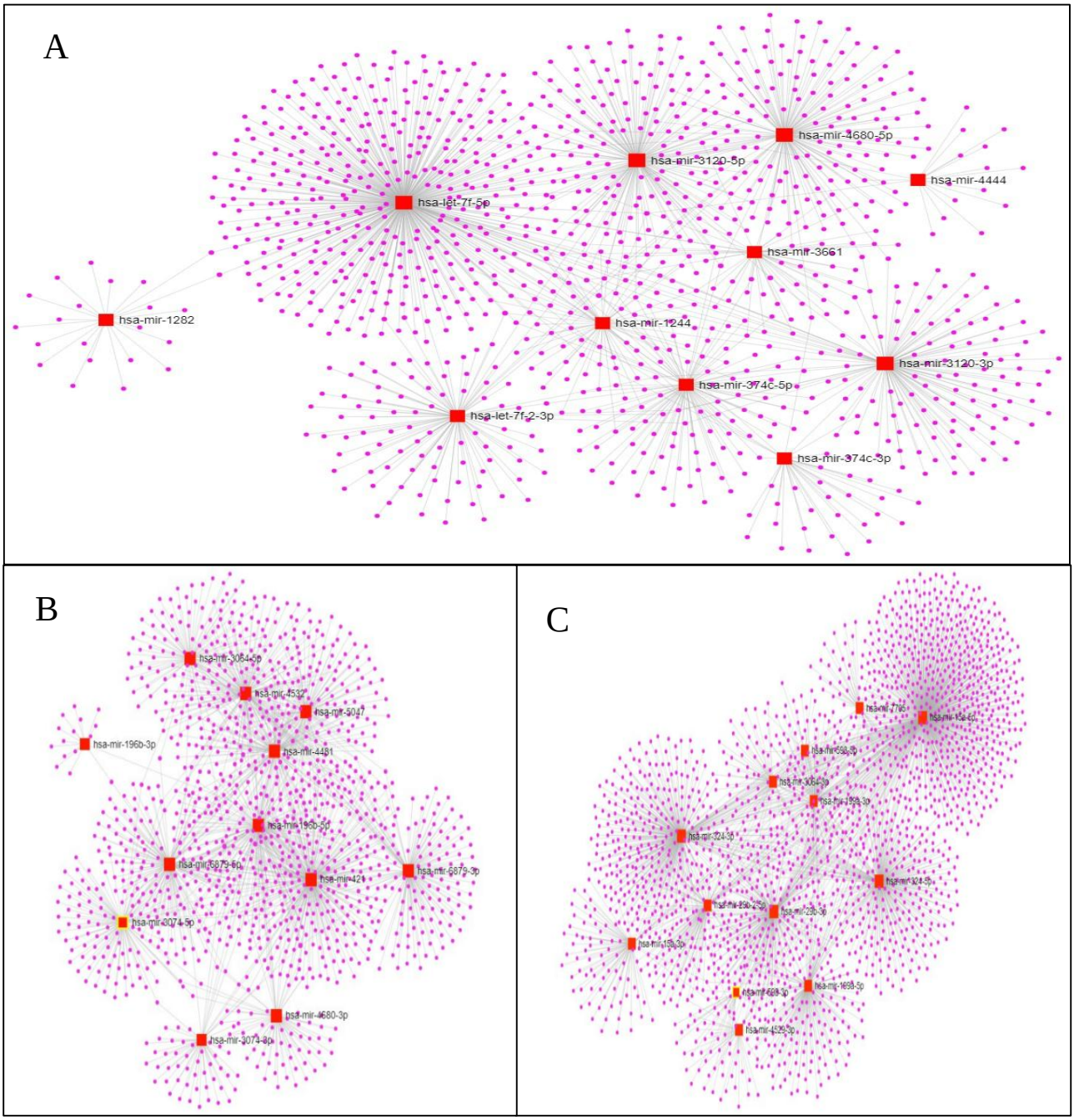


Figure 4.11: The predicted target genes of potential miRNAs. (A) for common observe miRNAs; (B) for uniquely observed Changthangi miRNA; (C) for uniquely observed Muzzafarnagri miRNA.

4.10 Reference genes primer specificity

This study evaluated the expression stability of 18 RGs in various sheep tissues, including rumen, heart, lung, liver, intestine, spleen, kidney, testis, skin, and muscle. The specificity of each primer pair was evident by its specific amplification in agarose gel, in addition to the identification of a single peak in the melting curve analysis. The five-point standard curve, which was produced by 10-fold serial dilution of cDNA revealed that the amplification efficiencies were more than 95% for all the genes. Table 4.4 provides information about the PCR efficiency, slope, correlation coefficient (R2) for each primer pair for the 18 investigated genes. The average Ct values for different genes ranged from 21.29 (*RPL19*) to 31.75 (*TBP*). Based on the mean Ct values, the genes could be categorized into 3 types: Highly expressed (*RPL19*, *RPS28*, RPL0, *RPS15*, *RPS9*) with Ct values ranging from 21.29 to 23.88, moderately expressed with Ct values varying from 25.42 to 27.94 (*PPIB*, *ACTB*, *B2M*, *PPIA*, *PGK1*, *GAPDH*, *BACH1*, *UXT*), and lowly expressed with Ct values between 29.76 to 31.75 (*YWHAZ*, *SDHA*, *HMBS*, *HPRT1*, *TBP*) (Table 4.5).The expression profiles of the potential RGs in various samples are displayed as the mean Ct values obtained from qPCR along with SD of the Ct in a box-whisker plot (Figure 4.12).

Table 4.4: PCR efficiency, slope, and correlation coefficient (R2) for each primer pair for the studied reference genes

Gene	Slope	PCR efficiency	R^2
ACTB	-3.14	-97.91	0.998
BACH1	-3.11	-97.90	1.000
B2M	-3.00	-97.85	1.000
GAPDH	-3.23	-97.92	0.992
HMBS	-3.12	-97.91	0.999
HPRT1	-2.75	-97.02	0.999
PGK1	-3.00	-97.87	0.987
PPIA	-3.01	-97.85	0.997

PPIB	-3.16	-97.92	0.997
RPLP0	-3.01	-97.84	0.991
RPL19	-3.05	-98.87	0.999
RPS9	-3.12	-97.86	0.996
RPS15	-3.14	-97.94	0.998
RPS28	-3.00	-97.86	0.998
SDHA	-3.14	-97.91	0.994
TBP	-3.05	-97.86	0.995
UXT	-3.11	-97.86	0.999
YWHAZ	-3.19	-97.93	0.996

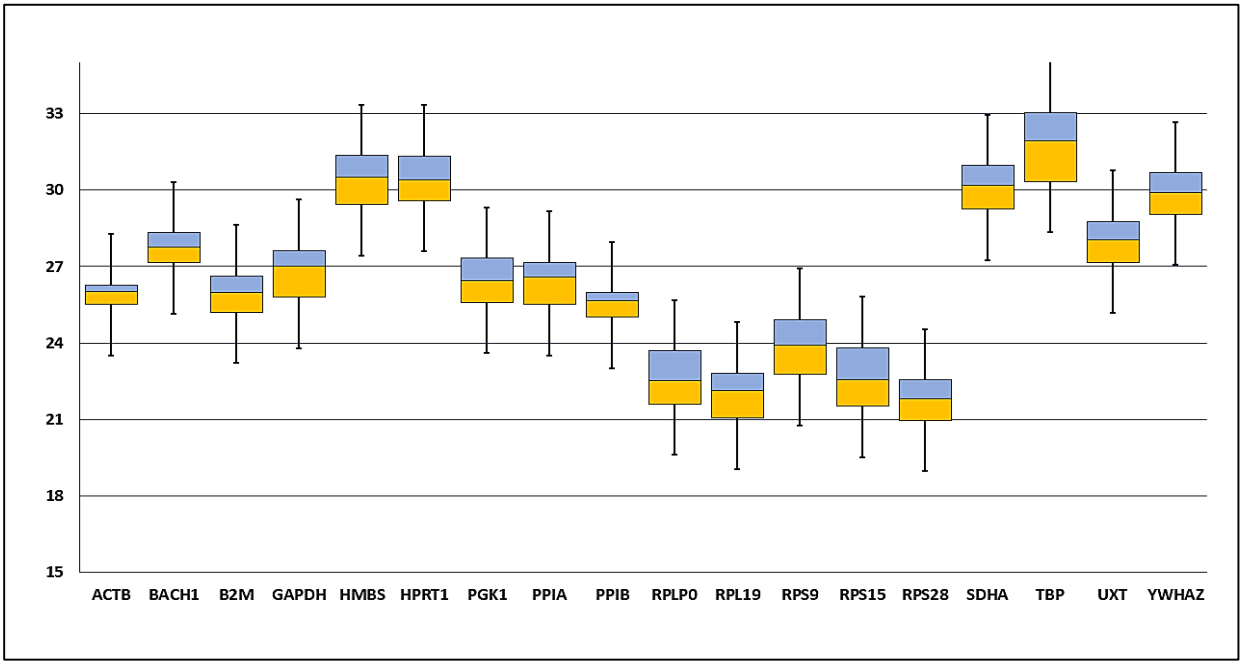


Figure 4.12: Box-whisker plot showing the Ct values of 18 potential reference genes. The whiskers depict maximum and minimum values, while, the median is shown as a line across the box.

Table 4.5: Ct values of individual reference genes in different ovine tissues

S.no	SAMPLE NAME	*ACTB*	*BACH1*	*B2M*	*GAPDH*	*PPIB*	*UXT*	*RPS28*	*RPS15*	*YWHAZ*	RPLP0	*PPIA*	*HPRT1*	*HMBS*	*RPS9*	*TBP*	*PGK1*	*RPL19*	*SDHA*
1	Skin1	24.61	27.53	25.16	24.31	25.22	27.16	19.58	22.08	30.96	20.27	26.55	29.7	28.6	21.95	29.33	24.66	18.47	28.04
2	Skin2	23.95	25.43	24.47	25.82	24.96	26.4	18.58	21.17	31.04	20.63	26.51	29.4	29.8	24.93	29.23	25.52	18.38	28.35
3	Skin3	26.05	28.24	26.39	27.47	25.66	26.93	22.5	23.81	31.36	23.77	26.84	30.66	31.67	24.03	32.77	26.94	22.39	31.94
4	Skin4	26.83	28.3	27.59	24.98	26.05	28.91	22.85	23.68	31.13	23.11	26.18	31.95	30.59	23.57	31.77	26.73	22.18	30
5	Skin5	24.3	27.34	25.17	27.04	23.75	28.04	21.35	21.19	29.56	21.57	25.83	29.37	31.15	22.39	30.38	27.3	21.21	29.78
6	Muscle1	25.62	27.43	23.29	27.19	25.33	26.73	20.87	22.11	30.75	21.89	27.22	30.19	29.06	22.75	29.78	24.66	21.59	28.43
7	Muscle2	26.1	28.68	25.41	26.37	25.76	28.52	22.9	21.64	31.41	20.57	27.43	31.34	31.59	25.48	32.65	28.37	22.8	31.57
8	Muscle3	26.42	28.79	25.51	24.87	25.81	27.24	22.15	23.55	30.84	22.89	27.55	30.7	32.2	23.59	32.71	27.3	22.38	30.47
9	Muscle4	26.19	28.09	26.84	24.54	25.88	27.48	22.74	23.39	29.57	22.98	25.5	29.88	32.03	24.6	33.24	25.85	20.23	28.84
10	Muscle5	25.07	26.15	23.32	25.12	23.84	28.61	20.01	21.8	29.22	20.86	25.69	30.41	28.16	24.84	29.66	25.65	18.9	28.11
11	Kidney1	25.49	27.35	25.22	27.68	25.61	28.95	20.17	21.98	30.95	21.81	26.86	30.54	31.75	24.61	33.28	26.05	20.62	29.69
12	Kidney2	25.96	28.17	25.93	27.44	25.59	29.24	22.9	24.69	30.09	22.54	26.59	32.46	31.38	25.08	32.32	26.09	22.39	31.19
13	Kidney3	26	28.03	25.29	27.12	25.55	28.13	21.27	23.85	28.26	23.96	25.01	29.27	30.44	24.13	31.91	27.32	21.72	29.73
14	Kidney4	25.9	27.04	25.93	26.97	25.69	29.08	22.75	24.22	29.28	23.83	25.85	29.11	29.68	24.96	33.42	28.32	20.45	30.42
15	Kidney5	25.57	29.48	25.72	27.95	25.19	27.94	19.01	23.61	29.61	23.12	25.53	31.72	31.12	23.55	31.82	28.02	21.9	29.64
16	Intestine1	26.76	26.56	26.99	26.66	24.68	27.99	21.74	20.1	29.56	21.6	25.44	31.66	29.32	24.77	31.87	27.84	19.66	30.89
17	Intestine2	25.96	28.45	26.7	28.05	25.91	28.04	22.64	23.39	31.14	21.66	27.49	32.47	31.69	23.29	32.32	28.45	20.5	31.25
18	Intestine3	25.9	28.08	26.2	27.74	26.34	26.34	21.92	21.01	30.49	22.36	27.27	32.5	31.61	23.88	33.46	27.31	21.22	31.84
19	Intestine4	26.46	27.67	26	26.79	26.51	27.19	21.58	22.65	28.86	22.97	27.58	31.97	29.94	24.52	31.18	28.55	22.09	30.1
20	Intestine5	26.19	26.8	26.14	27.94	25.78	27.16	22.23	24.62	29.03	21.52	27.15	29.69	30.09	24.56	29.94	28.11	21.79	30.15
21	Liver1	24.93	25.71	24.2	25.62	23.22	28.55	19.34	20.75	29.62	20.02	24.52	32.49	31.67	23.09	30.32	26.52	21.37	31.89
22	Liver2	26.21	27.61	26.32	27.71	25.54	26.01	20.03	22.86	27.35	23.17	25.81	31.34	30.77	21.75	29.36	25.1	20.23	30.16
23	Liver3	26.05	28.48	26.47	26.29	26.54	28.76	22.36	24.07	28.53	21.63	27.52	31.27	29.86	23.81	33.15	26.02	21.87	29.33
24	Liver4	26.53	28.17	26.81	27.22	26.49	27.4	22.05	23.65	29.71	23.16	25.32	31.07	29.31	23.16	32.98	27.33	21.75	30.56
25	Liver5	26.16	26.75	26.76	28	26.36	28.78	21.86	21.42	30.64	21.12	27.16	28.91	31.41	22.06	32.7	26.08	22.57	31.88
26	Rumen1	26.14	25.83	26.57	27.92	25.65	27.42	21.36	22.43	30.69	22.22	26.41	29.54	28.89	21.89	30.96	26.06	19.04	29.11
27	Rumen2	25.63	27.26	23.05	24.63	23.96	26.5	20.88	22.13	29.27	21.04	24.88	29.34	28.5	21.21	29.41	25.58	22.18	29.52
28	Rumen3	27.29	29.84	28.02	27.6	26.44	28.87	22.53	23.69	30.29	24.09	27.95	31.34	30.72	23.97	32.73	24.89	21.82	28.98
29	Rumen4	25	28.05	26.96	27.07	24.69	28.8	21.87	22.25	28.92	21.89	24.35	30.03	30.14	22.79	31.2	27.88	21.36	29.1
30	Rumen5	25.97	27.66	26.82	25.79	25.97	29.73	21.35	24.54	30.95	24.27	26.81	30.83	31.04	24.54	31.12	25.9	21.34	31.8
31	Lung1	25.44	26.79	24.56	24.8	23.15	26.37	19.97	20.15	27.62	20.05	24.13	29.03	28.57	22.59	29.17	24.91	19.37	28.05
32	Lung2	27.68	27.16	25.21	26.16	23.86	26.19	20.52	20.58	28.04	21.18	24.19	30.83	29.37	21.29	29.68	25.19	19.79	29.62
33	Lung3	26.23	29.09	26.31	27.63	26.21	28.51	21.56	21.00	30.43	21.34	26.09	31.68	29.58	24.63	33.27	25.1	21.57	30.85
34	Lung4	25.76	28.92	26.27	25.18	25.95	29.31	20.57	23.83	29.93	24.1	26.72	32.72	31.18	25.25	33.11	26.38	22.34	30.96
35	Lung5	25.97	28.01	26.65	27.89	25.64	29.55	20.45	24.4	29.88	22.64	27.07	30.99	31.64	25.14	32.44	26.38	21.92	31.24
36	Heart1	24.8	27.19	25.56	24.12	24.02	26.85	20.49	22.06	28.88	22.00	25.08	28.95	28.86	22.56	30.05	26.36	20.82	29.00
37	Heart2	27.23	28.15	27.94	26.51	26.61	28.3	21.24	22.09	30.4	22.48	26.73	30.27	28.79	22.07	32.19	24.2	20.20	30.83
38	Heart3	26.53	29.05	26.36	28.1	26.17	27.27	22.91	21.27	30.82	21.74	27.41	32.33	30.02	24.7	33.46	26.27	22.47	30.5
39	Heart4	25.83	25.64	25.44	25.33	24.76	27.96	21.41	24.37	30.56	23.17	27.02	29.55	30.88	23.8	33.16	26.65	22.47	28.25
40	Heart5	25.47	27.76	25.44	26.12	24.69	28.71	21.83	24.13	29.88	23.82	27.45	30.4	31.25	25.25	31.71	25.13	22.33	29.02
41	Testis1	25.73	28.29	25.07	26.64	25.67	29	20.87	21.49	29.12	23.71	27.37	29.73	30.91	26.8	33.06	24.43	20.69	30.17
42	Testis2	26.99	29.04	27.29	27.81	26.5	28.9	20.84	22.49	31.16	23.88	26.59	31.22	29.64	27.67	35.53	27.77	20.66	31.75
43	Testis3	25.88	27.73	25.31	27.66	25.83	28.31	21.86	24.31	27.53	24.26	26.43	28.93	28.97	25.53	33.4	26.69	21.6	30.61
44	Testis4	25.8	27.69	26.57	27.14	25.32	29.54	19.9	24.43	30.23	23.94	27.01	29.24	30.96	25.8	33.08	27.77	21.95	30.17
45	Testis5	23.99	26.99	24.56	26.99	25.76	27.43	21.32	20.88	27.54	24.09	26.55	29.69	28.54	21.59	29.5	25.22	19.55	29.21
46	Spleen1	24.44	25.48	24.41	27.06	24.55	26.73	21.86	21.94	28.32	22.33	24.5	30.04	31.39	23.41	30.31	27.29	21.45	30.32
47	Spleen2	26.24	28.92	27.04	24.46	26.52	28.04	19.7	20.61	30.42	21.85	27.32	29.76	29.72	21.38	31.52	27.84	22.75	30.88
48	Spleen3	24.34	27.29	23.88	26.79	25.26	27.01	20.72	23.41	27.99	23.81	25.28	29.58	30.29	23.82	31.93	27.72	22.41	30.91
49	Spleen4	24.75	27.77	25.63	27.58	25.74	27.99	21.64	23.15	29.91	23.57	25.41	29.63	31.85	25.55	31.89	26.65	22.84	31.66
50	Spleen5	26.44	27.07	26.13	27.2	25.21	28.6	22.44	24.88	30.56	23.61	27.04	30.56	30.91	25.42	32.42	26.89	22.95	32.27

