Muscle transcriptome provides the first insight into the dynamics of gene expression with progression of age in sheep

Reena Arora1, Naveen Kumar Siddaraju2, S. S. Manjunatha2, S. Sudarshan2, Mohamed Nadeem Fairoze2, Ashish Kumar1, Pooja Chhabra1, Mandeep Kaur1, R. M. Sreesujatha2, Sonika Ahlawat1 & Ramesh Kumar Vijh1

The dynamic synergy of genes and pathways in muscles in relation to age affects the muscle characteristics. Investigating the temporal changes in gene expression will help illustrate the molecular mechanisms underlying muscle development. Here we report the gene expression changes in skeletal muscles through successive age groups in Bandur, a meat type sheep of India. RNA sequencing data was generated from the longissimus thoracis muscles from four age groups, ranging from lamb to adult. Analysis of 20 highest expressed genes common across the groups revealed muscle protein, phosphorylation, acetylation, metal binding and transport as significant functions. Maximum differentiation was observed after 2.5–3 years on transition from lambs to adult. Transcriptional regulation by the TFAP2 transcription factors, IL-6 signaling and PI3K/AKT signaling pathways were enriched in younger animals. The gene-protein network demarcated key interactive genes involved in muscle development and proliferation that can be used as candidates for future research on improvement of muscle characteristics.

Livestock sector plays an important role in the rural economy of India. Sheep and goats contribute substantially to the livelihood of the small and marginal farmers. Small ruminants provide 21.89% of the total meat produced, of which 8.36% is from sheep1. Meat quality is an amalgamation of the effects of nutrition, environment and genetic capability. Factors like growth and development of skeletal muscle directly affect the quality and quantity of meat. The palatability of meat is influenced by tenderness and fat content, which are in turn affected by nutrition, muscle characteristics, post mortem events, genetics and the age of the animal2. Several studies have evaluated the impact of age on the tenderness and fatty acid profile of muscles in various species3,4. The dynamic synergy of genes and pathways in muscles in relation to age of the animal has also been explored in several species5–7. In sheep however, there is dearth of related information on lambs and mature animals.

Deep sequencing techniques have significantly broadened our knowledge of global gene expression patterns and regulatory mechanisms in several tissues. Transcriptomic studies have identified regulatory factors involved in muscle growth and meat quality in livestock species8,9. The genetic mechanisms underlying growth and development of skeletal muscles of sheep have been investigated using next generation sequencing techniques5. Differentially expressed genes for contrasting characteristics have been detected in small as well as large ruminants8,10. Such efforts have led to identification of biomarkers for meat quality11.

Consumer preferences for Bandur sheep in India led to the investigation that established tenderness and higher backfat content in muscles of Bandur sheep as compared to local sheep found in the same area12. Our previous studies have attempted to understand the regulatory molecular mechanisms underlying the muscle traits in Bandur sheep breed in comparison to the local sheep12,13. These studies have identified the genes and pathways that may be associated with the muscling traits. Previous studies on goat14 and pigs15 have provided some insight on the temporal pattern of expression of genes through developmental stages. However, lack of information on the effect of age on the expression of genes in skeletal muscles of sheep has prompted this investigation. The aim of the present study was to compare the gene expression pattern in skeletal muscles through successive age

1Animal Biotechnology Division, ICAR-National Bureau of Animal Genetic Resources, G T Road By-Pass, P O Box 129, Karnal 132001, Haryana, India. 2Karnataka Veterinary Animal and Fisheries Sciences University, Bangalore 560024, India. *email: rejagati@gmail.com
groups in Bandur sheep, a meat type breed of India. The temporal changes in gene expression will help illustrate the molecular mechanisms underlying muscle development in sheep.

