*Journal of Cell and Molecular Research (2024) 16 (1), 13-24* DOI: 10.22067/jcmr.2024.86067.1085

Research Article

# Differentially Expressed Genes and Related Pathways in the Ovine Muscle Transcriptome pre and postnatal Using the Integration of Microarray and RNA-seq Data

Fahime Mohammadi[1], Mojtaba Tahmoorespur[1], Ali Javadmanesh[1,2] *

[1] *Department of Animal Science, Faculty of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran*
[2] *Industrial Biotechnology Research Group, Research Institute of Biotechnology, Ferdowsi University of Mashhad, Mashhad, Iran*

Received 25 December 2023 Accepted 18 September 2024

## Abstract

The formation of muscle myofibrils as well as the growth and hypertrophy of the muscle are controlled by various genes. Bioinformatics tools could also be used to integrate and analyze heterogeneous data sets. In this study, the DEGs, gene network, GO and biological pathways have been investigated by incorporating the data obtained by expression array and RNA-Seq, related to the muscle tissue transcriptome of Texel sheep respectively before and after birth. The microarray expression profile was extracted from the GEO database, and the RNA-Seq expression profile was extracted from the ArrayExpress database. DEGs were identified with limma and sva software packages in R environment and a gene network was drawn with STRING, an application in Cytoscape software. The clustering and gene ontology were done with CytoCluster and ClueGO applications. The results showed a significant difference between the juvenile and 70-day embryonic stages the expression of 103 genes, between the adult and juvenile stages the expression of 28 genes and between the adult and 70-day embryonic stages the expression of 62 genes. By constructing the gene network between these DEGs, a total of 37 selected genes were identified. The results revealed the function of these genes in cell proliferation, protein synthesis, formation and organization of myofibrils, muscle contraction, and lipid metabolism. By integrating the expression data, this study provided a general view of the differences in transcriptomes in the muscle tissue of sheep. Also, the selected genes such as BUB1, RFC2, KIAA0101, RAD51, CKS2, and UQCRB were identified for the first time being reported as effective genes for myogenesis.

***Keywords*:** DEGs, Gen Ontology, Physiological pathway, Muscle tissue, Sheep

## Introduction

Skeletal muscle growth is a step-by-step and exponential process of differentiation, development, and maturation, which is regulated by molecular regulatory networks and cellular signaling pathways (Bassel-Duby and Olson, 2006). The investigation of gene expression is done with different methods, such as Northern Blot, Real-time PCR, expression arrays, Expressed Sequence Tag, and RNA sequencing (RNA-Seq) (Flintoft, 2008). So far, there have been many reports on the use of RNA-Seq and microarray techniques to evaluate the muscle tissue transcriptome in sheep (Betti et al., 2022b; Mohammadi et al., 2019; Zhang et al., 2013). Zhang *et al.*, 2013 used the RNA-Seq method for a comparative analysis of the muscle transcriptome in Dorper and Small-Tailed Han Sheep breeds using the Illumina platform. Finally, 40481 and 38851 coding single nucleotide polymorphisms (cSNPs) were identified in Han and Dorper breeds, respectively and additionally, 123678 novel transcript units were identified (Zhang et al., 2013). Also, the comparative transcriptome profile of the Longissimus muscle tissue between the Qianhua merino and small tail Han breeds was studied by Son *et al.*, 2016. The RNA-Seq results showed that 960 genes were differentially expressed of which, 463 genes were related to muscle growth and development and were involved in biological processes such as skeletal muscle tissue growth and muscle cell differentiation (Sun et al., 2016b). With the generation and accumulation of large sets of data, meta-analysis has become popular for combining the results of several studies to increase statistical power (Bakhshalizadeh et al., 2021 and 2024). Also, various studies have been published in which they have evaluated the transcriptome by using different

* Corresponding author's e-mail address: javadmanesh@um.ac.ir

bioinformatics methods in integrating RNA-seq and microarray data. In research by Ma et al. (2017), a hierarchical Bayesian model has been proposed for integrating data obtained from microarray and RNA-seq methods regarding breast cancer. The result indicated the improvement of detection capability and significant pathways in the identified biomarkers from the proposed Bayesian model. This finding was repeated in several experimental datasets and it was confirmed that merging two datasets improves the accuracy and detection capability (Ma et al., 2017). Another bioinformatics study was conducted by Castillo et al. (2017) to integrate data sets obtained through microarray with RNA-seq methods to identify breast cancer profiles. In this research, the microarray data set was analyzed through a pipeline in R, and the expression values were calculated. The RNA-Seq data set was also analyzed through the intended pipeline in the Linux and the expression values were calculated. Then the expression values of both sets were integrated into the R and DEGs were identified. The test results indicated the high accuracy of the integration method (Castillo et al., 2017). Zhang et al. (2017), also accomplished a meta-analysis study to investigate the carcinogenic risk of Bisphenol A (BPA), which is an endocrine-disrupting chemical and is associated with many diseases, including heart attacks and diabetes. In this research, the gene expression data sets of microarray and RNA-seq were integrated and after calculating the DEGs and analyzing the gene network and their ontology, finally 85 DEGs (68 genes with overexpression and 17 genes with downexpression) were detected (Zhan et al., 2018). The results of all these studies indicated the high accuracy of the integration method and improved accuracy and statistical strength of the results. Therefore, it can be stated that in the field of in silico studies (using different servers and software packages), it is possible to point out major genes that play key roles in the body's vital functions and pathways with greater accuracy. The present study is devoted to investigating the differentially expressed genes (DEGs) of the Longissimus Dorsi muscle of Texel sheep during the different stages of growth and development, using bioinformatics methods to integrate the data obtained from microarray and RNA-Seq methods related to the transcriptome of ovine skeletal muscle tissue. This integration and normalization of data was mainly to increase sample size and homogeneity of data.

## Materials and Methods

### Collecting and Analysis of Microarray Data

The gene expression microarray data of the Longissimus Dorsi muscle of Texel sheep at 70-days embryonic stage (E) with three repetitions were extracted from the GEO database with accession number GSE23563 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE23563). Quality control, correction, normalization and calculation of expression values, and integration of annotation file with the calculated expression value file to obtain the symbol genes were done in the R environment and using limma Biobase, GEOquery software package (Saedi et al., 2022).

### Collecting and Analysis of RNA-Seq Data

RNA-Seq data of muscle tissue from six samples with two replications in both juvenile (J) (aged 6 to 10 months) and adult (A) individuals was extracted from ArrayExpress database (https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-3838). Analysis of RNA-Seq data was done in the Ubuntu Linux environment. Quality control and trimming of RNA-Seq data were conducted by fastQC and Trimmomatic-0.36 software, respectively (Betti, et al., 2022a). Also, Mapping the reads, calculating the readings values and the expression values of RNA-Seq data was done using the HISAT2 2.2.0, Bowtie2, Samtools-1.9, HTSeq-0.6.1, and DESeq2 package, respectively (Castillo et al., 2017).

### Integration of Normalized Expression Values and DEGs Calculation in R

To integrate the expression values of microarray with RNA-seq data; limma and sva packages along with two data sets were called in R software and the aggregate command was executed for each of the data based on the approach described by Castillo et al (Castillo et al., 2017). To remove the non-biological effects, the batch effect was also removed. To calculate the DEGs, the samples were first grouped and then paired comparisons between age groups juvenile individuals vs 70-day embryonic stage (J vs E), adult individuals vs juvenile individuals (A vs J), and also adult individuals vs 70-day embryonic stage (A vs E). The Venn diagram tool was used to visualize common genes between three stages (Taheri et al., 2023).