4.11 geNorm analysis for reference gene selection

According to geNorm analysis, reference genes were ranked based on their expression stability values (M), which should be below the threshold value of 1.5. In general, a lower M value indicates greater expression stability, whereas a higher M value denotes lower expression stability. Stability index for all 18 reference genes in combined group analysis was within acceptable range. The RGs were ranked from most stable to least stable (Figure 4.13) based on average expression stability values as follows *ACTB = B2M > PPIB > BACH1 > PPIA > UXT > RPS28 > YWHAZ > HPRT1 > SDHA > HMBS > RPL19 > TBP > GAPDH > RPLP0 > PGK1 > RPS15 > RPS9*. GeNorm analysis indicated that *ACTB* (0.894), *B2M* (0.898), and *BACH1* (0.943) were the most stable genes, while *RPS9* (1.371), *RPS15* (1.346), and *PGK1* (1.325) were the least stable genes.

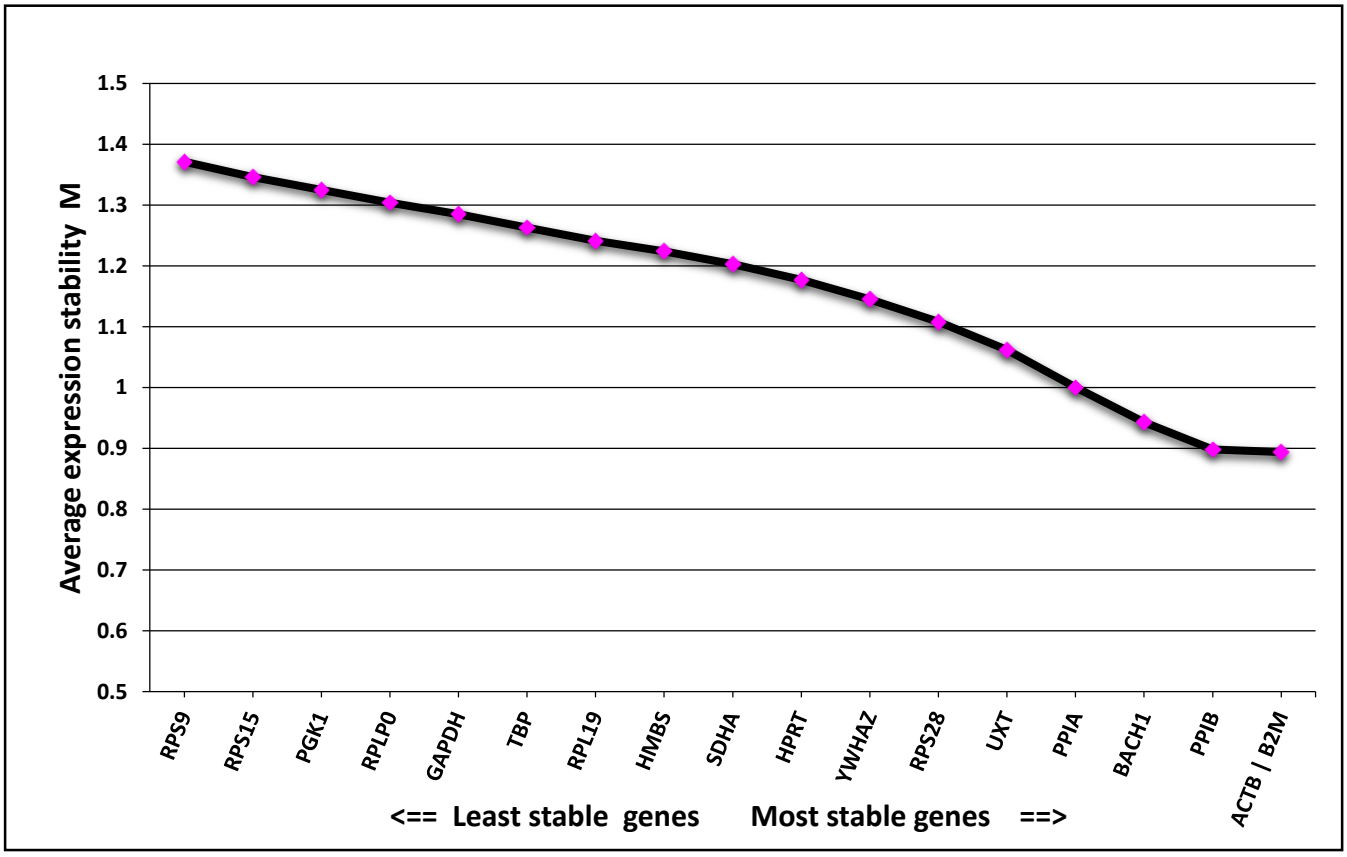


Figure 4.13: The average expression stability (M value) of 18 genes based on geNorm analysis Stability values are decreasing from left to right, ending with the most stable genes.

4.12 NormFinder analysis

The suitability of the gene to be used as a reference gene increases as the stability value decreases in the NormFinder analysis. Figure 4.14 shows the gene stability values obtained from NormFinder. RGs were ranked according to their stability value in the following order: *PPIB>ACTB>BACH1>UXT>B2M>HMBS>RPS28>PPIA>SDHA>RPL19>HPRT1>TBP>YWHAZ>GAPDH>* RPL0 *>RPS15>PGK1>RPS9*. According to NormFinder, *PPIB*, *ACTB*, and *BACH1* were observed to have lower stability values than others, making them the most suitable normalizing genes. *RPS9*, *PGK1*, and *RPS15* have higher stability values of 0.689, 0.685, and 0.665, respectively, which make them the least stable genes. NormFinder additionally provided information on each gene intra- and inter-group variance. *BACH1* and *PPIB* genes showed the best combinations (stability value = 0.340), as shown in Figure 4.15, with minimal inter- and intra-group variances.

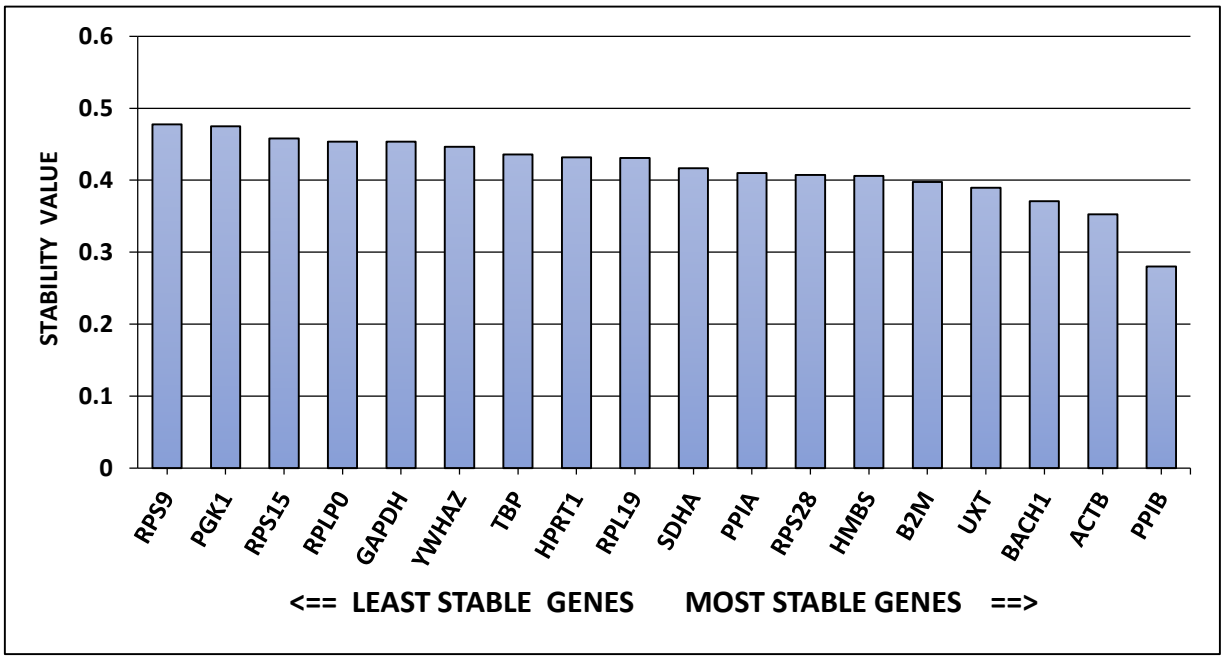


Figure 4.14: Plot made using candidate gene stability values obtained from NormFinder algorithm.

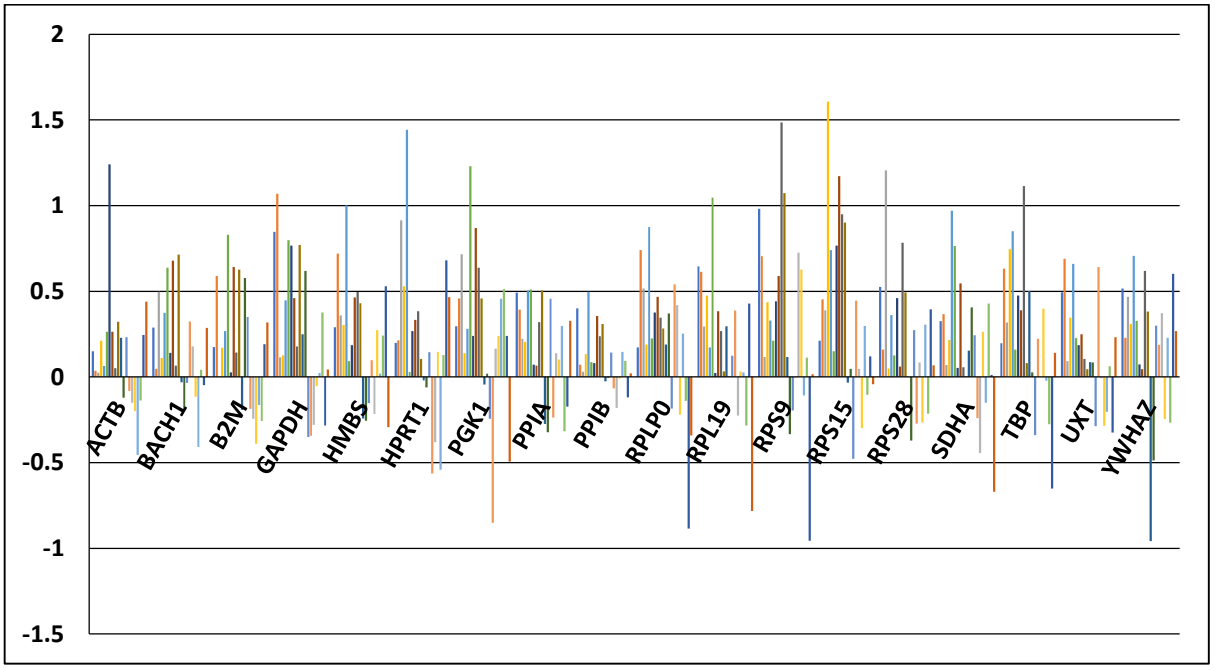


Figure 4.15: Inter- and intra-group variation analysis of reference genes using NormFinder

4.13 BestKeeper analysis

The stability of individual RGs was assessed using BestKeeper analysis based on the raw Ct data. SD value was observed to be <1 for 13 out of the 18 studied genes. Five genes, namely*PGK1*, RPL0, *RPS9*, *RPS15*, and *TBP* had SD >1, hence they are not appropriate for normalization of qPCR data. *ACTB*, *PPIB*, and *BACH1* genes were the most stable genes, while *TBP*, *RPS15*, and *RPS9* genes were the least stable (Table 4.6). Furthermore, inter-gene relationships were estimated for 18 RGs pairs. *PPIA*/*PPIB* (0.657), *PPIB*/*B2M* (0.638), RPL0/*RPS15* (0.635), *B2M*/*ACTB* (0.633), and *PPIB*/*BACH1* (0.585) showed strong correlation coefficient (r). Higher value of the reference gene pairs suggests that their expression patterns are similar in different sheep tissues (Table 4.7).