**Results**

Quality control and filtering of raw data resulted in an average of 93,434,064 reads for each library. The processed reads were aligned to the *Ovis aries* reference genome Oar_v4.0 (SAMN00116405). An overall mapping of 94.6% was observed across samples of all age groups (Table S1) while 79–90% of the reads were uniquely mapped. Comparison of the transcript profile revealed 85.4% transcripts common to all groups. The unique transcripts in each group accounted for only 1–1.8% (Fig. 1). The total number of known genes discovered with a minimum threshold of RPKM > 0.01, was 11045, 11168, 10907 and 10514 in group 1, 2, 3 and 4, respectively (Table S2).

**Abundantly expressed genes and their functions.** Twenty genes common across all four groups, with expression >2000 RPKM are shown in Fig. 2. Analysis of the 20 highest expressed genes common across the groups revealed muscle protein, phosphorylation, acetylation, metal binding and transport as significant (p ≤ 0.05) functions. The major biological processes associated with these genes included skeletal muscle contraction, regulation of muscle contraction, muscle tissue morphogenesis, transition between fast and slow fiber, etc. Significant cellular component were troponin and myosin complex, myofibril actin filament and muscle thin filament tropomyosin while actin filament binding, structural constituent of muscle, calcium ion binding were the major molecular functions common in all age groups (Fig. 3). Among these genes, the expression of ACTA1, COX3, TPM2, MYL2, CTYB2, MYH7, ND4 and TNNC1 decreased with progression of age. On the other hand, expression of TNNI2, TPM1 and EN03 genes increased with age. Enriched pathways associated with these...
genes were muscle contraction, glucose metabolism, respiratory electron transport and creatine metabolism (Table S3).

The expression profile of some genes previously associated with meat quality in livestock species like HSPB1 and HSPB8 conspicuously increased after 3 years of age. Expression of DNAJB5 showed a mild increase whereas that of HSPA6 decreased with age. Transcription of CAPN3 also exhibited a very slight increase with advancement of age in investigated animals. Other genes like FABP3, CAPN1 and CAST did not express prominent difference across the groups. Similarly, collagen genes are also known to influence meat quality. Although the expression of some collagen genes, namely, COL15A1, COL23A1, COL4A2, COL6A2, COL1A1, COL4A1, COL1A2 and COL3A1 was observed to be quite low (≤ 7.3 RPKM), their expression was highest in young animals and decreased considerably with age.

Differentially expressed genes and enriched pathways. Differential expression of genes was analyzed across all combinations of the four age groups. Considerably larger number of genes was observed to be differentially expressed between Group 1 versus 3 and Group 1 versus 4 only, which were therefore used for further analysis (Fig. 4). There were 301 significantly differentially expressed genes in Group 1 versus 3 (p_adj < 0.05; FDR < 0.05), followed by 141 in Group 1 versus 4. The up- and down-regulated genes observed in Group 1 versus 3 were 222 and 79, respectively, whereas 98 up-regulated and 43 down-regulated genes were identified in Group 1 versus 4 (Table S4). Gene ontology analysis revealed negative regulation of cell proliferation, response to hypoxia, positive regulation of GTPase activity, negative regulation of cysteine-type endopeptidase activity involved in apoptotic process, calcium ion binding etc., as significant functions (p_adj < 0.05) associated with differentially expressed genes in Group 1 versus 3 (Fig. 5a). Significant pathways associated with the up-regulated genes were transcriptional regulation by the TFAP2 (AP-2) family of transcription factors, IL-6 signaling, PI3K/AKT signaling, ECM proteoglycans, muscle contraction etc. (Table S5), while down-regulated genes were related with transcriptional regulation by RUNX1 and RUNX3. Similarly, enriched functions in Group 1 versus 4 included regulation of transcription from RNA polymerase II promoter, myoblast fusion, regulation of epithelial to mesenchymal transition and protein phosphatase regulator activity (Fig. 5b). The top canonical pathways identified again were transcriptional regulation by the TFAP2 (AP-2) family of transcription factors, IL-6 signaling, PI3K/AKT signaling. Other important pathways included MAPK1 and MAPK3 activation, DAG and IP3 signaling and PKA activation (Table S6).