### Analysis of the Gene Network and Ontology

To identify the relationship among the identified DEGs, the gene network was constructed using STRING 1.5.1, an application in Cytoscape 3.7.1 (http://apps.*cytoscape*.org/apps/stringapp) (Shannon et al., 2003). Next, due to the vastness of the gene

network, each network was clustered using MCODE 1.6.1 and CytoCluster 2.1.0 applications, and significant clusters (P-value < 0.05) were identified. Gene ontology analysis and biological pathways involved among the genes were investigated using the ClueGO 2.5.9 and CluePedia 1.5.9 applications in Cytoscape (Zeraatpisheh et al., 2023). Gene ontology enrichment analysis between DEGs of different stages was performed in three categories biological process (BP), molecular (MF) function, and cellular component (CC).

## Results and Discussion

### Calculation Expression Values of the Microarray Data

Data quality control and normalization were completed. To calculate expression values, the GPL10778 platform annotation file with Probe ID 1042520 was used. Only 9289 probe IDs were included in the result of this study and from those, 7918 gene symbols were identified.

### Quality Control, Mapping the Reads on the Reference Genome, and Calculating the Values of RNA-seq

Generally, the quality of all reads was fairly good, and the average Phred quality score of reads was above 25. The statistical information of the readings before and after the trimming process is reported in Table 1. To align and connect the reads, the reference genome of sheep (Ovis_aries.Oar_v3.1) and HISAT2 software (2.0.3-beta) were used. The results showed that more than 90% of the reads were mapped on the reference genome. In total, the number of Ensembl_Genes based on which the reading values were calculated by HTSeq software was 27056. After removing IDs with zero readings in all samples, 10855 IDs remained (Table 1).

### Normalization of Data and Calculation of Expressed Values of RNA-seq Data

Normalization and calculation of expressed values of RNA-seq data were done using the DESeq2 package. After merging them with the annotation file to obtain gene symbols from 10855 Ensembl ID, finally, 9417 common genes were identified among all individuals.

### Integration of Normalized Expressed Values of Microarray and RNA-seq Data

Both the expressed values of microarray and RNA-seq data, which were calculated in the previous steps, were integrated by R software to calculate the DEGs. Then, the batch effects (which were related to the different techniques of producing two sets of data) were removed and their graphs were plotted. The PCA plot before removing the batch effects, the boxplot, and the PCA plot of the merged data after removing the batch effects were plotted and have been shown in Figures 1, 2, and 3, respectively.

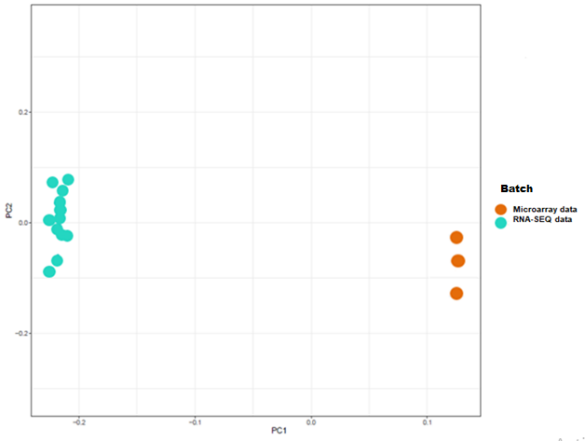


**Figure 1.** PCA plot of both microarray data (orange color) and RNA-seq data (green color) before removing batch effects.

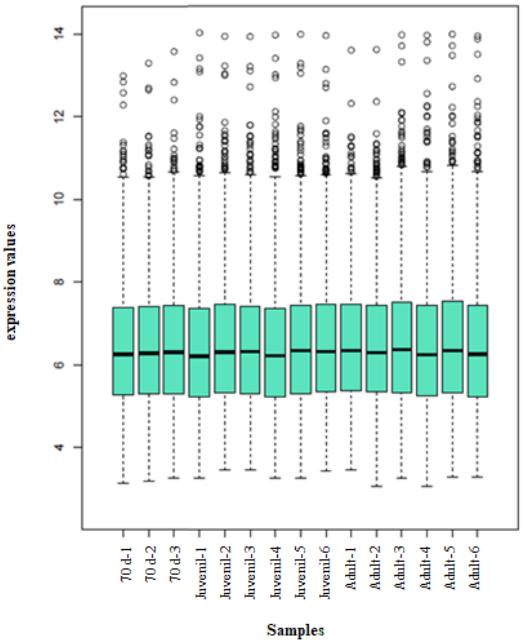


**Figure 2.** The box diagram of the distribution of gene expression values which is obtained from the microarray and RNA-seq methods after integration and removing the batch effects. The green areas include the first quartile, the middle quartile, and the third quartile.

**Table 1.** Statistical information of the reads (mapped on the reference genome and calculated by HTSeq)

| Samples | Total seq. of raw data | GC% | Total seq. of trimmed data | overall alignment rate% | Coverage | Ensembl Gene | Total counts |
|---|---|---|---|---|---|---|---|
| J[1]-F[2]-1 | 8647125 | 45& 45 | 6622025 | 83.14 | 34.92 | 35771 | 2973 |
| J-F-2 | 13368944 | 46 & 45 | 10132560 | 94.46 | 56.23 | 66540 | 4195 |
| J-F-3 | 13756065 | 46 & 45 | 10454415 | 94.38 | 62.52 | 68088 | 4247 |
| J-F-4 | 8685734 | 45& 45 | 6641245 | 84.72 | 35.93 | 37226 | 2995 |
| J-F-5 | 13374250 | 46 & 45 | 10156487 | 94.52 | 61.30 | 70882 | 4253 |
| J-F-6 | 13831915 | 46 & 45 | 10539372 | 94.49 | 63.17 | 69362 | 4257 |
| A[3]-F-1 | 19697023 | 46 & 46 | 15023092 | 94.04 | 79.39 | 5882 | 113446 |
| A-F-2 | 19792913 | 46 & 46 | 10350142 | 94.22 | 61.86 | 4981 | 78175 |
| A-M[4]-3 | 9016506 | 46 & 46 | 7220274 | 91.87 | 41.51 | 3937 | 63936 |
| A-M-4 | 5771315 | 45& 45 | 4633682 | 89.71 | 26.54 | 3077 | 39784 |
| A-M-5 | 9285314 | 46 & 46 | 7445816 | 91.85 | 43.09 | 3968 | 65933 |
| A-M-6 | 5852296 | 45& 45 | 4703164 | 89.99 | 27.03 | 3061 | 39651 |

[1] Juvenile

[2] Female

[3] Adult

[4] Male

## DEGs Calculation

The number of DEGs between J vs E stages was 103 (FDR- adj.P.Val < 0.05), of which 64 genes had overexpression (logFC > 1) and 39 genes had down expression (logFC < -1) (Table 2). In the comparison between A vs J individuals, 28 differential genes with FDR P-adj value (FDR-adj.P.Val) < 0.05) were identified, among which 11 genes were up- (logFC > 1) and 17 genes were downregulated (logFC < -1). Also, in the comparison between A vs E stages, 62 different genes (FDR-adj.P.Val < 0.05) were identified, among which 37 genes showed up-regulation (logFC > 1) and 25 genes showed down-regulation pattern (logFC < -1). However, no common differentially expressed gene was detected among these three age stages (Figure 4 and Table 3).