Table 4.6: Analysis parameter-based quantitative cycling points (CP) of 18 potential RGs

	ACTB	*BACH1*	*B2M*	*GAPDH*	*PPIB*	*UXT*	*RPS28*	*RPS15*	*YWHAZ*	RPL0	*PPIA*	*HPRT1*	*HMBS*	*RPS9*	*TBP*	*PGK1*	*RPL19*	*SDHA*
n	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50
geo Mean [CP]	25.8	27.64	25.79	26.6	25.41	27.93	21.28	22.63	29.74	22.45	26.3	30.51	30.33	23.84	31.72	26.48	21.26	30.16
AR Mean [CP]	25.82	27.66	25.82	26.62	25.43	27.95	21.31	22.68	29.77	22.48	26.32	30.53	30.35	23.88	31.76	26.5	21.29	30.18
min [CP]	23.95	25.43	23.05	24.12	23.15	26.01	18.58	20.1	27.35	20.02	24.13	28.91	28.16	21.21	29.17	24.2	18.38	28.04
max [CP]	27.68	29.84	28.02	28.1	26.61	29.73	22.91	24.88	31.41	24.27	27.95	32.72	32.2	27.67	35.53	28.55	22.95	32.27
std dev [+/- CP]	0.62	0.8	0.92	0.99	0.7	0.85	0.9	1.21	0.93	1.06	0.88	0.96	0.99	1.19	1.25	1	0.98	0.97
CV [% CP]	2.42	2.91	3.55	3.72	2.75	3.03	4.22	5.35	3.12	4.7	3.33	3.15	3.26	4.98	3.95	3.76	4.61	3.2
min [x-fold]	-3.61	-4.63	-6.69	-5.56	-4.8	-3.79	-6.5	-5.79	-5.26	-5.38	-4.51	-3.02	-4.5	-6.17	-5.86	-4.85	-7.34	-4.34
max [x-fold]	3.67	4.59	4.68	2.84	2.29	3.48	3.09	4.74	3.17	3.53	3.13	4.64	3.66	14.26	14.01	4.2	3.24	4.32
std dev [+/- x-fold]	1.54	1.75	1.89	1.99	1.62	1.8	1.87	2.32	1.9	2.08	1.83	1.95	1.98	2.28	2.38	2	1.97	1.95

N =Number of samples, geo Mean =geometric mean of CP, Mean [CP]=arithmetic mean of CP, min [CP] and max [CP]=extreme values of CP, Std dev [±CP]=standard deviation of the CP), CV [%CP] =coefficient of variation expressed as a percentage on the CP values, min [x-fold] and max [x-fold] = extreme values of expression levels expressed as absolute x-fold over or under the coefficient, std dev[±x-fold] = standard deviation of the absolute regulation coefficients.

Table 4.7: Analysis of repeated pair-wise gene correlation using the BestKeeper index

	ACTB	*BACH1*	*B2M*	*GAPDH*	*PPIB*	*UXT*	*RPS28*	*RPS15*	*YWHAZ*	RPL0	*PPIA*	*HPRT1*	*HMBS*	*RPS9*	*TBP*	*PGK1*	*RPL19*	*SDHA*
BACH1	0.456	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
p-value	0.001	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
B2M	0.633	0.504	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
p-value	0.001	0.001	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
GAPDH	0.169	0.164	0.258	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
p-value	0.24	0.254	0.071	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
PPIB	0.455	0.585	0.638	0.384	-	-	-	-	-	-	-	-	-	-	-	-	-	-
p-value	0.001	0.001	0.001	0.006	-	-	-	-	-	-	-	-	-	-	-	-	-	-
UXT	0.165	0.293	0.396	0.137	0.271	-	-	-	-	-	-	-	-	-	-	-	-	-
p-value	0.253	0.039	0.004	0.341	0.057	-	-	-	-	-	-	-	-	-	-	-	-	-
RPS28	0.392	0.276	0.396	0.275	0.419	0.173	-	-	-	-	-	-	-	-	-	-	-	-
p-value	0.005	0.052	0.004	0.053	0.002	0.228	-	-	-	-	-	-	-	-	-	-	-	-
RPS15	0.174	0.216	0.264	0.171	0.326	0.397	0.35	-	-	-	-	-	-	-	-	-	-	-
p-value	0.227	0.131	0.064	0.235	0.021	0.004	0.013	-	-	-	-	-	-	-	-	-	-	-
YWHAZ	0.23	0.232	0.338	0.032	0.345	0.298	0.158	0.085	-	-	-	-	-	-	-	-	-	-
p-value	0.109	0.105	0.016	0.824	0.014	0.036	0.272	0.558	-	-	-	-	-	-	-	-	-	-
RPL0	0.167	0.334	0.316	0.263	0.448	0.32	0.27	0.635	-0.094	-	-	-	-	-	-	-	-	-
p-value	0.247	0.018	0.025	0.065	0.001	0.023	0.058	0.001	0.517	-	-	-	-	-	-	-	-	-
PPIA	0.282	0.375	0.341	0.23	0.657	0.315	0.288	0.3	0.539	0.24	-	-	-	-	-	-	-	-
p-value	0.047	0.007	0.015	0.107	0.001	0.026	0.042	0.035	0.001	0.094	-	-	-	-	-	-	-	-
HPRT1	0.384	0.433	0.313	0.2	0.27	0.163	0.208	0.021	0.292	-0.02	0.234	-	-	-	-	-	-	-
p-value	0.006	0.002	0.027	0.163	0.058	0.259	0.148	0.887	0.04	0.886	0.102	-	-	-	-	-	-	-
HMBS	0.022	0.219	0.225	0.234	0.176	0.291	0.26	0.335	0.329	0.244	0.262	0.329	-	-	-	-	-	-
p-value	0.881	0.126	0.116	0.103	0.222	0.04	0.069	0.017	0.02	0.088	0.066	0.02	-	-	-	-	-	-
RPS9	0.099	0.215	0.137	0.255	0.253	0.521	0.201	0.383	0.228	0.399	0.32	0.192	0.338	-	-	-	-	-
p-value	0.492	0.133	0.345	0.074	0.076	0.001	0.161	0.006	0.111	0.004	0.024	0.181	0.016	-	-	-	-	-
TBP	0.455	0.536	0.54	0.353	0.62	0.535	0.456	0.36	0.377	0.473	0.429	0.282	0.414	0.612	-	-	-	-
p-value	0.001	0.001	0.001	0.012	0.001	0.001	0.001	0.01	0.007	0.001	0.002	0.047	0.003	0.001	-	-	-	-
PGK1	0.025	0.084	0.19	0.227	0.136	0.112	0.244	0.23	0.077	0.137	0.001	0.141	0.305	0.239	0.313	-	-	-
p-value	0.862	0.563	0.186	0.113	0.345	0.439	0.088	0.108	0.595	0.343	0.992	0.329	0.031	0.094	0.027	-	-	-
RPL19	0.208	0.414	0.167	0.178	0.3	0.29	0.411	0.446	0.178	0.345	0.292	0.2	0.511	0.22	0.443	0.379	-	-
p-value	0.147	0.003	0.247	0.216	0.035	0.041	0.003	0.001	0.216	0.014	0.039	0.163	0.001	0.125	0.001	0.007	-	-
SDHA	0.285	0.264	0.356	0.408	0.381	0.307	0.31	0.14	0.27	0.24	0.187	0.398	0.482	0.306	0.488	0.407	0.505	-
p-value	0.045	0.064	0.011	0.003	0.006	0.03	0.029	0.331	0.057	0.093	0.193	0.004	0.001	0.031	0.001	0.003	0.001	-
BestKeeper vs.	***ACTB***	***BACH1***	***B2M***	***GAPDH***	***PPIB***	***UXT***	***RPS28***	***RPS15***	***YWHAZ***	**RPL0**	***PPIA***	***HPRT1***	***HMBS***	***RPS9***	***TBP***	***PGK1***	***RPL19***	***SDHA***
coeff. of corr. [r]	0.5	0.606	0.638	0.472	0.709	0.573	0.596	0.609	0.431	0.589	0.58	0.44	0.573	0.597	0.829	0.422	0.641	0.627
p-value	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.002	0.001	0.001	0.001	0.001	0.001	0.001	0.002	0.001	0.001

4.14 Comparative ΔCt analysis for reference gene selection

The comparative ΔCt analysis carries out pair-wise comparison and calculates the SD. The average SD was used to evaluate the expression stability of the studied genes. Lower the SD, higher the stability ranking of the corresponding gene. The ΔCt analysis showed the following gene expression stability rankings: *ACTB* (0.625), *PPIB* (0.699), *BACH1* (0.805), *UXT* (0.847), *PPIA* (0.875), *RPS28* (0.899), *B2M* (0.917), *YWHAZ* (0.929), *HPRT1* (0.961), *SDHA* (0.966), *RPL19* (0.981), *HMBS* (0.988), *GAPDH* (0.991), *PGK1* (0.998), RPL0 (1.057), *RPS9* (1.190), *RPS15* (1.212), and *TBP* (1.253) (Figure 4.16). Thus, *ACTB*, *PPIB*, *BACH1* were the most stable genes. On the other hand, *TBP*, *RPS15*, *RPS9* were the least stable genes.

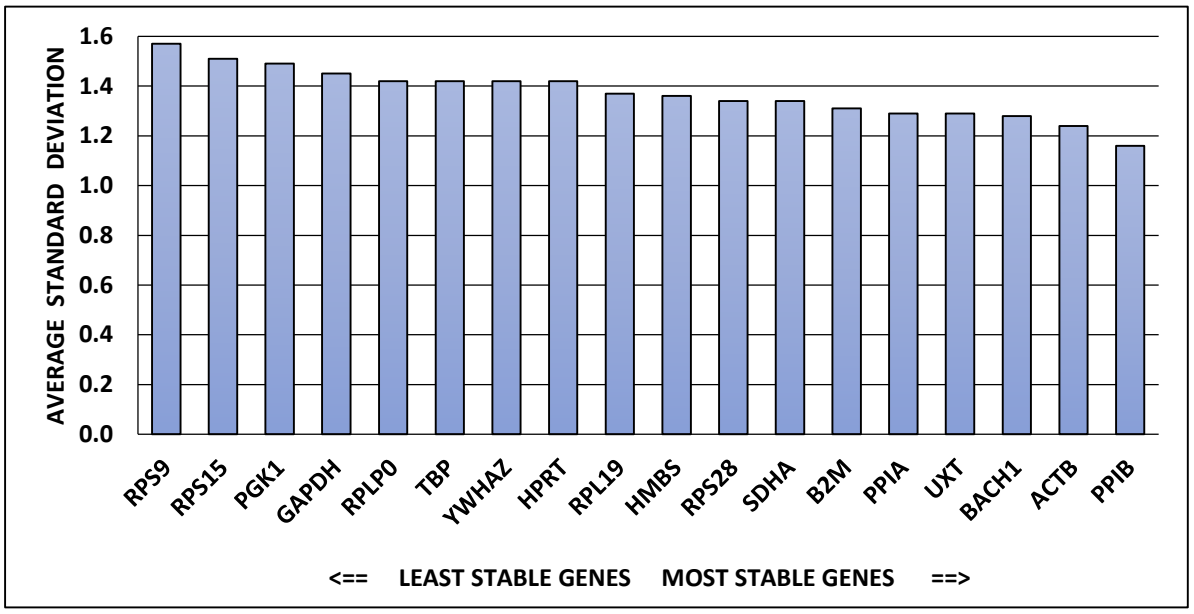


Figure 4.16: Expression stability of reference genes based on ΔCt analysis

4.15 Comprehensive ranking of genes by RefFinder

The RefFinder tool determined the comprehensive ranking of the genes based on combined analysis of different algorithms, namely geNorm, NormFinder, BestKeeper, and ΔC_t (Table 4.8). According to this comprehensive analysis, the stability values of candidate reference genes were ranked as follows: *ACTB* (1.41) >*PPIB* (1.570) >*BACH1* (3.22) >*B2M* (3.98) >*UXT* (4.43), >*PPIA* (5.00) >*RPS28* (7.2) >*SDHA* (8.37) >*HPRT1* (9.95) >*YWHAZ* (10) >*HMBS* (10.17) >*RPL19* (10.72) >*TBP* (13.82) >*GAPDH* (14.23) > RPL0 (14.49) >*PGK1* (15.47) >*RPS15* (17) >*RPS9* (17.48) (Figure 4.17, Table 4.8). Among the 18 genes, *ACTB*, *PPIB*, and *BACH1* genes were the most stable reference genes, and *RPS9*, *RPS15* and *PGK1* genes were the least stable genes.

In this study, different algorithms resulted in minor differences in the ranking orders of the RGs due to variation in the statistical approach being used to calculate the ranking. The top 10 genes ranked by four algorithms were used for the venn diagram analyses (Figure 4.18). A total of 8 genes of the top ten stable reference genes overlapped in this study, including *PPIB*, *ACTB*, *B2M*, *BACH1*, *PPIA*, *UXT*, *RPS28* and *SDHA*. *HPRT1* and *YWHAZ* were found in BestKeeper, ΔCt, and geNorm but not in NormFinder.