Validation of RNAseq data by RT-qPCR. Differentially expression of genes in Group 1 versus Group 3 and Group 1 versus Group 4 was validated by RT-qPCR. The magnitude of relative expression of the selected genes namely ACTA1, FABP3, COL1A1, COL1A2, HSPA6, HSPB1, IL17A, MYOD, RUNX1 and SOCS3 was observed to be similar to that derived from RNAseq data (Figs. 6 and 7).
Gene—protein interactions. To identify highly connected genes from the set of differentially expressed ones, a gene—protein network was constructed for each comparative Group 1 versus 3 and Group 1 versus 4 (Figs. 8 and 9). The central genes with 5 or more interactions identified for Group 1 versus 3 included COL1A2, SOCS3, IL17A, COL1A1, IFIT1, NR4A1, RUNX1, CDKL2, CBLB and VIPR1. Whereas ESR1, CCNE1, MYOD1, HSPB1, COL1A1, NR4A2, TWIST1, ISG15, PER1 and TLE3 were the most interactive hub genes identified in the network for Group 1 versus 4. Among these genes COL1A2, IL17A, COL1A1, NR4A1, VIPR1, ESR1, CCNE1, TWIST1 and ISG15 were up-regulated, while the remaining were down-regulated. COL1A1 was common in both the networks. Some other important genes identified in the network were HSPA6, LPAR1, EIF5A and PPARGC1A.
Figure 6. Relative expression of differentially expressed genes between Group 1 and Group 3 by quantitative PCR. Group 1 = 2 tooth stage or 1 year; Group 3 = 6 tooth stage or 2.5–3 years.

Figure 7. Relative expression of differentially expressed genes between Group 1 and Group 4 by quantitative PCR. Group 1 = 2 tooth stage or 1 year; Group 4 = 8 tooth stage or 3.5–4 years.
Muscle traits relevant to meat quality are modulated by age, nutrition, genetics and environmental factors which in turn affect the structure and composition of muscle fibres as well as intramuscular fat. Differential expression of genes in muscles between different breeds in relation to feed, carcass and muscle properties has been examined by several studies\(^8\)–\(^10,12\). It is well established that muscle characteristics are affected by age\(^16\). However, investigating the simultaneous change in gene expression with age in skeletal muscles of sheep has not garnered much interest.

Figure 8. Highly connected differentially expressed genes identified by comparison of age Groups 1 and 3. The top 10 highly connected genes have been highlighted (red: up-regulated; green: down-regulated). Group 1 = 2 tooth stage or 1 years; Group 3 = 6 tooth stage or 2.5–3 years.

Figure 9. Highly connected differentially expressed genes identified by comparison of age Groups 1 and 4. The top 10 highly connected genes have been highlighted (red: up-regulated; green: down-regulated). Group 1 = 2 tooth stage or 1 years; Group 4 = 8 tooth stage or 3.5–4 years.

Discussion
Muscle traits relevant to meat quality are modulated by age, nutrition, genetics and environmental factors which in turn affect the structure and composition of muscle fibres as well as intramuscular fat. Differential expression of genes in muscles between different breeds in relation to feed, carcass and muscle properties has been examined by several studies\(^8\)–\(^10,12\). It is well established that muscle characteristics are affected by age\(^16\). However, investigating the simultaneous change in gene expression with age in skeletal muscles of sheep has not garnered much interest.
The present study therefore, explored the pattern of gene expression with progression of age in skeletal muscles of Bandur sheep. The sarcomere proteins, mainly alpha-actinins, myosin heavy chains and Z disc form the basic units of muscle fibres and are coded by ACTA1, ACTN2, ACTN3, MYH4, MYH7 and MYOZ genes. The troponin complex modulates the interaction between actin and myosin during muscle contraction. TNNT2 and TNNC2 encode for their respective troponin isoforms and are known to play a key role in muscle composition and aging. The genes mainly involved in muscle contraction and muscle fibre composition (ACTA1, ACTN2, ACTN3, ANKR1, TCP1, NRAP, MYH4, MYH8, TNNT2, TNNC2, MYOZ1, LDB3, TPM1) were abundantly expressed irrespective of age group in our study. Earlier investigations have revealed no significant effect of sex on muscle fibre type, however, the fibre diameter was affected. Hence, further research on gene expression in longissimus thoracis muscles, across both rams and ewes is warranted. Besides structural components of muscles the highly expressed genes also represented glycogen metabolism and storage (LDHA, PYGM, PGAM2) as well as energy metabolism (COX1, COX3, CKM, ND4). The calpains are another family of genes associated with muscle tenderization. The calpain-calpastatin activity was reported to increase with age in beef cattle. It is well established that the activity of calpains is enhanced during aging. A slight increase in the expression of CAPN3 with age was also detected in Bandur sheep. The domination of expression of these genes in our study reinforces their fundamental role in muscle physiology.