**Table 2.** The DEGs in pairwise comparison of age stags

| Time point | No. over-expressed gens | No. down expressed gens | Total |
|---|---|---|---|
| 70 d Embryonic - Juvenile | 64 | 39 | 103 |
| Juvenile -Adult | 11 | 17 | 28 |
| 70 d Embryonic - Adult | 37 | 25 | 62 |

**Table 3.** The Venn analysis of DEGs in pairwise comparison of age stags

| Time Point | No. gene | Genes |
|---|---|---|
| Juvenile – 70 d Embryonic | 58 | HMGCR, ARNT, MYBPH, CMBL, CKAP2L, PPM1K, KDM1A, SEC24D, LMO7, RDX, AKR1A1, VBP1, MYOZ1, EPB41L3, PGAM2, TRAF4, CD6, PON3, ASF1B, RCSD1, TRPS1, RAD51, MARCKSL1, RFC2, PCOLCE2, PGM1, BUB1, MYOT, NCAPG, CCND2, PHF19, AOX1, CA4, PPP1R3A, ACTN2, AMPD1, CA11, USP25, DDC, CHURC1, MUSTN1, ADAMTSL4, SQLE, MPP7, PLAT, ETFDH, GHR, CHMP4C, UBE2C, ABR, TMEM159, COX7A1, MOCS1, NR3C1, ALDH2, EEF1A2, TCF7, SCHIP1 |
| Adult - Juvenile | 12 | ZBTB47, ARL6IP5, HIPK3, AKAP9, RBM25, COQ5, LRPAP1, MKLN1, HMOX2, FUBP1, FAM175A, IPO5 |
| Adult – 70 d Embryonic | 19 | PIGU, PTPN14, AFF1, TTLL1, PRKAB2, ACAN, SOD2, PRKCZ, MFN2, CAV1, MTUS1, NAMPT, PHKG2, AKAP7, PLN, INTS10, PPP1R3C, C1orf21, CDK2AP1 |
| Adult – Juvenile | 9 | PPIB, ANKRD40, ZFAND5, XIRP2 DLD, CKM, PDK4, PDLIM3, PPP1R1A |
| Juvenile – 70 d Embryonic | | |
| Adult – 70 d Embryonic | 36 | SFRP2, CASQ1, ACTN3, VCAN, FGL2, UCP3, EMILIN2, SNAP25, MLLT11, MX1, HN1, VEGFA, CCNB1, NNAT, KIAA0101, TMOD4, AMOT, NOSTRIN, ALDH1A1, VSNL1, ACOT7, FBP2, KLF9, CA2, MBP, AGTR2, MFAP2, SLC2A4, AASS, MYH1, CKS2, ISYNA1, IGFBP2, ERBB3, SYBU, HSD11B1 |
| Juvenile – 70 d Embryonic | | |
| Adult – 70 d Embryonic | 7 | NFAT5, PECAM1, UQCRB, CTSL1, CTNNB1, ABI3BP, COX4I1 |
| Adult - Juvenile | | |

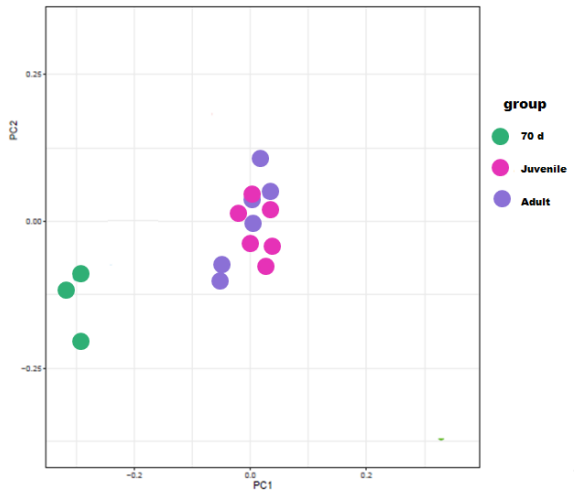


**Figure 3.** PCA plot of both microarray and RNA-seq data after integrating and removing batch effects. The data related to each age stage is marked with a separate color. A: Microarray data is shown in orange and RNA-seq data is shown in green. B: The green, pink, and purple colors represent the data of 70-day embryonic, juvenile, and Adult stages, respectively.

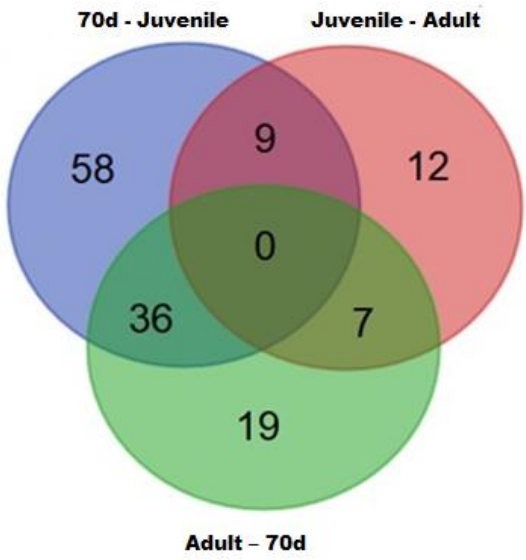


**Figure 4.** Venn diagram of pairwise comparison of DEGs identified among three age stages.

## Gene Network, Gene Ontology and Biological Pathways of DEGs

### Gene Network, Gene Ontology, and Biological Pathways Between DEGs of J vs E Stages

Among the 103 DEGs identified between J vs E stages, finally, the gene interactive network was plotted between 74 genes (Figure 5). Three significant gene clusters (P-value < 0.05) were also identified. The first significant cluster with a P-Value of 2.779E-5 had 11 nodes, including CCND2, UBE2C, BUB1, RAD51, CKS2, CCNB1, NCAPG, ASF1B, RFC2, KIAA0101 and CKAP2L genes, which all of them showed down expression (Figure 6). Various studies have shown that all these genes somehow play a role in the cell cycle, cell proliferation (Ewen et al., 1993; Rother et al., 2007), mitosis, and meiosis (Sartor et al., 1992), in regulating and modulating the structure of chromatin nucleosomes and DNA packaging (Geiman et al., 2004; Simpson et al., 2006). These pathways are related to mitosis and consequently to the growth and development of muscle tissue during embryonic, newborn, and mature individuals. Therefore, it is expected that the genes involved in the above pathways have higher expression during the embryonic period and the expression of these genes decreases during the postnatal period compared to the embryonic period. The second significant cluster with P-Value of 1.050E-4 had 11 nodes, including CKM, MYH1, AMPD1, ACTN3, CASQ1, MYOZ1, MYOT, TMOD4, ACTN2, XIRP2 and PDLIM3 genes, all of which had over expression. The main function of this cluster is mainly in skeletal and striated muscle function. Therefore, they play a role in connecting the actin filaments to sarcomere, muscle contraction (Ribeiro et al., 2014), cell energy homeostasis, ATP hydrolysis and converting chemical energy into mechanical ones (Vikenes et al., 2009), improving the nucleotide purine cycle, maintaining calcium concentration in the sarcoplasmic reticulum during muscle contraction (Terentyev et al., 2003), regulating the activity of Z-plane sarcomeres (Beggs et al., 1992; Salmikangas et al., 1999). The increased expression of these genes may be due to increased muscle activity or muscle contractions and hypertrophy, therefore, the over-expression of these genes in the juvenile stage is expected due to muscle tissue enlargement. Among the genes of this cluster, ACTN3 and MYOZ1 genes have already been reported as genes related to muscle hypertrophy (Gaeini, 2013). AMPD1, CASQ1, ACTN2, and PDLIM3 genes from this cluster have been identified and reported as differential genes in comparative transcriptome profiling of longissimus muscle tissues from Qianhua Mutton Merino and Small Tail Han sheep (Sun et al., 2016a). Also, The MYOT gene has been previously reported as a DEG in the Transcriptome analysis of skeletal muscle at prenatal stages in Polled Dorset versus Small-tailed Han sheep By microarray method (Liu et al., 2015). The third significant cluster with a P-value of 0.010. had 4 nodes included AOX1, ALDH2, ALDH1A1 and

AKR1A1 genes, all of which showed over expression. Genes that have been related to the third significant cluster involved in glucose metabolism (Lefrançois-Martinez et al., 2004), aldehyde dehydrogenase (Ohta et al., 2004), and adipogenesis (Sigruener et al., 2007).