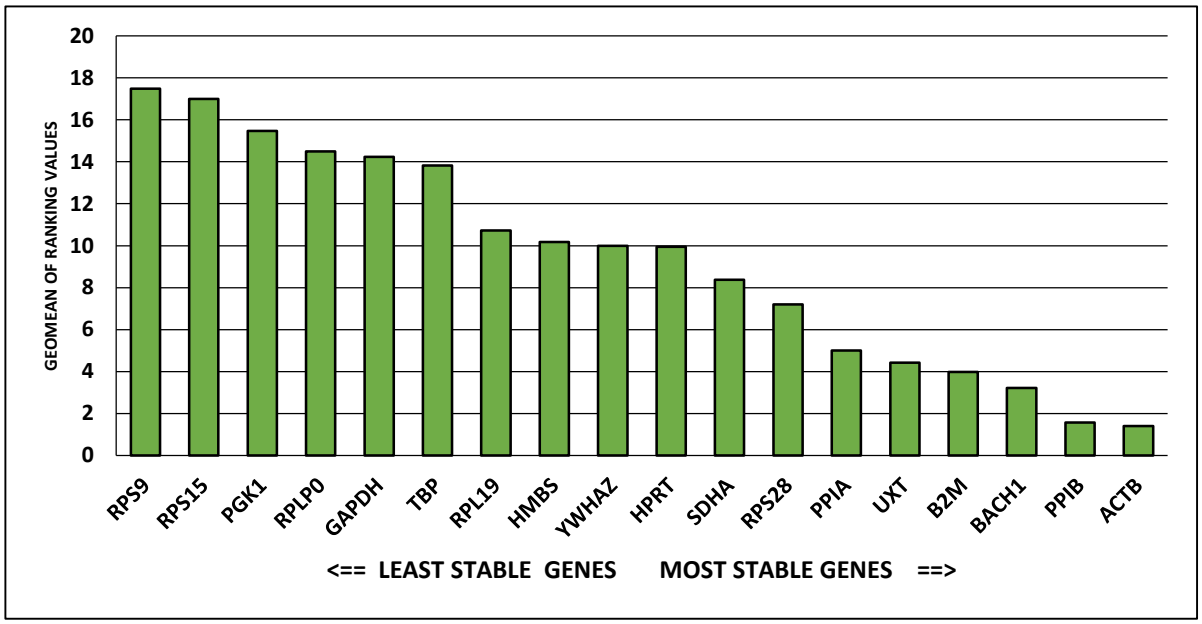


Figure 4.17: Comprehensive ranking of RGs based on RefFinder analysis

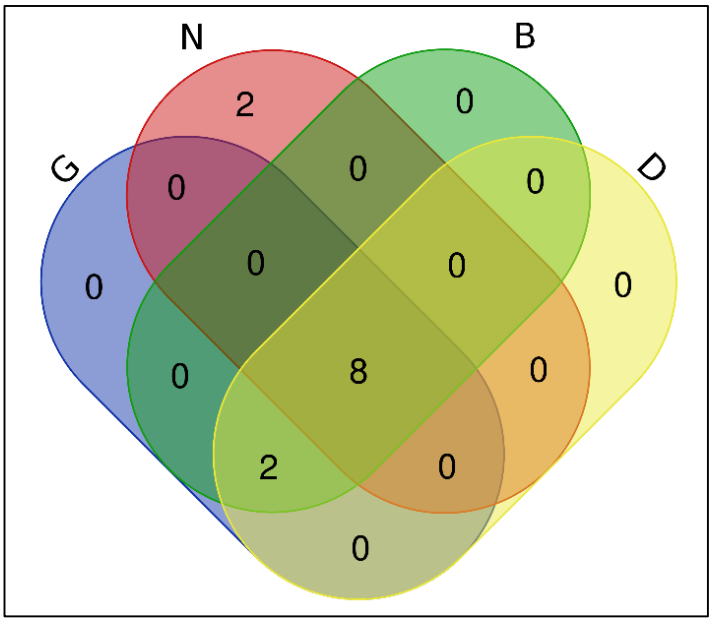


Figure 4.18: Top 10 stable genes identified by geNorm, NormFinder, BestKeeper and comparative ΔCt, are represented in a Venn diagram (G= GeNorm, N= NormFinder, B= BestKeeper, D= ΔCt)

Table 4.8: Overall ranking of suitable reference genes

Ranking	geNorm M value	NormFinder Stability value	BestKeeper Standard Deviation	Delta Ct Average SD	Overall ranking by RefFinder
1	*ACTB* (0.894)	*PPIB* (0.280)	*ACTB* (0.62)	*ACTB* (0.625)	*ACTB* (1.41)
2	*B2M* (0.894)	*ACTB* (0.352)	*PPIB* (0.7)	*PPIB* (0.699)	*PPIB* (1.57)
3	*PPIB* (0.898)	*BACH1* (0.370)	*BACH1* (0.8)	*BACH1* (0.805)	*BACH1* (3.22)
4	*BACH1* (0.943)	*UXT* (0.389)	*UXT* (0.85)	*UXT* (0.847)	*B2M* (3.98)
5	*PPIA* (1.00)	*B2M* (0.397)	*PPIA* (0.88)	*PPIA* (0.875)	*UXT* (4.43)
6	*UXT* (1.062)	*HMBS* (0.406)	*RPS28* (0.9)	*RPS28* (0.899)	*PPIA* (5.00)
7	*RPS28* (1.108)	*RPS28* (0.407)	*B2M* (0.92)	*B2M* (0.917)	*RPS28* (7.20)
8	*YWHAZ* (1.145)	*PPIA* (0.409)	*YWHAZ* (0.93)	*YWHAZ* (0.929)	*SDHA* (8.37)
9	*HPRT1* (1.177)	*SDHA* (0.416)	*HPRT1* (0.96)	*HPRT1* (0.961)	*HPRT1* (9.95)
10	*SDHA* (1.203)	*RPL19* (0.430)	*SDHA* (0.97)	*SDHA* (0.966)	*YWHAZ* (10.00)
11	*HMBS* (1.224)	*HPRT1* (0.431)	*RPL19* (0.98)	*RPL19* (0.981)	*HMBS* (10.17)
12	*RPL19* (1.241)	*TBP* (0.435)	*GAPDH* (0.99)	*HMBS* (0.988)	*RPL19* (10.72)
13	*TBP* (1.263)	*YWHAZ* (0.446)	*HMBS* (0.99)	*GAPDH* (0.991)	*TBP* (13.82)
14	*GAPDH* (1.285)	*GAPDH* (0.453)	*PGK1* (1)	*PGK1* (0.998)	*GAPDH* (14.23)
15	*RPLP0* (1.304)	*RPLP0* (0.454)	*RPLP0* (1.06)	*RPLP0* (1.057)	*RPLP0* (14.49)
16	*PGK1* (1.325)	*RPS15* (0.458)	*RPS9* (1.19)	*RPS9* (1.190)	*PGK1* (15.47)
17	*RPS15* (1.346)	*PGK1* (0.474)	*RPS15* (1.21)	*RPS15* (1.212)	*RPS15* (17.00)
18	*RPS9* (1.371)	*RPS9* (0.477)	*TBP* (1.25)	*TBP* (1.253)	*RPS9* (17.48)

4.16 RNA Sequencing data validation

We randomly selected 10 genes (*TYR*, *TYRP1*, *DCT*, *OCA2*, *PMEL*, *SLC45A2*, *KRTAP6*, *KRTAP7*, *KRTAP13*, and *KRT19*) and used qPCR to evaluate their relative expression in Changthangi and Muzzafarnagri sheep. The qPCR data was normalized by using the *PPIB* and *ACTB* as reference genes. Our RNAseq results were supported by the qPCR expression profiles of these genes, which showed a similar pattern except for minor differences in the magnitude (Figure 4.21).

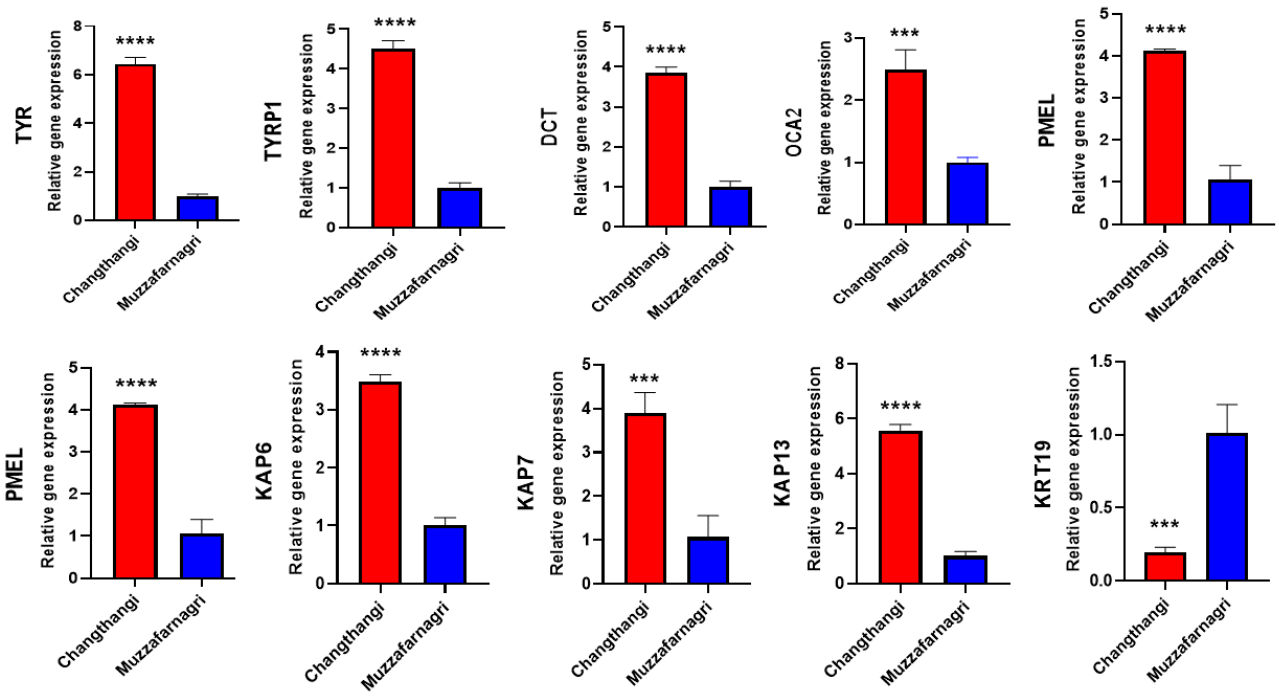


Figure 4.19: Gene expression pattern of differentially expressed genes between Changthangi and Muzzafarnagri sheep by qRT-PCR.

CHAPTER–5

Discussion

DISCUSSION

In this study, comparative skin transcriptome analysis was carried out for Changthangi and Muzzafarnagri sheep which are characterized by contrasting phenotypes and acclimatization to diverse habitats. Changthangi sheep produces fine fiber and thrives in cold deserts of Ladakh, while, Muzzafarnagri sheep produces coarse fleece and is adapted to the tropical climate. To identify molecular drivers associated with coat color in sheep, care was taken to ensure different coat colors in the investigated breeds, black in Changthangi and white in Muzzafarnagri sheep. The most conspicuous set of genes which were up-regulated in Changthangi sheep were associated with GO terms such as melanin biosynthesis, developmental pigmentation, tyrosine metabolism, melanosome, and monophenol monooxygenase activity. These include *PMEL*, *TYRP1*, *TYR*, *SLC45A2*, *DCT*, *OCA2*, and *MLANA*. We validated the higher expression of these genes in Changthangi sheep in comparison to Muzzafarnagri sheep by qPCR as well. Melanin pigment is produced by melanocytes in many tissues including skin and hair follicles, thorough a well-orchestrated mechanism involving many biochemical events. The amount and type of melanin synthesized by melonocytes is governed by genetic as well as environmental factors. In addition to its role in coloring of skin and hair, melanin also contributes to photoprotection and thermoregulation (D'Mello *et al.,* 2016). According to Reactome pathway shown in Figure 4.6, pheomelanin (red) and eumelanin (black) are the two types of pigments produced by melanocytes within the melanosomes. The starting material for melanin biosynthesis is l-Tyrosine, which is oxidized to form L-Dopaquinone by the action of *TYR*, *OCA2*, and *SLC45A2*. In the absence of glutathione/cysteine, L-Dopaquinone leads to synthesis of eumelanin by the enzymatic action of *TYRP1* and *TYRP2 (DCT)*. However, if glutathione/cysteine is present, then pheomelanin pigment is synthesized (Bhat *et al.,* 2019).

Among the upregulated genes observed in Changthangi sheep, *OCA2* is involved in sorting and trafficking melanosomes, *TYR* and *DCT* are critical for determining the quality and quantity of melanin produced, and *PMEL* and *MLANA* code for important structural proteins (Yamaguchi *et al.,* 2007). The silver homolog premelanosome protein PMEL is central to the structure and function of premelanosomes and is associated with pigmentation. A proteolytic fibrillary matrix is formed in melanosomes by *PMEL*, which acts as a scaffold to deposit melanin (Solano *et al.,* 2000*).* In many species such as dog, mouse, horse, and chicken, dilution in coat color has been associated with mutation in the *PMEL* gene (Knaust *et al.*, 2020). *PMEL* and *TYR* have been linked to pigmentation patterns across dogs, cats, horses, cows, pigs, sheep, ferrets, minks, camelids, and rabbits (Strain, 2015). Solute carrier family 45 member 2 (*SLC45A2)* regulatesmelanin synthesis

by being involved in transportation and processing of *TYR*, *TYRP1* and *DCT* in cells. The importance of this gene in pigmentation can be understood from the albino phenotype observed in dogs that have mutation in *SLC45A2* (Wijesena & Schmutz, 2015). A previous study in Chinese sheep has also documented the role of *TYR*, *TYRP1*, *DCT*, *SLC45A2*, *PMEL*, and *MLANA* in regulation of coat color in four different phenotypes including white, light brown, cyan gray and black coat animals. These genes were suggested as candidates for selecting sheep for natural coat colors (Yao *et al.*, 2019). Recently, tyrosine metabolism and melanogenesis were identified to be enriched pathways in skin transcriptome analysis of Minxian black fur sheep and here again; *OCA2, DCT, TYRP1* and *PMEL* were identified to be involved in pigmentation (Shi *et al.,* 2021). Sufficient scientific literature to support the role of melanogenesis related genes in governing skin coloration is available from other species such as pigs, donkeys, goats, chicken, and ducks (Xu *et al.,* 2019; Wang *et al.,* 2020; Bhat *et al.,* 2019; Li *et al.,* 2015; Arora *et al.,* 2023). For instance, significantly higher expression of *MLANA*, *TYR*, *PMEL*, and *TYRP1* has been reported in black skin relative to white skin in Chinese Wuzhishan pigs. Similarly, the expression of *TYR* and *TYRP1* genes was observed in pigmented skin tissues, but not in non-pigmented tissue in spot ducks. Neither melanocyte progenitor cells nor melanocytes were found in unpigmented areas of spot ducks (Li *et al.,* 2015). Upregulated expression of many melanogenesis specific genes namely, *TYR, MLANA, OCA2, SLC24A5*, *TRPM1*, and *KITLG* has been documented in the skin of donkeys with black or chestnut coat colors as compared to gray skinned animals (Wang *et al.,* 2020). Even in Pashmina goats from the Ladakh region of India, significantly higher expression of genes related to melanin production (*TYR, TRP1, TRP2, DCT*, and *PMEL*) has been reported in black and brown skin samples in comparison to the white skinned goats (Bhat *et al.,* 2019). Recently, a study on comparative gene expression profiling of muscle of black chicken Kadaknath and broiler has identified significantly higher expression of *PMEL, TYR, DCT, OCA2*, and *MLANA* in the former providing scientific evidence for accumulation of melanin in skin, muscle and other organs (Arora *et al.,* 2023).

Wool from sheep is an essential agricultural commodity used in textiles and clothing. The wool fiber is made up of many keratins (KRTs) and keratin-associated proteins (KRTAPs) whose spatial arrangement and chemical bonding in the inter-filamentous matrix determines the physical attributes of fiber (Gong *et al.*, 2016). In this study, we observed higher expression of 4 *KRTAPs* (*KRTAP6, KRTAP7, KRTAP13,* and *KRTAP21*) in Changthangi sheep skin relative to Muzzafarnagri sheep. Super Merino sheep with fine wool have been reported to have higher expression of *KRTAP6* in their skin as compared to Small Tail Han sheep with coarse wool (Zhang *et al.,* 2017). In addition to determining the physico-chemical properties of wool fiber, *KRTAP6*has

also been shown to be associated with differences in crimp wool fibre (Gong *et al.,* 2016). Recent study by Ahlawat et al. (2020) has documented the involvement of *KRTAP6* and *KRTAP7* genes in governing the pashmina fiber quality in Changthangi goats. Moreover, variations in *KRTAP7* have been reported to influence the wool yield and mean staple length in some sheep breeds and their crosses in a study from Pakistan (Ullah *et al.,* 2020). *In situ* hybridization analysis has revealed a strong expression of the *KRTAP7* gene in the cortical layer, hair matrix of the primary follicles, inner root sheath, and secondary hair follicles in Cashmere goats (Jin *et al.,* 2011). *KRTAP13* is an important keratin-associated protein whose expression significantly varies during transformation of wool follicle as has been substantiated in Tibetan sheep by comparison of gene expression profiles of wool follicle bulb samples at different stages (anagen, catagen and late telogen) (Liu *et al.,* 2015). Li et al., 2019 have identified sequence variations in *KRTAP21* in Merino X Southdown-cross sheep and also reported their association with wool yield, thereby suggesting *KRTAP21* to be a potential gene for improving the wool production in sheep.