Major pathways identified by differential analyses also had relevance to myogenesis. During muscle development, the transition from epithelial to mesenchymal cells is known to be regulated by TGFβ. Studies on mice have demonstrated that cytokines of the IL-6 family play a crucial role in regulating myogenesis. It is interesting to note that this activity is in turn dependent on MAPK and NFP signaling. Recent studies on Bandur sheep have also reported the enrichment of the PI3K-Akt and MAPK pathways that are associated with hypertrophy as well muscle differentiation. These pathways were enriched in the young animals.

As expected, lesser differentiation of gene expression age groups. Differential expression was prominent after 2.5–3 years in Bandur sheep reflecting transition from lamb to adult sheep. Major differences observed between lambs and adult sheep muscles were evident in the expression of genes associated with muscle structure, growth and lipid metabolism. The gene-protein network demarcated important interactive hub genes that may be involved in age related muscle development in sheep. The hub genes enriched in younger animals included COL1A2, COL1A1, TWIST1, NRAAI, VIPR1, CCNE1, ESR1, IL17A and ISG15. It is worth mentioning that these highly connected genes identified in our study are known to be involved in muscle growth and regeneration. Collagens form the major connective tissue that imparts structural strength to muscles. The crosslinking between the collagen molecules increases with age resulting in increased toughness of meat in adult animals. The expression of COL1A1 and COL1A2 was observed to be decreased in bovine fetuses that were myostatin deficient. A decrease of collagen expression in older animals was also observed in mice. Several studies have associated the significance of expression of collagen type 1 gene in muscle composition and regulation of intra muscular fat in pork. Age affects the structure and distribution of collagen in the extracellular matrix of skeletal muscle, which is also reflected in the change in expression of collagen precursor genes in skeletal muscles of Bandur sheep. TWIST1 is another gene encoding a transcription factor that is involved in the early growth of skeletal muscles but its function in adults is less explored. It is also known to inhibit DNA binding of MYOD1 during myotube formation. Other genes with higher expression in lambs that are associated with muscle growth include NRAAI and VIPR1. Experimental evidence using NRAAI knockout mice suggests its association with growth and proliferation of muscle cells. Studies on mouse myoblasts revealed NRAAI as an important factor regulating myoblast differentiation. VIPR1 is a G-protein coupled receptor expressed in the skeletal muscles that may be involved in muscle atrophy/hypertrophy. Further, experimental evidence supports the role of the estrogen receptor (ESR1) in muscle strength by modulation of myosin regulatory light chain in mice. The ESR1 gene regulates expression via the estrogen signaling pathway. The IL17A gene, on the other hand has been reported to activate PPARγ which is involved in increased adipogenesis. In addition, NRAA receptor family in conjunction with CRT2 and PPARy assists in myogenesis and adipogenesis via the CREB pathway. Thus, most of the highly connected genes identified in Bandur lamb muscles were associated with growth, proliferation and adipogenesis.