Ontology enrichment analysis for DEGs between J vs E stages showed that 29 BP were enriched in a significant pathway (P-adj Bonf > 0.05), which mainly related to angiotensin pathway and its responses, cell development in striated muscle, myofibril assembly, sarcomere organization, calcineurin regulation, biosynthesis process of ribonucleotide biphosphate, pyruvate, and purine. Also, 10 MF were significantly enriched (P-adj Bonf > 0.05), which are mainly related to carbonate dehydratase activity, carbon-oxygen lyase activity, hydro-lyase activity, iron-sulfur cluster binding, and flavin adenine D binding nucleotide (FAD) and SNARE binding (the main role of SNARE proteins is to mediate the fusion of vesicles with the target membrane). Eight CC were significantly enriched (P-adj Bonf > 0.05), which were mainly related to contractile fiber, sarcomere, myofibril, I band, and Z plates (Figure S1).

**Gene Network and Ontology in DEGs of A vs J individuals**

The gene network and biological pathways between DEGs of A vs J individuals were not identified, which may be due to not only the small number of DEGs (28 genes) identified between these two age stages but also the lower levels of ontological information of the sheep species compared to other species (Javadmanesh, 2020).

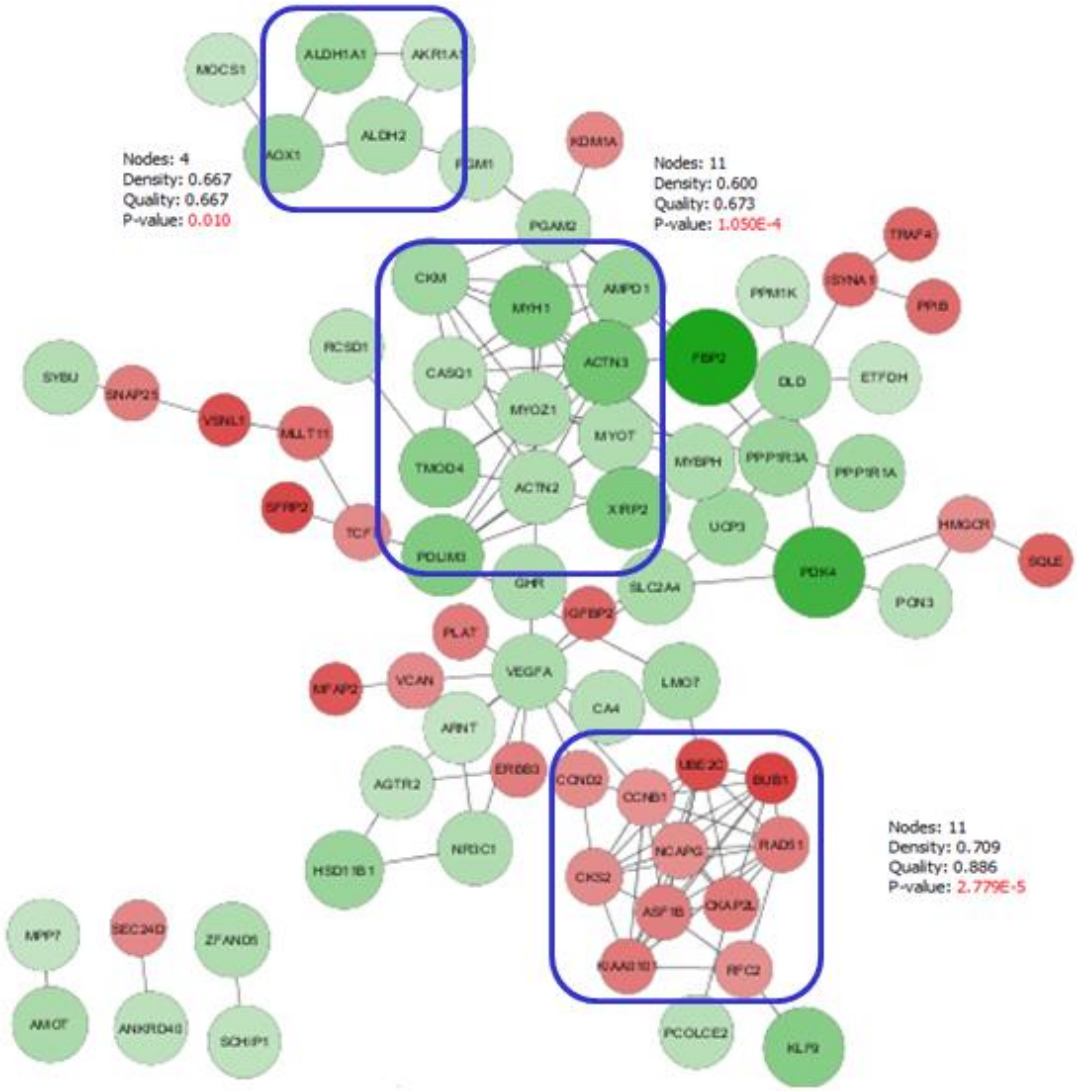


**Figure 5.** Significant clusters of interactive network between DEGs in 70-day embryonic and juvenile stages. In this network, genes with large and dark green nodes have the highest expression and smaller and red nodes have the lowest expression.