We observed interaction of *KRTAP6* with many *HOX* genes (*HOXA9, HOXA13, HOXA11, HOXC8, HOX*C9, and *HOXD11*) in the interaction network of the up-regulated genes in Changthangi sheep. Basically, *HOX* genes constitute a family of transcription factors that are pivotal in determining cell fate and identity during embryogenesis and govern regional properties of different tissues (Jave-Suarez *et al.,* 2002). Although the role of *HOX* genes in skin development is relatively unexplored, their importance in skin patterning have been known since long based on their expression in feather and dermal papillae of hair follicles in chicken and mice, respectively (Yu *et al.,* 2017). The role of Wnt/b-catenin signaling in growth and regeneration of hair follicles is unequivocal (Choi *et al.,* 2013). A recent study has shown that *HOXC* gene controls Wnt signaling pathway and has the potential to activate dormant dermal papilloma cells to support hair follicle regeneration in mice models (Yu *et al.,* 2018). Evidence to supplement the relationship of *HOX* genes and *KRTAPs* in hair growth is available from mice, wherein sequence-specific interaction has been documented between *KRTAP16* and *HOXC13* (Pruett *et al.,* 2004). Interestingly, *HOX* genes, namely *HOXC13, HOXC8*, and *HOXC12* have been identified as important determinants of wool fineness in sheep (Awgulewitsch *et al.,* 2003). Thus, the *HOX* genes identified to be significantly abundantly expressed in Changthangi sheep skin in this study become intriguing candidates for future research.

Ladakh, the home tract of the Changthangi sheep is an inhospitable terrain characterized by extremely cold weather, lower oxygen levels and higher ultraviolet radiation (UVR) (Singh *et al.,* 2013). UVR leads to systemic oxidative stress through generation of reactive oxygen species, which can profoundly impact the health of humans and animals inhabiting the high altitude regions.

In vitro experiments have proved that the blue-black eumelanin pigment with its antioxidant properties provides protection against the oxidative damage to DNA, lipids, and proteins. Thus, higher melanization helps to curtail the UVR associated oxidative stress (Shih *et al.,* 2020). Path-breaking research aiming to decipher the molecular mechanism of acclimatization to high altitudes in Tibetans has also suggested that skin pigmentation is an adaptation to counteract the ill-effects of UVR. Greater melanin production and darker baseline skin tone in Tibetans help to protect the skin from UV damage (Yang *et al.,* 2022). Since we observed significantly higher expression of many genes of the melanogenesis pathway in Changthangi sheep, it is plausible to speculate that similar mechanisms may be involved in their adaptation to the tough terrains of Ladakh.

The gene expression data generated through the high-throughput sequencing platforms is routinely validated by quantitative real time PCR (qPCR) owing to the sensitivity and reproducibility of the technique (Arora *et al.,* 2021). There are many factors that determine the quality of qPCR results, including primer designing, integrity of purified RNA, presence/absence of various contaminants, quality of cDNA, inclusion of sufficient biological/technical replicates in the experiment, and data normalization using appropriate reference genes. With the availability of robust softwares for primer designing and commercial kits for wet lab experimentation, consistency in many of these factors can be maintained (Hruz *et al.,* 2011). Normalization continues to be a challenge since the expression of the selected RG should be consistent under different experimental treatments, across tissues, and physiological stages (Bustin *et al.,*2009). Every biological study involving a variety of tissues, physiological phases, cell types, disease states, and developmental stages faces a significant challenge in determining the best RG whose expression remains constant without evident variations among samples. In most cases, reference genes like *ACTB* and *GAPDH* are selected at random without appropriate validation. Unfortunately, the utilisation of such experimental RGs may not offer precise normalisation and may lead to suspicion on the accuracy of gene expression estimation (Perez *et al.,* 2008).This suggests that it is crucial to validate the selected RGs for every study.

Over the last ten years, numerous studies have been conducted in sheep with the goal to identify RGs for different experimental conditions involving various tissues to identify a panel of stable normalizers for reliable PCR results (Zang *et al.,* 2011; Jiang *et al.,* 2015; Vorachek *et al.,* 2013; Sonowal *et al.,* 2022; Schulze *et al.,* 2017). This study is an attempt to comprehensively assess 18 RGs from different functional classes (*ACTB, BACH1, B2M, GAPDH, HMBS, HPRT1, PGK1, PPIA, PPIB, RPLP0, RPL19, RPS9, RPS15, RPS28, SDHA, TBP, UXT,* and *YWHAZ*) in 10 different ovine tissues (muscle, skin, kidney, liver, intestine, rumen, lung, testis, heart, and spleen) to identify a panel of RGs that show stable expression across tissues. Four alternative approaches,

namely NormFinder, geNorm, BestKeeper, and ΔCt were employed to examine the stability of gene expression. NormFinder gives ranking to the candidate genes based on the similarity of their expression profiles. It allows comparisons of intra and intergroup variability, as well as the calculation of expression stability values (Andersen *et al.,* 2004). geNorm relies on the determination of the gene-stability value (M) based on the average pair-wise variation of one gene with other investigated genes and also utilizes the pair-wise variation of genes to estimate the standard deviation of the logarithmically transformed expression ratios (Vandesompele *et al.*, 2002). To determine the optimal RGs, BestKeeper uses repeated pair-wise correlation analysis of all pairs of potential RGs, and determines the geometric mean of the most stable genes (Pfaffl *et al.,* 2004). The comparative ΔCt technique calculates the average standard deviation values by comparing the relative level of expression of a RG with those of other RGs (Silver *et al.,* 2006). Different software programs use different algorithms to determine gene expression stability, which may be the reason for the difference in results obtained from these algorithms (Zhu *et al.,* 2015). Thus, majority of recent publications advice using a comprehensive ranking based on multiple algorithms to shortlist the best endogenous control genes (Tiwari *et al.,* 2022; Kaur *et al.,* 2023). Fortunately, all the tested algorithms in our study resulted in similar trends in the ranking of genes. This study recommends a panel of RGs comprising of *ACTB*, *PPIB*, *BACH1*, and *B2M* as normalizers in ovine studies involving tissues such as rumen, heart, lung, liver, intestine, spleen, kidney, testis, skin, and muscle.

ACTB, *PPIB*, *BACH1*, and *B2M* encode proteins that are important in cellular homeostasis. For instance, beta-actin, which is produced by the *ACTB* gene, is necessary for a variety of cytoplasmic processes, including the control of cell shape and motility, as well as nuclear processes, control of gene expression, cell division, and proliferation (Cuvertino *et al.,* 2017). *PPIB* encoding for cyclophilin B protein belongs to the family of *PPIA*ses (peptidyl-prolyl cis-trans isomerases), which control protein folding and maturation by catalysing the cis-trans isomerization of proline imidic peptide bonds (Gothel and Marahiel, 1999). *BACH1* is an essential transcription factor that controls immune function, cell cycle, haematopoiesis, heme homeostasis, and reactive oxygen species (ROS) generation (Zhang *et al.,* 2018). *B2M* is an important component of MHC class I molecules, which presents the antigens of viruses or tumour cells to killer T cells. It also functions as a self-recognition marker to inhibit autologous killer T cell attacks (Dydensborg *et al.,* 2006).

Many scientific groups have identified suitable RGs for livestock species under different experimental conditions. Consistent with our results, *ACTB* has been identified as an endogenous control in cattle milk somatic cells and mammary epithelial cells across lactation stages (Jatav *et al.,* 2016) and peripheral blood mononuclear cells (PBMCs) of high altitude adapted Ladakhi yak

and Zanskar ponies (Tiwari *et al.,* 2022). It has also been identified as the most stable RG for qPCR studies in zebu cattle and their crosses in muscle, liver, and jejunum tissues (Coelho *et al.,* 2022).There are many studies in sheep where *ACTB* has been reported to have the most stable expression in diverse tissues such as mandibular condyle (Jiang *et al.,* 2015), cerebrum and spleen (Garcia-Crespo *et al.,* 2005), pulmonary artery and lung (Passmore *et al.,* 2009), and longissimus thoracis muscle (Lozano and Barragan, 2022).

PPIB has been shown to be the best RG for precise normalization of qPCR data related to intramuscular fat (IMF) deposition in goat longissimus dorsi and biceps femoris muscles across different developmental stages (Zhu *et al.,* 2015). Similarly, to normalise the qPCR data for validation of the comparative transcriptome analysis results of the longissimus thoracis muscles of Changthangi and Barbari goat breeds, *ACTB* and *PPIB* were used as the RGs (Kumar *et al.,* 2021). Because of lowest inter-sample variation, *PPIB* was reported to be a suitable RG in peripheral whole blood of humans suffering from inflammatory diseases (Pachot *et al.,* 2004). Recently, both *ACTB* and *PPIB* were recommended as RGs for comparative expression analyses of cord blood samples from neonates representing different gestational age groups as well as for the comparison of full-term or preterm neonates with adult peripheral blood samples in human subjects (Hieronymus *et al.,* 2021).The third most stable RG identified in our study was *BACH1* whose suitability as a normalizer in the bovine PBMCs for qPCR studies on tick and fly-borne diseases has been substantiated lately (Kaur *et al.,* 2023).

B2M has been commonly used for normalization of qPCR data in ovine bone and mesenchymal stem cells (Schulze *et al.,* 2017), mice haematopoietic stem cells (Matsuzaki *et al.,* 2015), cattle whole blood, PBMCs, cumulus cells and oocytes (Lozano-Villegas *et al.,* 2021; Kishore *et al.,* 2013; Caetano *et al.,* 2019), bubaline heat stressed mammary explants and mammary epithelial cells (Sodhi *et al.,* 2013), porcine gastrointestinal tissues and liver (Ryan *et al.,* 2010), and bovine PBMCs for studies related to vector-borne haemoparasitic diseases (Kaur *et al.,* 2023). Recently, *B2M* was reported to be the most reliable RG for gene expression studies involving lamb testis cells under sheep pox virus infection, since its expression was observed to be most stable in infected and control groups (Sonowal *et al.,* 2022). Both *ACTB* and *B2M* have been identified as the most stable reference genes in whole blood of Gyr, Brahman, and Romosinuano cattle breeds (Lozano-Villegas *et al.,* 2021) and PBMCs of Ladakhi donkey and double humped camel (Tiwari *et al.,* 2022).

An important criterion for a reference gene is constitutive expression in all cell types under different biological conditions. We observed high variability in the expression of some genes, namely *RPS15, RPS9,* and *PKG1* in the 10 ovine tissues investigated in this study. Despite reports

of constitutive and stable expression of *RPS15* and *RPS9* in bovine blood and mammary gland (Kishore *et al.,* 2013, Tiwari *et al.,* 2022), our investigation does not recommend their use as normalizers in different ovine tissues since they exhibited maximum inter-sample variation. *PGK1* has been previously documented to be an unreliable internal control gene in rat skeletal muscle owing to highest expression variability and lowest stability (Wang *et al.,* 2019). Unstable expression of *PGK1* across different investigational paradigms in choroid plexus of rat brain has also been reported, rendering it unsuitable for qPCR analysis (Ho and Patrizi, 2021).

Our study screened 18 potential RGs (*ACTB, BACH1*, *B2M, GAPDH, HMBS, HPRT1*, *PGK1*, *PPIA, PPIB*, *RPLP0, RPL19, RPS9, RPS15, RPS28, SDHA*, *TBP*, *UXT*, and *YWHAZ)* to identify stably expressed normalizer genes in 10 different sheep tissues (muscle, skin, kidney, liver, intestine, rumen, lung, testis, heart, and spleen) so that they can be used to achieve reproducibility and accuracy in qPCR results. Four well-accepted stability testing algorithms: geNorm, NormFinder, BestKeeper, and ΔCt were employed for this purpose. Good agreement with respect to the ranking of genes was observed between the four algorithms since 8 out of the top 10 genes were common. Ultimately, a comprehensive ranking of the genes was achieved using RefFinder, which is a statistical platform that integrates the output of the abovementioned algorithms. *ACTB*, *PPIB, BACH1* and *B2M* were the most stable RGs, whereas, *RPS9, RPS15*, and *PGK1* were the least preferred RGs in our investigation.

Put together, this study identified genes and molecular pathways associated with fiber quality, coat color and adaptation to high altitude in Changthangi sheep by comparison of the global gene expression profiles with Muzzafarnagri sheep from the plains of India.Additionally, we recommend a panel of robust RGs comprising of *ACTB, PPIB, BACH1* and *B2M* that showed stable expression in 10 ovine tissues to be used as normalizers in experimental designs involving these tissues.

CHAPTER–6

Summary and Conclusions

SUMMARY AND CONCLUSION

- The skin transcriptomes of Changthangi and Muzzafarnagri sheep were compared to identify differentially expressed genes and molecular pathways contributing to variation in fiber quality and coat color.
- The top 25 genes most abundantly expressed genes in both breeds were involved in translation, collagen fibril organization, ribosome, ribonucleoprotein complex, and RNA binding.
- A total of 149 genes showed significantly higher expression and 2139 genes exhibited significantly lower expression in Changthangi sheep as compared to Muzzafarnagri sheep.
- The enriched KEGG pathways for the up-regulated genes in Changthangi sheep were tyrosine metabolism, metabolic pathways, and retinol metabolism.
- Melanosome and melanin pigment synthesis were the most enriched GO terms for the up-regulated dataset in Changthangi sheep.
- Eleven genes associated with fiber quality and coat color (*TYRP1, MLANA, TYR, SLC45A2, PMEL, DCT, OCA2, KRTAP6, KRTAP7, KRTAP13,* and *KRTAP21*) showed significantly higher expression in Changthangi sheep.
- The enriched GO terms for the down-regulated genes included 181 terms for biological processes, 98 terms for cellular components, and 104 terms for molecular functions. Signal transduction, cell differentiation, and intracellular signal transduction were enriched in the BP category, cytoplasm, cytosol, and nucleoplasm in the CC category and ATP binding, metal ion binding, and identical protein binding in the MF category
- 22,581 ncRNAs were commonly found in both Changthangi and Muzzafarnagri, 12,174 ncRNAs were uniquely observed in Changthangi, and 19,122 ncRNAs were uniquely observed in Muzzafarnagri sheep.
- Comprehensive assessment of 18 potential reference genes (*ACTB, BACH1, B2M, GAPDH, HMBS, HPRT1, PGK1, PPIA, PPIB, RPLP0, RPL19, RPS9, RPS15, RPS28, SDHA, TBP, UXT,* and *YWHAZ)* was carried out in 10 ovine tissues (muscle, skin, kidney, liver, intestine, rumen, lung, testis, heart, and spleen) from 5 sheep to identify stably expressed genes across different tissues.
- The ranking of genes was determined using four algorithms, namely, geNorm, NormFinder BestKeeper, and Delta Ct (ΔCt). Similar trend in ranking was obtained

using these approaches. Finally, comprehensive ranking of the potential RGs was achieved using RefFinder, which integrates the output obtained from different algorithms.