Notable hub genes in older or mature animals included MYOD1, RUNX1, TLE3, SOCS3, CBLB, HSPB1, IFIT1, CDLK2, NRAA2 and PER1. Myoblast differentiation is orchestrated by an assortment of muscle-specific regulatory factors. MYOD1, a myogenic regulatory factor and an important marker for myoblast development, facilitates myogenesis during embryogenesis. MYOD1 along with RUNX1 and TFAP1 transcription factors, has been associated with myoblast regeneration or proliferation. Although the function of MYOD1 in adult muscle is less investigated, there are some reports that have established its role in growth and proliferation in adult muscles. RUNX1 has been suggested to contribute to myotube generation in humans and its expression is activated by muscle damage. TLE is a family of transcriptional co-factors that are involved in cell differentiation. The transcriptional activity of MyoD has also been reported to be regulated by TLE3. Another important gene SOCS3, identified in this study has been linked to myogenic differentiation. Previous reports suggest the role of CBLB gene in skeletal muscle atrophy in mice by negative regulation of growth factors during cell differentiation and development.

The small heat shock proteins (Hsps) long associated with heat stress, are being increasingly implicated in cellular senescence and apoptosis. The HSPB1 is highly expressed in muscle tissues and codes for Hsps. The Hsps are anti-apoptotic and prevent the muscles from denaturation, which has implications in meat tenderness. Several studies have reported the involvement of HSPB1 in meat tenderness in cattle. Age related expression of genes associated with meat quality detected considerable increase in the expression of HSPB1 and HSPB8 in older age (3.5–4 years) in Bandur sheep. A similar increase in hspb1 protein was observed in the skeletal muscles of chicken. Although much information on HSPB8 and aging are not available, it is known to be involved in protein damage through translational arrest and autophagy. Previous studies on sheep have reported no...
significant effect of sex on the tenderness of *longissimus thoracis* muscles. Therefore, it may be speculated that a similar expression of hsp70 could be expected across both sexes. 

The role of some of the identified hub genes is not very well defined in skeletal muscles. These include the interferon stimulated genes *ISG15* and *IFIT1* that are suspected to be involved in muscular diseases. Although inadequate information is available on *CCNE1* (Cyclin E1) and Cyclin Dependent Kinase (CDK), little information is available on *CDK7*. Normalization of reads was done as reads per kilobase million (RPKM) and reads with RPKM values > 0.01 were included in the study. The CLC transcriptomics analysis tool was used for differential expression analysis. The reads were mapped against the ovine genome assembly (Oar_v4.0). The RNA sequencing data have been deposited in the NCBI Bio Project PRJNA416678 with accession numbers SAMN16191735-750. Normalization of reads was done as reads per kilobase million (RPKM) and reads with RPKM values > 0.01 were included in the study. The CLC transcriptomics analysis tool was used for differential expression analysis across the different age groups.

**Conclusion**

This study is an attempt to gain an insight into the dynamics of gene expression with progression of age in sheep skeletal muscles. Our results highlight the temporal changes in gene expression in skeletal muscles from lambs to adult sheep. Noteworthy demarcation between lambs and adult sheep was observed in the expression of genes associated with muscle growth, adipogenesis and apoptosis. The enrichment of PI3K-Akt and MAPK pathways in lambs suggests their relevance in muscle growth. This differential expression analysis will contribute towards understanding the genetic basis of physiological changes in muscles with age. Several highly connected genes identified in our study are known to be involved in muscle development and proliferation and may serve as candidates for future research in myogenesis.

**Materials and methods**

**Samples.** Healthy animals were selected in coordination with the sheep rearers and butchers from Mandya district, Karnataka. As the animals were selected from the field, they were considered healthy if they did not show any visible signs of infection or ailment and were inspected by a veterinary officer. Care was taken to select unrelated male animals from flocks reared under the same management and feed. Since no records of the animals were maintained, the age of the animals was determined by information provided by the rearers as well as the dentine pattern. A two-tooth stage confirmed the age about 12 months (Group 1), 4-tooth between 18 and 24 months (Group 2), 6-tooth between 30 and 36 months (Group 3) and 8-tooth between 42 and 48 months (Group 4). Four animals in each group were selected. The animals were slaughtered according to standard commercial 'halal' procedures.