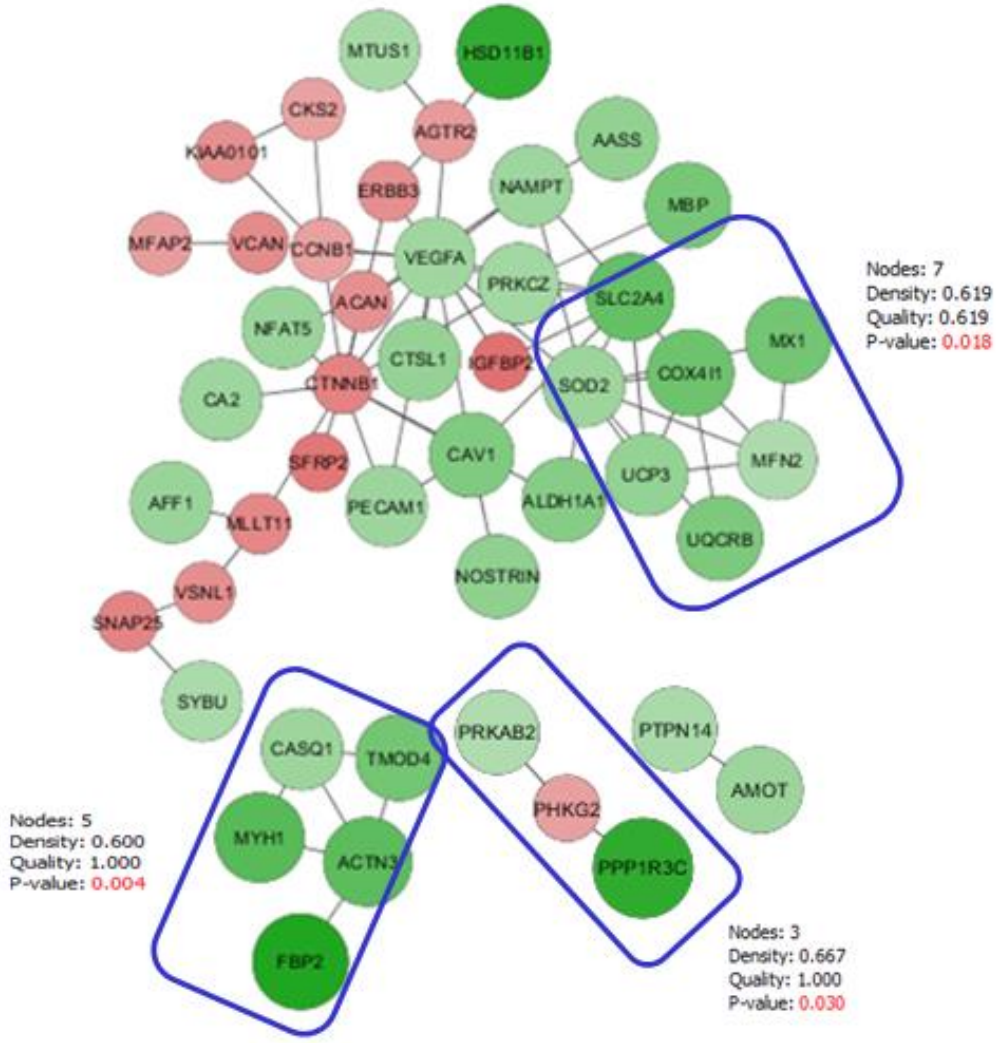


**Figure 6.** Significant clusters of an interactive network of DEGs between 70-days embryonic and adult stages. In this network, genes with large and dark green nodes have the overexpression, and smaller and red nodes have the down expression.

In ontology analysis, one MF was significantly enriched (P-adj Bonf > 0.05) which was related to collagen binding. Also, five CC were significantly enriched (P-adj Bonf > 0.05), which were related to the Golgi-Cys apparatus, cell-cell contact zone, contractile fiber, I band, and Z plates (Figure S3).

Enrichment analysis for DEGs between A vs J individuals showed 20 BP significantly enriched (P-adj Bonf > 0.05), which were mainly related to cell death in response to oxidative stress, bone regeneration, bone erosion, endothelial cell differentiation and growth, endothelial barrier creation, acetyl-CoA biosynthesis process from pyruvate, Thioester metabolism/biosynthesis process, nucleoside bisphosphate metabolism/biosynthesis, metabolic process/purine nucleoside bisphosphate biosynthesis, metabolic process Acyl-CoA, metabolic process/biosynthesis of acetyl-CoA, metabolic process/biosynthesis of ribonucleoside bisphosphate (Figure S2).

**Gene Network, Gene Ontology and Biological Pathways between DEGs of A vs E stages**

Among the 62 DEGs identified between A vs E stages, finally, the gene interactive network was plotted between 47 genes (Figure 6). The first significant cluster with a P-Value of 0.004 had five nodes including CASQ1, TMOD4, ACTN3, MYH1, and FBP2 genes, all of which had overexpression which is impressive in muscle contraction, hypertrophy, and regulation of calcium ion concentration in the sarcoplasmic reticulum (Almenar-Queralt et al., 1999; Ribeiro et al., 2014; Terentyev et al., 2003) as well as participating in regulation the process of gluconeogenesis (Rakus et al., 2005). The second significant cluster with a P-Value of 0.018 had seven nodes that included

SLC2A4, MX1, COX4I1, SOD2, MFN2, UQCRB, and UCP3 genes, and all of them had overexpression which played a role in the easy release of insulin-dependent glucose in adipose tissue and striated muscles (skeletal and heart muscle), induction of interferon in macrophages and adjustment the cellular immunity in cell (Horisberger, 1992), production of mitochondrial cytochrome c oxidase enzyme and adjustment the equilibrium of the proton electrochemical (Hüttemann et al., 2001), production of mitochondrial superoxide dismutase 2 (Zelko et al., 2002) in the mitochondrial respiratory chain and electron transport (Jung et al., 2011), The expression of GTPase proteins that are embedded in the outer membrane of mitochondria and are necessary for mitochondrial fusion and also energy balance in the mitochondria (Boss et al., 1998; Filadi et al., 2018). The SLC2A4 gene from this cluster has been identified and reported as differential genes in comparative transcriptome profiling of longissimus muscle tissues from Qianhua Mutton Merino and Small Tail Han sheep (Sun et al., 2016a). The third significant cluster with a P-Value of 0.030 also had three nodes including PRKAB2, PHKG2, and PPP1R3C genes, where PPP1R3C and PRKAB2 genes had overexpression and PHKG2 gene had down expression. These genes are involved in cell energy supply through the expression of phosphorylase b kinase enzyme (CHEUNG et al., 2000). This enzyme activates glycogen phosphorylase b to maintain the normal level of glucose between meals by breaking down the glycogen (Huang et al., 1993). PPP1R3C gene from this cluster has been identified and reported as differential genes in comparative transcriptome profiling of longissimus muscle tissues from Qianhua Mutton Merino and Small Tail Han sheep (Sun et al., 2016a).

Ontology enrichment analysis for DEGs between adult and 70 days embryonic stages showed that 25 BP were significantly enriched (P-adj Bonf > 0.05). These processes are mainly related to the bone-building process, positive regulation of organ and muscle tissue growth, angiotensin pathway and responses, regulation of striated muscle contraction, calcineurin-mediated signaling, calcineurin-NFAT signaling cascade and its regulation, regulation of receptor signaling pathway through JAK-STAT, calcium-mediated signaling, and cellular carbohydrate biosynthesis. The significant MF of enrichment (P-adj Bonf > 0.05) was related to the regulation of cation channel activity. A CC was also significantly enriched (P-adj Bonf > 0.05) which was related to contractile fiber (Figure S3). Therefore, it is very important to evaluate genes at the level of the transcriptome, which has significant differentially expressed levels in different stages of development, as well as to determine the amount and mode of action of these regulatory factors on the control of the function of a tissue.