- A panel of reference genes, namely *ACTB, PPIB, BACH1* and *B2M* were identified as the most stable reference genes for genes expression studies in diverse ovine tissues including skin.
- RNA Seq results were validated for 10 selected genes (*TYR, TYRP1, DCT, OCA2, PMEL, SLC45A2, KRTAP6, KRTAP7, KRTAP13,* and *KRT19*) by qPCR using *ACTB* and *PPIB* as reference genes.

Conclusion

- Genes associated with pigmentation contribute to the coat color of Changthangi sheep and may also be involved in adaptation to the high-altitude environment. These represent interesting candidates for future research in regulating adaptive mechanisms.
- This study recommends a panel of robust reference genes comprising of *ACTB*, *PPIB*, *BACH1* and *B2M* that showed stable expression across to be used as normalizers in qPCR experiments involving diverse tissues, namely muscle, skin, kidney, liver, intestine, rumen, lung, testis, heart, and spleen.

Bibliography

BIBLIOGRAPHY

20th Livestock Census. BAHS-basic animal husbandry statistics: Department of Animal Husbandry, Dairying & Fisheries, Ministry of Agriculture, Government of India. Available online from www.dahd.nic.in (2019).

Ahlawat, S., Arora, R., Sharma, R., Sharma, U., Kaur, M., Kumar, A., Singh, K.V., Singh, M.K. and Vijh, R.K., 2020. Skin transcriptome profiling of Changthangi goats highlights the relevance of genes involved in Pashmina production. *Scientific reports*, *10*(1), p.6050.

Ahlawat, S., Arora, R., Sharma, U., Sharma, A., Girdhar, Y., Sharma, R., Kumar, A. and Vijh, R.K., 2021. Comparative gene expression profiling of milk somatic cells of Sahiwal cattle and Murrah buffaloes. *Gene*, *764*, p.145101.

Ahmad, N. and Mukhtar, H., 2004. Cytochrome p450: a target for drug development for skin diseases. *Journal of investigative dermatology*, *123*(3), pp.417-425.

Andersen, C.L., Jensen, J.L. and Orntoft, T.F., 2004. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer research*, *64*(15), pp.5245-5250.

Anisimov, S.V., 2008. Serial Analysis of Gene Expression (SAGE): 13 years of application in research. *Current pharmaceutical biotechnology*, *9*(5), pp.338-350.

Arora, R., Fairoze, M.N., Kaur, M., Sharma, A., Girdhar, Y., Devatkal, S.K., Ahlawat, S. and Vijh, R.K., 2019. Transcriptome profiling of longissimus thoracis muscles identifies highly connected differentially expressed genes in meat type sheep of India. *PLoS One*, *14*(6), p.e0217461.

Arora, R., Sharma, R., Ahlawat, S., Chhabra, P., Kumar, A., Kaur, M., Vijh, R.K., Lal, S.B., Mishra, D.C., Farooqi, M.S. and Srivastava, S., 2023. Transcriptomics reveals key genes responsible for functional diversity in pectoralis major muscles of native black Kadaknath and broiler chicken. *3 Biotech*, *13*(7), pp.1-8.

Arora, R., Siddaraju, N.K., Manjunatha, S.S., Sudarshan, S., Fairoze, M.N., Kumar, A., Chhabra, P., Kaur, M., Sreesujatha, R.M., Ahlawat, S. and Vijh, R.K., 2021. Muscle transcriptome provides the first insight into the dynamics of gene expression with progression of age in sheep. *Scientific Reports*, *11*(1), p.22360.

Awgulewitsch, A., 2003. Hox in hair growth and development. *Naturwissenschaften*, *90*, pp.193-211.

Bai, T., Liang, B., Zhao, Y., Han, J., Pu, Y., Wang, C., Ma, Y. and Jiang, L., 2021. Transcriptome analysis reveals candidate genes regulating the skin and hair diversity of Xinji Fine-Wool sheep and Tan sheep. *Agriculture*, *12*(1), p.15.

Berihulay, H., Abied, A., He, X., Jiang, L. and Ma, Y., 2019. Adaptation mechanisms of small ruminants to environmental heat stress. *Animals*, *9*(3), p.75.

Bhat, B., Singh, A., Iqbal, Z., Kaushik, J.K., Rao, A.R., Ahmad, S.M., Bhat, H., Ayaz, A., Sheikh, F.D., Kalra, S. and Shanaz, S., 2019. Comparative transcriptome analysis reveals the genetic basis of coat color variation in Pashmina goat. *Scientific Reports*, *9*(1), p.6361.

Bumgarner, R., 2013. Overview of DNA microarrays: types, applications, and their future. *Current protocols in molecular biology*, *101*(1), pp.22-1.

Bustin, S.A., 2002. INVITED REVIEW Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems. *Journal of molecular endocrinology*, *29*, pp.23-39.

Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L. and Vandesompele, J., 2009. The MIQE Guidelines: M inimum Information for Publication of Q uantitative Real-Time PCR E xperiments.

Caetano, L.C., Miranda-Furtado, C.L., Batista, L.A., Pitangui-Molina, C.P., Higa, T.T., Padovan, C.C. and Rosa-e-Silva, A.C.J.D.S., 2019. Validation of reference genes for gene expression studies in bovine oocytes and cumulus cells derived from in vitro maturation. *Animal Reproduction*, *16*, pp.290-296.

Chan, P.P. and Lowe, T.M., 2016. GtRNAdb 2.0: an expanded database of transfer RNA genes identified in complete and draft genomes. *Nucleic acids research*, *44*(D1), pp.D184-D189.

Chin, C.H., Chen, S.H., Wu, H.H., Ho, C.W., Ko, M.T. and Lin, C.Y., 2014. cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC systems biology*, *8*(4), pp.1-7.

Choi, Y.S., Zhang, Y., Xu, M., Yang, Y., Ito, M., Peng, T., Cui, Z., Nagy, A., Hadjantonakis, A.K., Lang, R.A. and Cotsarelis, G., 2013. Distinct functions for Wnt/β-catenin in hair follicle stem cell proliferation and survival and interfollicular epidermal homeostasis. *Cell stem cell*, *13*(6), pp.720-733.

Coelho, T.C., Chalfun-Junior, A., Barreto, H.G., Duarte, M.D.S., Garcia, B.D.O., Teixeira, P.D., Gionbelli, T.R.S. and Ladeira, M.M., 2022. Reference gene selection for quantitative PCR in liver, skeletal muscle, and jejunum of Bos indicus cattle. *Revista Brasileira de Zootecnia*, *51*.

Cuvertino, S., Stuart, H.M., Chandler, K.E., Roberts, N.A., Armstrong, R., Bernardini, L., Bhaskar, S., Callewaert, B., Clayton-Smith, J., Davalillo, C.H. and Deshpande, C., 2017. ACTB loss-of-function mutations result in a pleiotropic developmental disorder. *The American Journal of Human Genetics*, *101*(6), pp.1021-1033.

D'Mello, S.A., Finlay, G.J., Baguley, B.C. and Askarian-Amiri, M.E., 2016. Signaling pathways in melanogenesis. *International journal of molecular sciences*, *17*(7), p.1144.

d'Ischia, M., Wakamatsu, K., Napolitano, A., Briganti, S., Garcia- Borron, J.C., Kovacs, D., Meredith, P., Pezzella, A., Picardo, M., Sarna, T. and Simon, J.D., 2013. Melanins and melanogenesis: methods, standards, protocols. *Pigment cell & melanoma research*, *26*(5), pp.616-633.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T.R., 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, *29*(1), pp.15-21.

Dydensborg, A.B., Herring, E., Auclair, J., Tremblay, E. and Beaulieu, J.F., 2006. Normalizing genes for quantitative RT-PCR in differentiating human intestinal

epithelial cells and adenocarcinomas of the colon. *American Journal of Physiology-Gastrointestinal and Liver Physiology*, *290*(5), pp.G1067-G1074.

Fan, R., Xie, J., Bai, J., Wang, H., Tian, X., Bai, R., Jia, X., Yang, L., Song, Y., Herrid, M. and Gao, W., 2013. Skin transcriptome profiles associated with coat color in sheep. *BMC genomics*, *14*, pp.1-12.

Farhadian, M., Rafat, S.A., Panahi, B. and Ebrahimie, E., 2022. Transcriptome signature of two lactation stages in Ghezel sheep identifies using RNA-Sequencing. *Animal Biotechnology*, *33*(2), pp.223-233.

Fiers, W., Contreras, R., Duerinck, F., Haegeman, G., Iserentant, D., Merregaert, J., Min Jou, W., Molemans, F., Raeymaekers, A., Van den Berghe, A. and Volckaert, G., 1976. Complete nucleotide sequence of bacteriophage MS2 RNA: primary and secondary structure of the replicase gene. *Nature*, *260*(5551), pp.500-507.

French, M.C., Littlejohn, R.P., Greer, G.J., Bain, W.E., McEwan, J.C. and Tisdall, D.J., 2006. Growth hormone and ghrelin receptor genes are differentially expressed between genetically lean and fat selection lines of sheep. *Journal of animal science*, *84*(2), pp.324-331.

Garcia-Crespo, D., Juste, R.A. and Hurtado, A., 2005. Selection of ovine housekeeping genes for normalisation by real-time RT-PCR; analysis of PrPgene expression and genetic susceptibility to scrapie. *BMC veterinary research*, *1*(1), pp.1-8.

Garcıa-Vallejo, J.J., Van Het Hof, B., Robben, J., Van Wijk, J.A.E., Van Die, I., Joziasse, D.H. and Van Dijk, W., 2004. Approach for defining endogenous reference genes in gene expression experiments. *Analytical biochemistry*, *329*(2), pp.293-299.

Gill, R.W. and Sanseau, P., 2000. Rapid in silico cloning of genes using expressed sequence tags (ESTs).

Glazar, P., Papavasileiou, P. and Rajewsky, N., 2014. circBase: a database for circular RNAs. *Rna*, *20*(11), pp.1666-1670.

Gong, H., Zhou, H., Forrest, R.H., Li, S., Wang, J., Dyer, J.M., Luo, Y. and Hickford, J.G., 2016. Wool keratin-associated protein genes in sheep—a review. *Genes*, *7*(6), p.24.

Gothel, S.F. and Marahiel, M.A., 1999. Peptidyl-prolyl cis-trans isomerases, a superfamily of ubiquitous folding catalysts. *Cellular and molecular life sciences CMLS*, *55*, pp.423-436.

Hamalainen, H.K., Tubman, J.C., Vikman, S., Kyrölä, T., Ylikoski, E., Warrington, J.A. and Lahesmaa, R., 2001. Identification and validation of endogenous reference genes for expression profiling of T helper cell differentiation by quantitative real-time RT-PCR. *Analytical biochemistry*, *299*(1), pp.63-70.

Harrow, J., Frankish, A., Gonzalez, J.M., Tapanari, E., Diekhans, M., Kokocinski, F., Aken, B.L., Barrell, D., Zadissa, A., Searle, S. and Barnes, I., 2012. GENCODE: the reference human genome annotation for The ENCODE Project. *Genome research*, *22*(9), pp.1760-1774.

Hatey, F., Tosser-Klopp, G., Clouscard-Martinato, C., Mulsant, P. and Gasser, F., 1998. Expressed sequence tags for genes: a review. *Genetics Selection Evolution*, *30*(6), pp.521-541.

HEARING, V.J., 2000. The melanosome: the perfect model for cellular responses to the environment. *Pigment Cell Research*, *13*, pp.23-34.

Hieronymus, K., Dorschner, B., Schulze, F., Vora, N.L., Parker, J.S., Winkler, J.L., Rösen-Wolff, A. and Winkler, S., 2021. Validation of reference genes for whole blood gene expression analysis in cord blood of preterm and full-term neonates and peripheral blood of healthy adults. *BMC genomics*, *22*(1), pp.1-12.

Ho, K.H. and Patrizi, A., 2021. Assessment of common housekeeping genes as reference for gene expression studies using RT-qPCR in mouse choroid plexus. *Scientific Reports*, *11*(1), p.3278.

Hruz, T., Wyss, M., Docquier, M., Pfaffl, M.W., Masanetz, S., Borghi, L., Verbrugghe, P., Kalaydjieva, L., Bleuler, S., Laule, O. and Descombes, P., 2011. RefGenes: identification of reliable and condition specific reference genes for RT-qPCR data normalization. *BMC genomics*, *12*, pp.1-14.

Huang, D.W., Sherman, B.T. and Lempicki, R.A., 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols*, *4*(1), pp.44-57.

Huggett, J., Dheda, K., Bustin, S. and Zumla, A., 2005. Real-time RT-PCR normalisation; strategies and considerations. *Genes & Immunity*, *6*(4), pp.279-284.

Hutchison III, C.A., 2007. DNA sequencing: bench to bedside and beyond. *Nucleic acids research*, *35*(18), pp.6227-6237.

Jain, M., Nijhawan, A., Tyagi, A.K. and Khurana, J.P., 2006. Validation of housekeeping genes as internal control for studying gene expression in rice by quantitative real-time PCR. *Biochemical and biophysical research communications*, *345*(2), pp.646-651.

Jassal, B., Matthews, L., Viteri, G., Gong, C., Lorente, P., Fabregat, A., Sidiropoulos, K., Cook, J., Gillespie, M., Haw, R. and Loney, F., 2020. The reactome pathway knowledgebase. *Nucleic acids research*, *48*(D1), pp.D498-D503.

Jatav, P., Sodhi, M., Sharma, A., Mann, S., Kishore, A., Shandilya, U.K., Mohanty, A.K., Kataria, R.S., Yadav, P., Verma, P. and Kumar, S., 2016. Identification of internal control genes in milk- derived mammary epithelial cells during lactation cycle of I ndian zebu cow. *Animal Science Journal*, *87*(3), pp.344-353.

Jave-Suarez, L.F., Winter, H., Langbein, L., Rogers, M.A. and Schweizer, J., 2002. HOXC13 is involved in the regulation of human hair keratin gene expression. *Journal of Biological Chemistry*, *277*(5), pp.3718-3726.

Jiang, X., Xue, Y., Zhou, H., Li, S., Zhang, Z., Hou, R., Ding, Y. and Hu, K., 2015. Evaluation of reference gene suitability for quantitative expression analysis by quantitative polymerase chain reaction in the mandibular condyle of sheep. *Molecular Medicine Reports*, *12*(4), pp.5633-5640.

Jin, M., Wang, L., Li, S., Xing, M.X. and Zhang, X., 2011. Characterization and expression analysis of KAP7. 1, KAP8. 2 gene in Liaoning new-breeding cashmere goat hair follicle. *Molecular biology reports*, *38*, pp.3023-3028.

Kalds, P., Zhou, S., Gao, Y., Cai, B., Huang, S., Chen, Y. and Wang, X., 2022. Genetics of the phenotypic evolution in sheep: a molecular look at diversity-driving genes. *Genetics Selection Evolution*, *54*(1), pp.1-27.