**RNA sequencing.** The *longissimus thoracis* muscles were collected in RNLater (Thermo Fisher Scientific Baltics, UAB, Vilnius, Lithuania) solution. RNeasy kit (Qiagen, Hilden, Germany) was used for extraction and purification of total RNA. Samples with RIN value ≥ 7.0 (Agilent Bioanalyzer) were selected for further processing. Libraries of four biological replicates from each age group were prepared by TruSeq RNA Library Prep Kit v2 (Illumina, San Diego, CA, USA). Paired end sequencing (150 bp) was performed on Illumina HiSeq-2000 Platform.

**Data analysis.** FastQC (v 0.11.5) was used for assessment of the quality of samples. The raw reads were trimmed and filtered using FastXTooKit. CLS Genomics Workbench 6.5.1 (CLC Bio, Aarhus, Denmark) was used for data analysis. The reads were mapped against the ovine genome assembly (Oar_v4.0). The RNA sequencing data have been deposited in the NCBI Bio Project PRJNA416678 with accession numbers SAMN16191735-750. Normalization of reads was done as reads per kilobase million (RPKM) and reads with RPKM values > 0.01 were included in the study. The CLC transcriptomics analysis tool was used for differential expression analysis across the different age groups.

**Validation by real time quantitative PCR (RT qPCR).** Ten differentially expressed genes were selected for RT qPCR analysis (Table S7). Primers for selected genes were designed by Primer 3 software or taken from published sequences. cDNA was synthesized from 2 µg of purified RNA, using Super Script III Reverse Transcriptase (Thermo Fisher Scientific, Carlsbad, CA), as per manufacturer’s protocol. Each sample was analyzed in triplicate qPCR reactions. The final reaction volume of 10 µl contained 2 µl of cDNA, 8 µl of qPCR master mix (5 µl of SYBR Green Real-Time master mix (Applied Biosystems, Vilnius, Lithuania), 0.3 µl of each primer, 2.4 µl of DNA/RNA-free water. The samples were run on QuantStudio 5 Real-Time PCR System (Applied Biosystems). Standard curve calculation using four points of cDNA serial dilutions was used to estimate the PCR efficiency.

**Statistical analysis.** Differentially expressed genes with log2 fold change ≥ 2.0 and p value (padj) < 0.05 were used for further analyses. DAVID, Consensus Pathway Data Base and Reactome were utilized for the functional and pathways analysis. Cytoscape ver 3.6.0 along with cytoHubba app was utilized for gene-protein network analysis. For the RT qPCR the mean cycle threshold (Ct) values of the genes were normalized to geometric mean of the reference genes *PPIB* and *β-ACTN* along with the 2−ΔΔCT method was used for data analysis.

**Ethics approval.** Animal samples were purchased from local butchers. All ethical norms and guidelines were followed, with approval from Institutional Animal Ethics Committee, ICAR-National Bureau of Animal Genetic Resources, Karnal, Haryana, India (F.No. NBAGR/IAEC/2017, dated 21.01.2017).
Acknowledgements

This work was financially supported by ICAR-Consortium Research Platform-Genomics (Animal Science). We are grateful to Director, ICAR-National Bureau of Animal Genetic Resources (NBAGR), Karnal and Indian Council of Agricultural Research (ICAR), New Delhi for providing necessary facilities.
Author contributions
RA: conceptualization and supervision. MNF, NKS, SS, MSS and SRM: resources. AK, MK and SA: data curation. AK, PC, RKV and RA: formal analyses; RA: writing—review & editing. All authors reviewed the manuscript.

Funding
This work was financially supported by ICAR-Consortium Research Platform—Genomics (Animal Science).

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-01848-5.

Correspondence and requests for materials should be addressed to R.A.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2021