## Conclusion

This study has investigated the expression profile of muscle genes by integrating the transcriptome data of Texel sheep muscle tissue (obtained from two different methods, microarray and RNA-Seq methods) at three important age points. By assessment the gene network among the DEGs of three age stages, 37 genes, including CCND2, UBE2C, BUB1, RAD51, CKS2, CCNB1, NCAPG, ASF1B, RFC2, KIAA0101, CKAP2L, CKM, MYH1, AMPD1, ACTN3, CASQ1, MYOZ1, MYOT, TMOD4, ACTN2, XIRP2, PDLIM3, AOX1, ALDH2, ALDH1A1, AKR1A1, FBP2, SLC2A4, MX1, COX4I1, SOD2, MFN2, UQCRB and UCP3, PRKAB2, PHKG2, PPP1R3C were selected. The gene ontology and enrichment analysis of the identified DEGs revealed the function of these genes in cell proliferation, protein synthesis, formation, and organization of myofibrils (in the embryonic stage), muscle contraction (in the juvenile stage) and lipid metabolism (in the adult stage). By integrating the expression data obtained from two different methods, this study provided a general view of the differences in transcriptomes in the muscle tissue of sheep at three important age points. Also, the selected genes such as BUB1, RFC2, KIAA0101, RAD51, CKS2, and UQCRB that were identified for the first time in this study can be a useful source for biological investigations of genes related to muscle growth and development in sheep. In the end, it can be said that the data integration method can complete many incomplete studies and provide useful results. However, for the final approval, laboratory validation and additional studies are needed.

**Acknowledgment**

This study was supported by the Ferdowsi University of Mashhad (Grant number 3/46613).

**Conflict of Interests**

None.

## References

Almenar-Queralt A., Lee A., Conley C. A., de Pouplana L. s. R. and Fowler V. M. (1999) Identification of a novel tropomodulin isoform,

skeletal tropomodulin, that caps actin filament pointed ends in fast skeletal muscle. Journal of Biological Chemistry 274:28466-28475.

Bakhshalizadeh S., Zerehdaran S. and Javadmanesh A. (2021) Meta-analysis of genome-wide association studies for somatic cells score trait in dairy cows. Journal of Ruminant Research 9:39-58.

Bakhshalizadeh S., Zerehdaran S., Hasanpur K. and Javadmanesh A. (2024) Identification of potential candidate genes associated with milk protein differences in Holstein cows: a meta-analysis integrating GWAS and RNA-Seq transcriptome. Canadian Journal of Animal Science. e-First https://doi.org/10.1139/cjas-2023-0108

Bassel-Duby R. and Olson E. N. (2006) Signaling pathways in skeletal muscle remodeling. Annual Review of Biochemistry 75:19-37.

Beggs A. H., Byers T., Knoll J., Boyce F., Bruns G. and Kunkel L. (1992) Cloning and characterization of two human skeletal muscle alpha-actinin genes located on chromosomes 1 and 11. Journal of Biological Chemistry 267:9281-9288.

Betti S. B., Tahmoorespur M. and Javadmanesh A. (2022a) Alternative Splicing Novel lncRNAs and Their Target Genes in Ovine Skeletal Muscles. Journal of Cell and Molecular Research 13:129-136.

Betti S. B., Tahmoorespur M. and Javadmanesh A. (2022b) Identification of lncRNAs expression and their regulatory networks associated with development and growth of skeletal muscle in sheep using RNA-Seq. Agriculture and Natural Resources 56:373–386-373–386.

Boss O., Giacobino J.-P. and Muzzin P. (1998) Genomic Structure of Uncoupling. Genomics 47:425-426.

Castillo D., Gálvez J. M., Herrera L. J., Román B. S., Rojas F. and Rojas I. (2017) Integration of RNA-Seq data with heterogeneous microarray data for breast cancer profiling. BMC Bioinformatics 18:1-15.

CHEUNG P. C., SALT I. P., DAVIES S. P., HARDIE D. G. and CARLING D. (2000) Characterization of AMP-activated protein kinase γ-subunit isoforms and their role in AMP binding. Biochemical Journal 346:659-669.

Ewen M. E., Sluss H. K., Sherr C. J., Matsushime H., Kato J.-y. and Livingston D. M. (1993) Functional interactions of the retinoblastoma protein with mammalian D-type cyclins. Cell 73:487-497.

Filadi R., Pendin D. and Pizzo P. (2018) Mitofusin 2: from functions to disease. Cell Death & Disease 9:1-13.

Flintoft L. (2008) Digging deep with RNA-Seq. Nature Reviews Genetics 9:568-568.

Gaeini A. (2013) Changes in ACTN3 gene expression and fiber type composition in flexor hallucis longus muscle after eight weeks progressive resistance training in Sprague-Dawley rats. Tehran University Medical Journal TUMS Publications 71:37-45.

Geiman T. M., Sankpal U. T., Robertson A. K., Chen Y., Mazumdar M., Heale J. T., Schmiesing J. A., Kim W., Yokomori K. and Zhao Y. (2004) Isolation and characterization of a novel DNA methyltransferase complex linking DNMT3B with components of the mitotic chromosome condensation machinery. Nucleic Acids Research 32:2716-2729.

Horisberger M. A. (1992) Interferon-induced human protein MxA is a GTPase which binds transiently to cellular proteins. Journal of Virology 66:4705-4709.

Huang C.-Y. F., Yuan C.-J., Livanova N. B. and Graves D. J. 1993. Expression, purification, characterization, and deletion mutations of phosphorylase kinase γ subunit: identification of an inhibitory domain in the γ subunit. *In* Reversible Protein Phosphorylation in Cell Regulation. Springer. 7-18.

Hüttemann M., Kadenbach B. and Grossman L. I. (2001) Mammalian subunit IV isoforms of cytochrome c oxidase. Gene 267:111-123.

Javadmanesh A. 2020. Genes and Genomes in Farm Animals: A Perpective. *In* Javadmanesh A. 2020. Proceedings of the 4th International and 16th Iranian Genetics Congress. , Tehran, Iran.

Jung H. J., Kim K. H., Kim N. D., Han G. and Kwon H. J. (2011) Identification of a novel small molecule targeting UQCRB of mitochondrial complex III and its anti-angiogenic activity. Bioorganic & Medicinal Chemistry Letters 21:1052-1056.

Lefrançois-Martinez A.-M., Bertherat J., Val P., Tournaire C., Gallo-Payet N., Hyndman D., Veyssière G., Bertagna X., Jean C. and Martinez A. (2004) Decreased expression of cyclic adenosine monophosphate-regulated aldose reductase (AKR1B1) is associated with malignancy in human sporadic adrenocortical tumors. The Journal of Clinical Endocrinology & Metabolism 89:3010-3019.

Liu N., He J., Yu W., Liu K., Cheng M., Liu J., He Y., Zhao J. and Qu X. (2015) Transcriptome analysis of skeletal muscle at prenatal stages in Polled Dorset versus Small-tailed Han sheep. Genetics and Molecular Research 14:1085-1095.

Ma T., Liang F., Oesterreich S. and Tseng G. C. (2017) A joint Bayesian model for integrating microarray and RNA sequencing transcriptomic data. Journal of Computational Biology 24:647-662.

Mohammadi F., Tahmoorespur M. and Javadmanesh A. (2019) Study of differentially expressed genes, related pathways and gene networks in sheep fetal muscle tissue in thin-and fat-tailed breeds. Animal Sciences Journal 32:301-312.

Ohta S., Ohsawa I., Kamino K., Ando F. and Shimokata H. 2004. Mitochondrial ALDH2 deficiency as an oxidative stress. *In* Mitochondrial Pathogenesis. Springer. 36-44.

Rakus D., Maciaszczyk E., Wawrzycka D., Ułaszewski S., Eschrich K. and Dzugaj A. (2005) The origin of the high sensitivity of muscle fructose 1, 6-bisphosphatase towards AMP. FEBS Letters 579:5577-5581.

Ribeiro J. E. A., Pinotsis N., Ghisleni A., Salmazo A., Konarev P. V., Kostan J., Sjöblom B., Schreiner C., Polyansky A. A. and Gkougkoulia E. A. (2014) The structure and regulation of human muscle α-actinin. Cell 159:1447-1460.