Kaur, M., Kumar, A., Siddaraju, N.K., Fairoze, M.N., Chhabra, P., Ahlawat, S., Vijh, R.K., Yadav, A. and Arora, R., 2020. Differential expression of miRNAs in

skeletal muscles of Indian sheep with diverse carcass and muscle traits. *Scientific Reports*, *10*(1), p.16332.

Kaur, R., Ahlawat, S., Choudhary, V., Kumari, A., Kumar, A., Kaur, M., Arora, R., Sharma, R. and Vijh, R.K., 2023. Validation of stable reference genes in peripheral blood mononuclear cells for expression studies involving vector-borne haemoparasitic diseases in bovines. *Ticks and Tick-borne Diseases*, *14*(4), p.102168.

Kchouk, M., Gibrat, J.F. and Elloumi, M., 2017. Generations of sequencing technologies: from first to next generation. *Biology and Medicine*, *9*(3).

Kishore, A., Sodhi, M., Khate, K., Kapila, N., Kumari, P. and Mukesh, M., 2013. Selection of stable reference genes in heat stressed peripheral blood mononuclear cells of tropically adapted Indian cattle and buffaloes. *Molecular and cellular probes*, *27*(3-4), pp.140-144.

Knaust, J., Weikard, R., Albrecht, E., Brunner, R.M., Günther, J. and Kühn, C., 2020. Indication of premelanosome protein (PMEL) expression outside of pigmented bovine skin suggests functions beyond eumelanogenesis. *Genes*, *11*(7), p.788.

Kozomara, A. and Griffiths-Jones, S., 2010. miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic acids research*, *39*(suppl_1), pp.D152-D157.

Kukurba, K.R. and Montgomery, S.B., 2015. RNA sequencing and analysis. *Cold Spring Harbor Protocols*, *2015*(11), pp.pdb-top084970.

Kumar, A., Kaur, M., Ahlawat, S., Sharma, U., Singh, M.K., Singh, K.V., Chhabra, P., Vijh, R.K., Yadav, A. and Arora, R., 2021. Transcriptomic diversity in longissimus thoracis muscles of Barbari and Changthangi goat breeds of India. *Genomics*, *113*(4), pp.1639-1646.

Lathe, W., Williams, J., Mangan, M., and Karolchik, D. (2008). Genomic data resources: challenges and promises. Nat. Educ. 1:2.

Li, L., Li, D., Liu, L., Li, S., Feng, Y., Peng, X. and Gong, Y., 2015. Endothelin receptor B2 (EDNRB2) gene is associated with spot plumage pattern in domestic ducks (Anas platyrhynchos). *PloS one*, *10*(5), p.e0125883.

Li, S., Zhou, H., Gong, H., Zhao, F., Wang, J., Liu, X., Hu, J., Luo, Y. and Hickford, J.G., 2019. Identification of the ovine keratin-associated protein 21-1 gene and its association with variation in wool traits. *Animals*, *9*(7), p.450.

Lin, J.Y. and Fisher, D.E., 2007. Melanocyte biology and skin pigmentation. *Nature*, *445*(7130), pp.843-850.

Liu, G., Liu, R., Tang, X., Cao, J., Zhao, S. and Yu, M., 2015. Expression profiling reveals genes involved in the regulation of wool follicle bulb regression and regeneration in sheep. *International Journal of Molecular Sciences*, *16*(5), pp.9152-9166.

Livak, K.J. and Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2− ΔΔCT method. *methods*, *25*(4), pp.402-408.

Lozano, E.S. and Barragan, I.S.R., 2022. Reference gene stability in the Longissimus thoracis et lumborum muscle of Colombian Creole sheep. *Semina: Ciências Agrárias*, *43*(3), pp.987-1006.

Lozano-Villegas, K.J., Rodríguez-Hernández, R., Herrera-Sánchez, M.P., Uribe-García, H.F., Naranjo-Gómez, J.S., Otero-Arroyo, R.J. and Rondón-Barragán, I.S., 2021. Identification of reference genes for expression studies in the whole-blood from three cattle breeds under two states of livestock weather safety. *Animals*, *11*(11), p.3073.

Lyahyai, J., Serrano, C., Ranera, B., Badiola, J.J., Zaragoza, P. and Martin-Burriel, I., 2009. Effect of scrapie on the stability of housekeeping genes. *Animal biotechnology*, *21*(1), pp.1-13.

Ma, Z., Guo, W., Guo, X., Wang, X. and Kang, L., 2011. Modulation of behavioral phase changes of the migratory locust by the catecholamine metabolic pathway. *Proceedings of the National Academy of Sciences*, *108*(10), pp.3882-3887.

Mahakapuge, T.A.N., Scheerlinck, J.P., Rojas, C.A., Every, A.L. and Hagen, J., 2016. Assessment of reference genes for reliable analysis of gene transcription by RT-qPCR in ovine leukocytes. *Veterinary Immunology and Immunopathology*, *171*, pp.1-6.

Malik, A.A., Khan, H.M., Sofi, A.H., Mir, M.S., Farooq, J., Sheikh, F.D., Mir, A.Q. and Abdullah, M., 2021. Wool characteristics of Changthangi sheep. *Small Ruminant Research*, *199*, p.106364.

Mandal, A., Singh, L.B. and Rout, P.K., 2000. The Muzaffarnagari sheep, a mutton breed in India. *Animal Genetic Resources/Resources genetiques animales/Recursos genéticos animales*, *28*, pp.19-25.

Matsuzaki, Y., Umemoto, T., Tanaka, Y., Okano, T. and Yamato, M., 2015. β2-Microglobulin is an appropriate reference gene for RT-PCR-based gene expression analysis of hematopoietic stem cells. *Regenerative Therapy*, *1*, pp.91-97.

Michailidou, S., Gelasakis, A., Banos, G., Arsenos, G. and Argiriou, A., 2021. Comparative transcriptome analysis of milk somatic cells during lactation between two intensively reared dairy sheep breeds. *Frontiers in Genetics*, *12*, p.700489.

Morioka, M.S., Kawaji, H., Nishiyori-Sueki, H., Murata, M., Kojima-Ishiyama, M., Carninci, P. and Itoh, M., 2020. Cap analysis of gene expression (CAGE): a quantitative and genome-wide assay of transcription start sites. In *Bioinformatics for Cancer Immunotherapy: Methods and Protocols* (pp. 277-301). New York, NY: Springer US.

Nygard, A.B., Jørgensen, C.B., Cirera, S. and Fredholm, M., 2007. Selection of reference genes for gene expression studies in pig tissues using SYBR green qPCR. *BMC molecular biology*, *8*(1), pp.1-6.

Pachot, A., Blond, J.L., Mougin, B. and Miossec, P., 2004. Peptidylpropyl isomerase B (PPIB): a suitable reference gene for mRNA quantification in peripheral whole blood. *Journal of biotechnology*, *114*(1-2), pp.121-124.

Parkinson, J. and Blaxter, M., 2009. Expressed sequence tags: an overview. *Expressed Sequence Tags (ESTs) Generation and Analysis*, pp.1-12.

Parreira, J.R. and de Sousa Araújo, S., 2018. Studying the animal transcriptome: state of the art and challenges in the context of animal and veterinary sciences. *Proteomics in Domestic Animals: from Farm to Systems Biology*, pp.421-446.

Passmore, M., Nataatmadja, M. and Fraser, J.F., 2009. Selection of reference genes for normalisation of real-time RT-PCR in brain-stem death injury in Ovis aries. *BMC molecular biology*, *10*, pp.1-8.

Peletto, S., Bertuzzi, S., Campanella, C., Modesto, P., Maniaci, M.G., Bellino, C., Ariello, D., Quasso, A., Caramelli, M. and Acutis, P.L., 2011. Evaluation of internal reference genes for quantitative expression analysis by real-time PCR in ovine whole blood. *International journal of molecular sciences*, *12*(11), pp.7732-7747.

Perez, R., Tupac-Yupanqui, I. and Dunner, S., 2008. Evaluation of suitable reference genes for gene expression studies in bovine muscular tissue. *BMC molecular biology*, *9*(1), pp.1-6.

Peters, D.G., O'Hare, E.H., Ferrell, R.E., Kassam, A.B., Yonas, H. and Brufsky, A.M., 1999. Comprehensive transcript analysis in small quantities of mRNA by SAGE-lite. *Nucleic acids research*, *27*(24), pp.e39-e44.

Pfaffl, M.W., Tichopad, A., Prgomet, C. and Neuvians, T.P., 2004. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper–Excel-based tool using pair-wise correlations. *Biotechnology letters*, *26*, pp.509-515.

Pruett, N.D., Tkatchenko, T.V., Jave-Suarez, L., Jacobs, D.F., Potter, C.S., Tkatchenko, A.V., Schweizer, J. and Awgulewitsch, A., 2004. Krtap16, characterization of a new hair keratin-associated protein (KAP) gene complex on mouse chromosome 16 and evidence for regulation by Hoxc13. *Journal of Biological Chemistry*, *279*(49), pp.51524-51533.

Pruett, N.D., Tkatchenko, T.V., Jave-Suarez, L., Jacobs, D.F., Potter, C.S., Tkatchenko, A.V., Schweizer, J. and Awgulewitsch, A., 2004. Krtap16, characterization of a new hair keratin-associated protein (KAP) gene complex on mouse chromosome 16 and evidence for regulation by Hoxc13. *Journal of Biological Chemistry*, *279*(49), pp.51524-51533.

Puech, C., Dedieu, L., Chantal, I. and Rodrigues, V., 2015. Design and evaluation of a unique SYBR Green real-time RT-PCR assay for quantification of five major cytokines in cattle, sheep and goats. *BMC veterinary research*, *11*(1), pp.1-14.

Qi, Y., Fu, S., He, X., Wang, B., Da, L., Te, R., Yuejun, M., Suzhen, S., Zhang, W. and Liu, Y., 2021. Preliminary comparison of skin transcriptome from sheep with different wool fibre diameters. *Animal Production Science*, *61*(7), pp.708-714.

Robinson, T.L., Sutherland, I.A. and Sutherland, J., 2007. Validation of candidate bovine reference genes for use with real-time PCR. *Veterinary immunology and immunopathology*, *115*(1-2), pp.160-165.

Rosenkranz, D., 2016. piRNA cluster database: a web resource for piRNA producing loci. *Nucleic acids research*, *44*(D1), pp.D223-D230.

Rout, M., Panigrahi, S., Pradhan, S., Routray, A. and ranjan Swain, B., 2018. Genetic basis of fecundity in sheep-a review. *The Pharma Innovation*, *7*(4), pp.314-316.

Ryan, M.T., Collins, C.B., O'Doherty, J.V. and Sweeney, T., 2010. Selection of stable reference genes for quantitative real-time PCR in porcine gastrointestinal tissues. *Livestock Science*, *133*(1-3), pp.42-44.

Sahu, A.R., Wani, S.A., Saxena, S., Rajak, K.K., Chaudhary, D., Sahoo, A.P., Khanduri, A., Pandey, A., Mondal, P., Malla, W.A. and Khan, R.I.N., 2018. Selection and validation of suitable reference genes for qPCR gene expression analysis in goats and sheep under Peste des petits ruminants virus (PPRV), lineage IV infection. *Scientific reports*, *8*(1), p.15969.

Sai Lakshmi, S. and Agrawal, S., 2008. piRNABank: a web resource on classified and clustered Piwi-interacting RNAs. *Nucleic acids research*, *36*(suppl_1), pp.D173-D177.

Sanger, F. and Coulson, A.R., 1975. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of molecular biology*, *94*(3), pp.441-448.

Sanger, F., Coulson, A., Barrell, B.G., Smith, A.J.H. and Roe, B.A., 1980. Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. *Journal of molecular biology*, *143*(2), pp.161-178.

Schmittgen, T.D. and Livak, K.J., 2008. Analyzing real-time PCR data by the comparative CT method. *Nature protocols*, *3*(6), pp.1101-1108.

Schmittgen, T.D. and Zakrajsek, B.A., 2000. Effect of experimental treatment on housekeeping gene expression: validation by real-time, quantitative RT-PCR. *Journal of biochemical and biophysical methods*, *46*(1-2), pp.69-81.

Schulze, F., Malhan, D., El Khassawna, T., Heiss, C., Seckinger, A., Hose, D. and Rösen-Wolff, A., 2017. A tissue-based approach to selection of reference genes for quantitative real-time PCR in a sheep osteoporosis model. *BMC genomics*, *18*(1), pp.1-9.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B. and Ideker, T., 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research*, *13*(11), pp.2498-2504.

Sharma, R., Ahlawat, S., Sharma, H., Sharma, P., Panchal, P., Arora, R. and Tantia, M.S., 2020. Microsatellite and mitochondrial DNA analyses unveil the genetic structure of native sheep breeds from three major agro-ecological regions of India. *Scientific reports*, *10*(1), p.20422.

Shendure, J., Balasubramanian, S., Church, G.M., Gilbert, W., Rogers, J., Schloss, J.A. and Waterston, R.H., 2017. DNA sequencing at 40: past, present and future. *Nature*, *550*(7676), pp.345-353.

Shi, R., Li, S., Liu, P., Zhang, S., Wu, Z., Wu, T., Gong, S. and Wan, Y., 2022. Identification of key genes and signaling pathways related to Hetian sheep wool density by RNA-seq technology. *PLoS One*, *17*(5), p.e0265989.

Shi, X., Wu, J., Lang, X., Wang, C., Bai, Y., Riley, D.G., Liu, L. and Ma, X., 2021. Comparative transcriptome and histological analyses provide insights into the skin pigmentation in Minxian black fur sheep (Ovis aries). *PeerJ*, *9*, p.e11122.

Shih, B.B., Farrar, M.D., Vail, A., Allan, D., Chao, M.R., Hu, C.W., Jones, G.D., Cooke, M.S. and Rhodes, L.E., 2020. Influence of skin melanisation and ultraviolet radiation on biomarkers of systemic oxidative stress. *Free Radical Biology and Medicine*, *160*, pp.40-46.

Shiraki, T., Kondo, S., Katayama, S., Waki, K., Kasukawa, T., Kawaji, H., Kodzius, R., Watahiki, A., Nakamura, M., Arakawa, T. and Fukuda, S., 2003. Cap analysis gene

expression for high-throughput analysis of transcriptional starting point and identification of promoter usage. *Proceedings of the National Academy of Sciences*, *100*(26), pp.15776-15781.

Silver, N., Best, S., Jiang, J. and Thein, S.L., 2006. Selection of housekeeping genes for gene expression studies in human reticulocytes using real-time PCR. *BMC molecular biology*, *7*(1), pp.1-9.

Singh, G.K., Chatterjee, M., Grewal, R.S. and Verma, R., 2013. Incidence and care of environmental dermatoses in the high-altitude region of Ladakh, India. *Indian Journal of Dermatology*, *58*(2), p.107.

Sodhi, M., Kishore, A., Khate, K., Kapila, N., Mishra, B.P., Kataria, R.S., Mohanty, A.K., Varshney, N. and Mukesh, M., 2013. Evaluating suitable internal control genes for transcriptional studies in heat- stressed mammary explants of buffaloes. *Journal of Animal Breeding and Genetics*, *130*(2), pp.106-117.