Rother K., Dengl M., Lorenz J., Tschöp K., Kirschner R., Mössner J. and Engeland K. (2007) Gene expression of cyclin-dependent kinase subunit Cks2 is repressed by the tumor suppressor p53 but not by the related proteins p63 or p73. FEBS letters 581:1166-1172.

Saedi N., Aminafshar M., Chamani M., Honarvar M. and Javadmanesh A. (2022) Expression pattern and network visualization of genes involved in milk persistency in bovine mammary tissue. Agriculture and Natural Resources 56(1): 23–34.

Salmikangas P., Mykkänen O.-M., Grönholm M., Heiska L., Kere J. and Carpén O. (1999) Myotilin, a novel sarcomeric protein with two Ig-like domains, is encoded by a candidate gene for limb-girdle muscular dystrophy. Human Molecular Genetics 8:1329-1336.

Sartor H., Ehlert F., Grzeschik K.-H., Müller R. and Adolph S. (1992) Assignment of two human cell cycle genes, CDC25C and CCNB1, to 5q31 and 5q12, respectively. Genomics (San Diego, Calif.) 13:911-912.

Sigruener A., Buechler C., Orso E., Hartmann A., Wild P., Terracciano L., Roncalli M., Bornstein S. and Schmitz G. (2007) Human aldehyde oxidase 1 interacts with ATP-binding cassette transporter-1 and modulates its activity in hepatocytes. Hormone and Metabolic Research 39:781-789.

Simpson F., van Bueren K. L., Butterfield N., Bennetts J. S., Bowles J., Adolphe C., Simms L. A., Young J., Walsh M. D. and Leggett B. (2006) The PCNA-associated factor KIAA0101/p15PAF binds the potential tumor suppressor product p33ING1b. Experimental Cell Research 312:73-85.

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

Sun L., Bai M., Xiang L., Zhang G., Ma W. and Jiang H. (2016a) Comparative transcriptome profiling of longissimus muscle tissues from Qianhua Mutton Merino and Small Tail Han sheep. Scientific Reports 6:33586.

Sun L., Bai M., Xiang L., Zhang G., Ma W. and Jiang H. (2016b) Comparative transcriptome profiling of longissimus muscle tissues from Qianhua Mutton Merino and Small Tail Han sheep. Scientific Reports 6:1-13.

Taheri S., Saedi N., Zerehdaran S. and Javadmanesh A. (2023) Identification of selection signatures in Capra hircus and Capra aegagrus in Iran. Animal Science Journal. 94(1): e13864.

Terentyev D., Viatchenko-Karpinski S., Györke I., Volpe P., Williams S. C. and Györke S. (2003) Calsequestrin determines the functional size and stability of cardiac intracellular calcium stores: mechanism for hereditary arrhythmia. Proceedings of the National Academy of Sciences 100:11759-11764.

Vikenes K., Andersen K. S., Melberg T., Farstad M. and Nordrehaug J. E. (2009) Long-term prognostic value of creatine kinase-myocardial band mass after cardiac surgery in low-risk patients with stable angina. Cardiology 113:122-131.

Zelko I. N., Mariani T. J. and Folz R. J. (2002) Superoxide dismutase multigene family: a comparison of the CuZn-SOD (SOD1), Mn-SOD (SOD2), and EC-SOD (SOD3) gene structures, evolution, and expression. Free Radical Biology and Medicine 33:337-349.

Zeraatpisheh, Y., Zerehdaran, S., & Javadmanesh, A. (2023). Investigation of metabolic pathways

related to the QTLs of parasite resistance trait in sheep genome using gene network and gene ontology. Veterinary Research & Biological Products, 36(1), 83-90. doi: 10.22092/vj.2022.357660.1941

Zhan S., Zhao W., Song T., Dong Y., Guo J., Cao J., Zhong T., Wang L., Li L. and Zhang H. (2018) Dynamic transcriptomic analysis in hircine longissimus dorsi muscle from fetal to neonatal development stages. Functional & Integrative Genomics 18:43-54.

Zhang C., Wang G., Wang J., Ji Z., Liu Z., Pi X. and Chen C. (2013) Characterization and comparative analyses of muscle transcriptomes in Dorper and small-tailed Han sheep using RNA-Seq technique. PloS one 8:e72686.

## Supplementary Figures

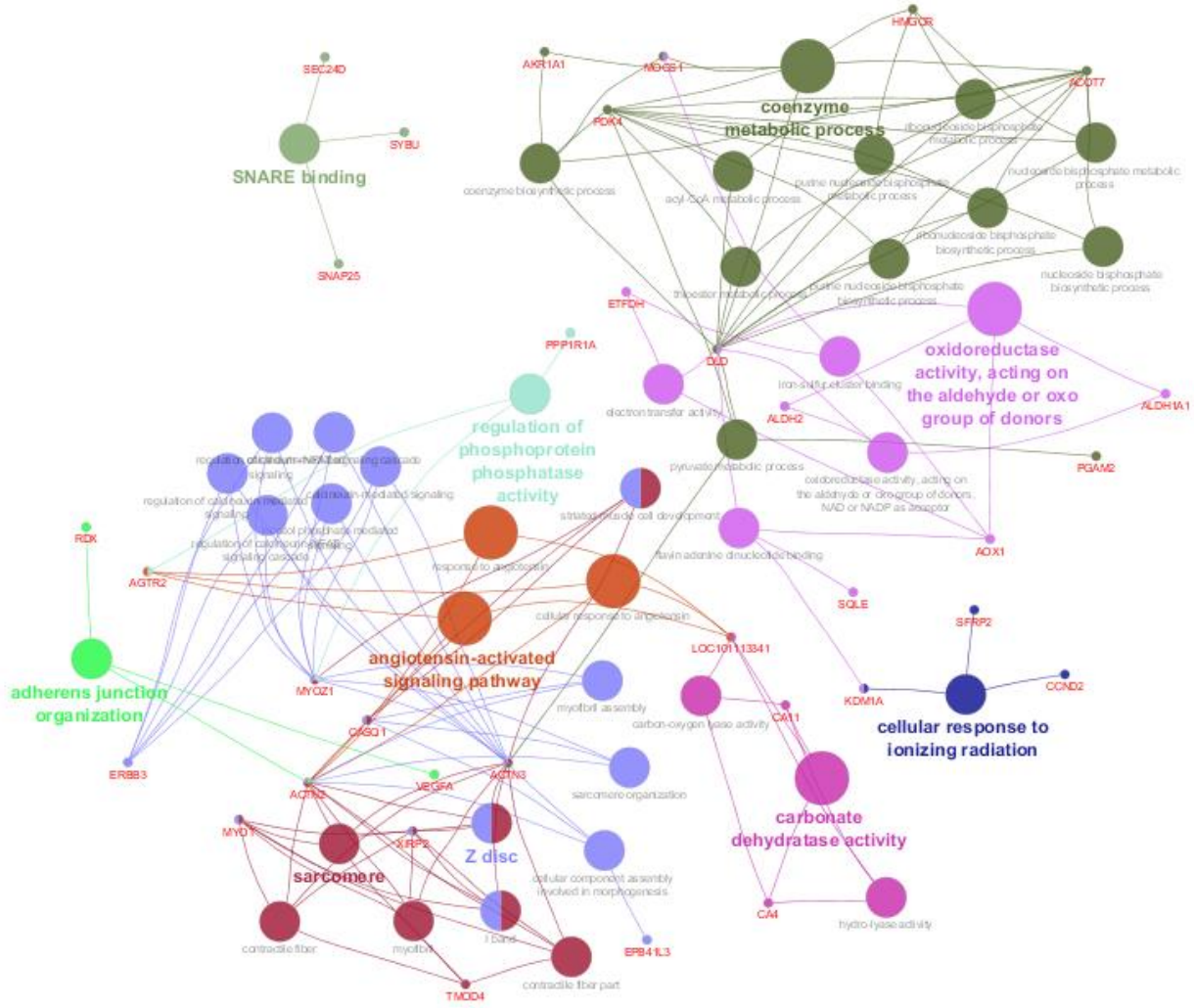


**Figure S1**. The ontology of the DEGs identified between the 70-day embryonic and juvenile stages. Small nodes are genes and large nodes are ontology phrases.

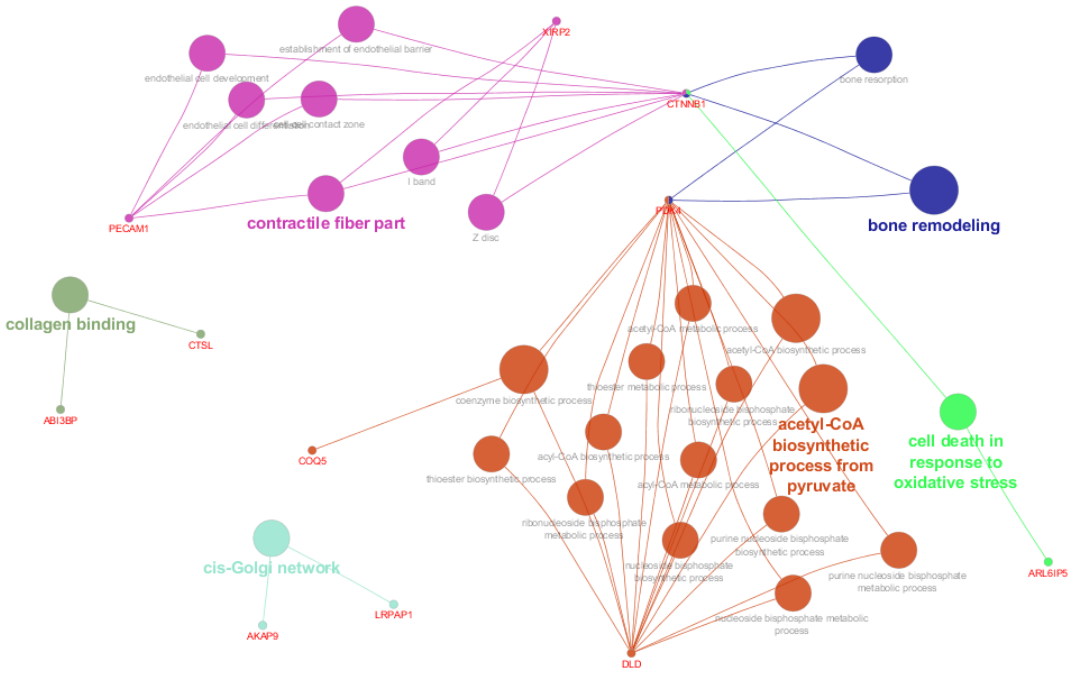


**Figure S2.** Ontology of DEGs identified between adult and juvenile stages. Small nodes of genes and large nodes of ontology are phrases

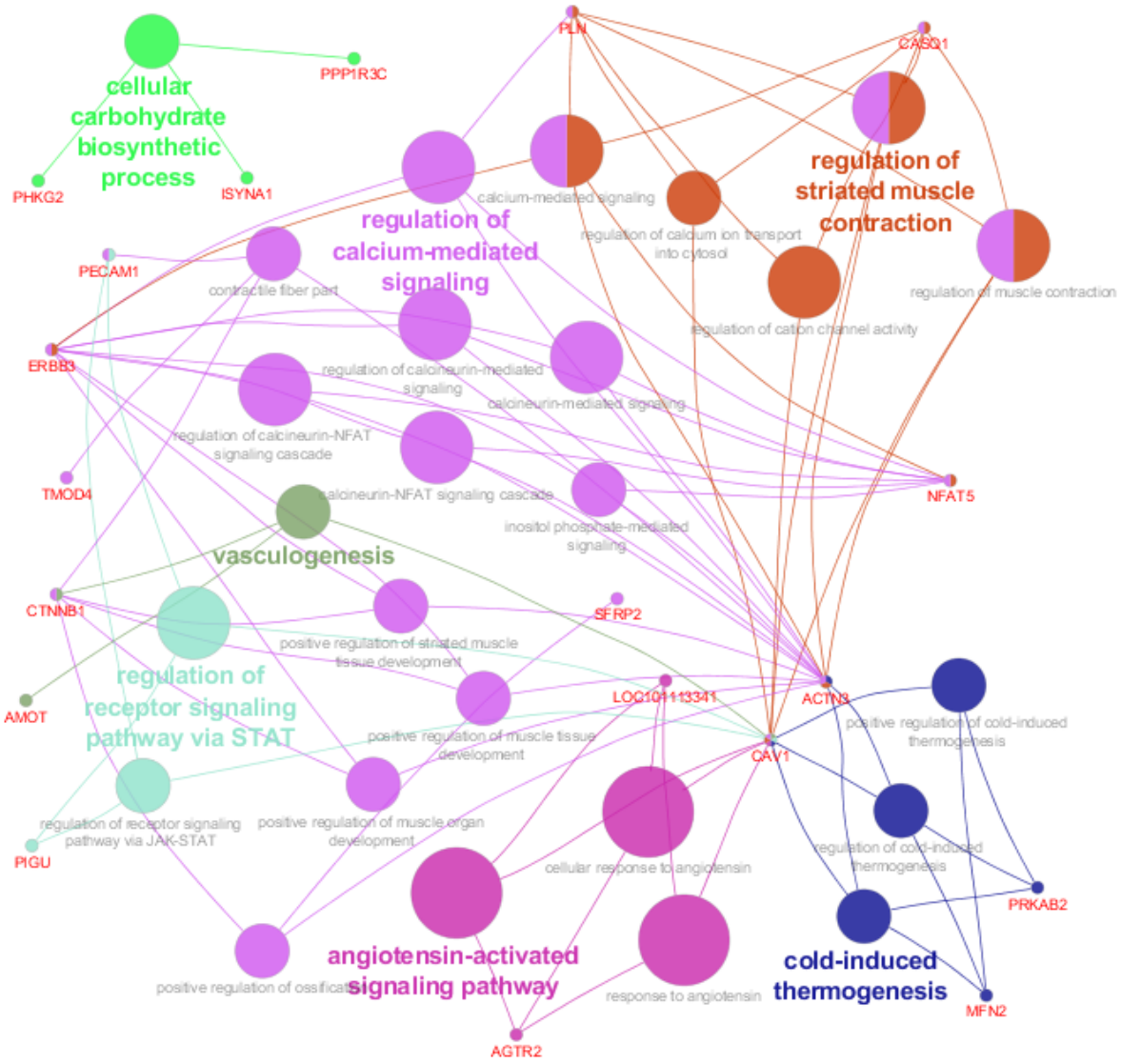


**Figure S3.** The ontology of the DEGs identified between the 70-day embryonic and adult stages. Small nodes are genes and large nodes are ontology phrases.