SOLANO, F., MARTINEZ- ESPARZA, M.A.R.Í.A., JIMENEZ- CERVANTES, C.E.L.I.A., HILL, S.P., LOZANO, J.A. and GARCÍA- BORRÓN, J.C., 2000. New insights on the structure of the mouse silver locus and on the function of the silver protein. *Pigment Cell Research*, *13*, pp.118-124.

Sonowal, J., Patel, C.L., Dev, K., Singh, R., Barkathullah, N., Malla, W.A., Gandham, R.K., Agarwal, R.K., Kumar, D., Saxena, S. and Kalaiselvan, E., 2022. Selection and validation of suitable reference gene for qPCR gene expression analysis in lamb testis cells under sheep pox virus infection. *Gene*, *831*, p.146561.

Strain, G.M., 2015. The genetics of deafness in domestic animals. *Frontiers in Veterinary Science*, *2*, p.29.

Szklarczyk, D., Morris, J.H., Cook, H., Kuhn, M., Wyder, S., Simonovic, M., Santos, A., Doncheva, N.T., Roth, A., Bork, P. and Jensen, L.J., 2016. The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. *Nucleic acids research*, p.gkw937.

Thellin, O., Zorzi, W., Lakaye, B., De Borman, B., Coumans, B., Hennen, G., Grisar, T., Igout, A. and Heinen, E., 1999. Housekeeping genes as internal standards: use and limits. *Journal of biotechnology*, *75*(2-3), pp.291-295.

Tian, Y., Du, J., Yang, X., Zeng, W., He, J., Zhao, B., Fu, X., Xu, X., Wu, W., Di, J. and Huang, X., 2022. Sheep IGFBP2 and IGFBP4 promoter methylation regulates gene expression and hair follicle development. *Electronic Journal of Biotechnology*, *59*, pp.46-54.

Tiwari, M., Sodhi, M., Verma, P., Vivek, P., Kataria, R.S., Niranjan, S.K., Bharti, V.K., Masharing, N., Gujar, G., Chanda, D. and Mukesh, M., 2022. Selection of species specific panel of reference genes in peripheral blood mononuclear cells of native livestock species adapted to trans-Himalayan region of Leh-Ladakh. *Scientific Reports*, *12*(1), p.18473.

Tuteja, R. and Tuteja, N., 2004. Serial analysis of gene expression: applications in human studies. *BioMed Research International*, *2004*, pp.113-120.

Ullah, F., Jamal, S.M., Ekegbu, U.J., Haruna, I.L., Zhou, H. and Hickford, J.G., 2020. Polymorphism in the ovine keratin-associated protein gene KRTAP7-1 and its association with wool characteristics. *Journal of animal science*, *98*(1), p.skz381.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M. and Rozen, S.G., 2012. Primer3—new capabilities and interfaces. *Nucleic acids research*, *40*(15), pp.e115-e115.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A. and Speleman, F., 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome biology*, *3*(7), pp.1-12.

Verma, P., Sharma, A., Sodhi, M., Tiwari, M., Vivek, P., Kataria, R.S., Nirajan, S.K., Bharti, V.K., Singh, P., Lathwal, S.S. and Sharma, V., 2022. Identification of internal reference genes in peripheral blood mononuclear cells of cattle populations adapted to hot arid normoxia and cold arid hypoxia environments. *Frontiers in Genetics*, *12*, p.730599.

Vorachek, W.R., Hugejiletu, Bobe, G. and Hall, J.A., 2013. Reference gene selection for quantitative PCR studies in sheep neutrophils. *International journal of molecular sciences*, *14*(6), pp.11484-11495.

Wan, Z., Yang, H., Cai, Y., Ma, J., Cheng, P., Wang, Z., Wang, F. and Zhang, Y., 2022. Comparative transcriptomic analysis of Hu sheep pituitary gland prolificacy at the follicular and luteal phases. *Genes*, *13*(3), p.440.

Wang, C., Li, H., Guo, Y., Huang, J., Sun, Y., Min, J., Wang, J., Fang, X., Zhao, Z., Wang, S. and Zhang, Y., 2020. Donkey genomes provide new insights into domestication and selection for coat color. *Nature communications*, *11*(1), p.6014.

Wang, H., Xue, L., Li, Y., Zhao, B., Chen, T., Liu, Y., Chang, L. and Wang, J., 2016. Distribution and expression of SLC45A2 in the skin of sheep with different coat colors. *Folia Histochemica et Cytobiologica*, *54*(3), pp.143-150.

Wang, L., Chen, X., Song, T., Zhang, X., Zhan, S., Cao, J., Zhong, T., Guo, J., Li, L., Zhang, H. and Wang, Y., 2020. Using RNA-Seq to identify reference genes of the transition from brown to white adipose tissue in goats. *Animals*, *10*(9), p.1626.

Wang, X., Zhao, H., Ni, J., Pan, J., Hua, H. and Wang, Y., 2019. Identification of suitable reference genes for gene expression studies in rat skeletal muscle following sciatic nerve crush injury. *Molecular Medicine Reports*, *19*(5), pp.4377-4387.

Wang, Z., Gerstein, M. and Snyder, M., 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nature reviews genetics*, *10*(1), pp.57-63.

Watson, J.D. and Crick, F.H., 1953. Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature*, *171*(4356), pp.737-738.

Wei, C., Wang, H., Liu, G., Wu, M., Cao, J., Liu, Z., Liu, R., Zhao, F., Zhang, L., Lu, J. and Liu, C., 2015. Genome-wide analysis reveals population structure and selection in Chinese indigenous sheep breeds. *BMC genomics*, *16*(1), pp.1-12.

Wijesena, H.R. and Schmutz, S.M., 2015. A missense mutation in SLC45A2 is associated with albinism in several small long haired dog breeds. *Journal of Heredity*, *106*(3), pp.285-288.

Xu, Q., Liu, X., Chao, Z., Wang, K., Wang, J., Tang, Q., Luo, Y., Zheng, J., Tan, S. and Fang, M., 2019. Transcriptomic analysis of coding genes and Non-Coding RNAs reveals complex regulatory networks underlying the black back and white belly coat phenotype in Chinese Wuzhishan pigs. *Genes*, *10*(3), p.201.

Yamaguchi, Y., Brenner, M. and Hearing, V.J., 2007. The regulation of skin pigmentation. *Journal of biological chemistry*, *282*(38), pp.27557-27561.

Yamamoto, M., Wakatsuki, T., Hada, A. and Ryo, A., 2001. Use of serial analysis of gene expression (SAGE) technology. *Journal of immunological methods*, *250*(1-2), pp.45-66.

Yang, Q., Yin, J., Li, G., Qi, L., Yang, F., Wang, R. and Li, G., 2014. Reference gene selection for qRT-PCR in Caragana korshinskii Kom. under different stress conditions. *Molecular biology reports*, *41*, pp.2325-2334.

Yang, Z., Bai, C., Pu, Y., Kong, Q., Guo, Y., Ouzhuluobu, Gengdeng, Liu, X., Zhao, Q., Qiu, Z. and Zheng, W., 2022. Genetic adaptation of skin pigmentation in highland Tibetans. *Proceedings of the National Academy of Sciences*, *119*(40), p.e2200421119.

Yao, L., Bao, A., Hong, W., Hou, C., Zhang, Z., Liang, X. and Aniwashi, J., 2019. Transcriptome profiling analysis reveals key genes of different coat color in sheep skin. *PeerJ*, *7*, p.e8077.

Yoo, S.M., Choi, J.H., Lee, S.Y. and Yoo, N.C., 2009. DOI QR Code Applications of DNA Microarray included Disease Diagnostics. *Journal of microbiology and biotechnology*, *19*(7), pp.635-646.

Yu, S., Liao, J., Tang, M., Wang, Y., Wei, X., Mao, L., Zeng, C. and Wang, G., 2017. A functional single nucleotide polymorphism in the tyrosinase gene promoter affects skin color and transcription activity in the black-boned chicken. *Poultry Science*, *96*(11), pp.4061-4067.

Yu, Z., Jiang, K., Xu, Z., Huang, H., Qian, N., Lu, Z., Chen, D., Di, R., Yuan, T., Du, Z. and Xie, W., 2018. Hoxc-dependent mesenchymal niche heterogeneity drives regional hair follicle regeneration. *Cell Stem Cell*, *23*(4), pp.487-500.

Zang, R., Bai, J., Xu, H., Zhang, L., Yang, J., Yang, L., Lu, J. and Wu, J., 2011. Selection of suitable reference genes for real-time quantitative PCR studies in Lanzhou fat-tailed sheep (Ovis aries). *Asian Journal of Animal and Veterinary Advances*, *6*(8), pp.789-804.

Zhang, L., Sun, F., Jin, H., Dalrymple, B.P., Cao, Y., Wei, T., Vuocolo, T., Zhang, M., Piao, Q. and Ingham, A.B., 2017. A comparison of transcriptomic patterns measured in the skin of Chinese fine and coarse wool sheep breeds. *Scientific reports*, *7*(1), p.14301.

Zhang, P., Si, X., Skogerbo, G., Wang, J., Cui, D., Li, Y., Sun, X., Liu, L., Sun, B., Chen, R. and He, S., 2014. piRBase: a web resource assisting piRNA functional study. *Database*, *2014*, p.bau110.

Zhang, X., Guo, J., Wei, X., Niu, C., Jia, M., Li, Q. and Meng, D., 2018. Bach1: function, regulation, and involvement in disease. *Oxidative medicine and cellular longevity*, *2018*.

Zhang, Y., Zhang, X.D., Liu, X., Li, Y.S., Ding, J.P., Zhang, X.R. and Zhang, Y.H., 2013. Reference gene screening for analyzing gene expression across goat tissue. *Asian-Australasian journal of animal sciences*, *26*(12), p.1665.

Zhao, H., Guo, T., Lu, Z., Liu, J., Zhu, S., Qiao, G., Han, M., Yuan, C., Wang, T., Li, F. and Zhang, Y., 2021. Genome-wide association studies detects candidate genes for wool traits by re-sequencing in Chinese fine-wool sheep. *BMC genomics*, *22*, pp.1-13.

Zhu, W., Lin, Y., Liao, H. and Wang, Y., 2015. Selection of reference genes for gene expression studies related to intramuscular fat deposition in Capra hircus skeletal muscle. *PloS one*, *10*(3), p.e0121280.

Appendices

APPENDIX- I

Chemicals, Enzymes, Molecular markers and Primers

S. No.	Chemicals, enzymes, weight markers, kits and primers	Company
1	1% PBS	Invitrogen
2	100% Ethanol	Omnis
3	50 bp	Fermentas
4	6x Loading Dye	Fermentas
5	Agarose	Sigma Aldrich
6	Chloroform	Sigma
7	DEPC (Diethyl polycarbonate)	Sigma Aldrich
8	dNTP mix	Thermo Scientific
9	DreamTaq MM	ThermoScientific
10	EtBr (Ethidium bromide)	Sigma Aldrich
11	First Strand Synthesis Reaction Buffer	New England Biolabs
12	High sensitivity DNA 1000 chip	Agilent Technologies
13	NEBNext Ultra RNA Library Prep Kit	New England Biolabs
14	Primers	Sigma
15	QuantiTech SYBR Green kit	QIAGEN
16	RevertAid First Strand cDNA Synthesis kit	Thermo Scientific
17	RevertAid M-MuL V RT	ThermoScientific
18	RiboLock RNase inhibitor	ThermoScientific
19	RNA ladder	Ambion
20	RNALater	Hi-media
21	RNase free water	ThermoScientific
22	RNaseZap	Ambion
23	RPE buffer	QIAGEN
24	RT buffer	Thermo Scientific
25	RT enzyme	Thermo Scientific
26	Second Strand Synthesis Reaction Buffer	New England Biolabs
27	Tris Base	Hi-media
28	TRIzol reagent	Ambion

APPENDIX- II

Common buffers, media and reagents used

1. Agarose gel (0. 5%, 0.8%, 1.0% and 2.0%)

Agarose gels (0.5, 0.8, 1.0 and 2.0%) were prepared, according to assay need and microwave oven was used to dissolve the gel. The gel was freshly prepared each time.

2. EDTA, 0.5 M (8.0)

Dissolve 18.6g of the sodium salt EDTA in 80ml of distilled water. Stir vigorously on a magnetic stirrer. Adjust the pH as with NaOH pellets. Adjust the final volume to 100ml with distilled water. Sterilize by autoclaving. This compound will not dissolve completelyuntil pH reaches near 8.0.

3. Ethidium bromide (EtBr)

Dissolve 1g of ethidium bromide in 100ml of water. Stir on a magnetic stirrer fora several hours to ensure that the dye has dissolved. Wrap the container with aluminum foiland store at 4°C.

4. Ethanol (70%, 75%, 80%)

Dissolve 70 ml, 75 ml and 80ml of absolute ethanol into 30 ml, 25 ml and 20ml of distilled water respectively and mix well.

5. Gel loading dye (6x)

Bromophenol blue 0.25 g 0.25%

Xylene cynol0.125 g 0.125%

Ficoll (Type 400) 15 g 15%

Distilled water up to 100 ml

Store in 10 ml aliquots at room temperature or preferably at 4°C.

6. Tris Acetic EDTA (TAE) buffer (50x)

Tris-base: 121g

Glacial acetic acid: 28.5ml

EDTA: 18.6g

Add double distilled autoclaved water to make volume up to 500 ml.

7. Tris Acetic EDTA (TAE) buffer (1x)

Dilute 20ml of Tris Acetic EDTA (50X) in 980ml nuclease free water and make the volume 1000 ml of Tris Acetic EDTA buffer (1x).

8. Tris - HCl, 1 M (pH - 7.4, 8.0 or 9.5)

Dissolve 12.1 g of Trizma - HCL in 80 ml of distilled water. Adjust the pH with NaOHand make the final volume up to 100 ml with distilled water and sterilize by autoclaving.

9. Tris - HCl, 10mM (pH - 7.4, 8.0 or 9.5)

Dilute 0.1ml Tris-HCl (1M) in 99.9ml nuclease free water and make the volume 100ml of10mM Tris- HCl.

10. DEPC treated water (0.1%)

Add 1ml DEPC in 1 liter distilled water and incubate overnight at 37°. On the next day,autoclave for 45 min. at 15 psi pressure at 120°C temperature

APPENDIX- III

List of Laboratory Equipment

S. No.	Equipment's	Brand name, Firm
1	Autoclave	Indfos
2	Bench top centrifuge	REMI
3	Bioanalyzer	Bio-Rad
4	Deep freezer (-20 °C)	Vest frost
5	Electronic Balance	Precisa
6	Gel Documentation System	UVP
7	Homogenizer	Glas-Col
8	Horizontal Electrophoresis Unit	Bio-Rad
9	Ice flake machine	Spectrum
10	Laminar air flow	ESCO
11	Micro centrifuge	Eppendorf
2	Micropipettes	Eppendorf, Brand
13	Microwave oven	KenStar
14	Mini cooler	Bangalore Genie
15	Nanodrop	ThermoScientific ND1000
16	pH meter	TIMPL
17	Qubit 3.0 fluorometer	ThermoScientific
18	Real-time PCR	QuantStudio 5
19	Micro centrifuge	DLAB
20	Refrigerated plate centrifuge	Eppendorf
21	Thermal cycler	Eppendorf
22	Ultra-low deep freezer (- 80 °C)	Haier
23	Vortex Shaker	REMI
24	Water bath	York Scientific
25	Water purifier	Aqua Max