

# Meta-analysis of RNA-Seq datasets highlights novel genes/pathways involved in fat deposition in fat-tail of sheep

Seyedeh Fatemeh Hosseini

University of Tehran

Mohammad Reza Bakhtiarizadeh (✉ [mrbakhtiari@ut.ac.ir](mailto:mrbakhtiari@ut.ac.ir))

University of Tehran

Abdolreza Salehi

University of Tehran

---

## Article

**Keywords:** Adipose tissue, Lipid metabolism, Comparative transcriptome, Gene expression

**Posted Date:** July 7th, 2022

**DOI:** <https://doi.org/10.21203/rs.3.rs-1770948/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

Fat-tail in sheep is considered as an important energy reservoir to provide energy as a survival buffer during harsh challenges. However, fat-tail is losing its importance in modern sheep industry systems and thin-tailed breeds are more desirable. Using comparative transcriptome analysis to compare fat-tail tissue between fat- and thin-tailed sheep breeds provides a valuable approach to study the complex genetic factors associated with fat-tail development. However, Transcriptomic studies often suffer from issues with reproducibility, which can be improved by integrating multiple studies based on a meta-analysis. Hence, for the first time, an RNA-Seq meta-analysis on sheep fat-tail transcriptomes was performed using six publicly available datasets. A total of 500 genes (221 up-regulated, 279 down-regulated) were identified as differentially expressed genes (DEGs). A jackknife sensitivity analysis confirmed the robustness of the DEGs. Moreover, QTL and functional enrichment analysis reinforced the importance of the DEGs in the underlying molecular mechanisms of fat deposition. Protein-protein interactions (PPIs) network analysis revealed the functional interactions among the DEGs and the subsequent sub-network analysis led to identify six functional sub-networks. According to the results of the network analysis, down-regulated DEGs in green and pink sub-networks (like collagen subunits IV, V and VI, integrins 1 and 2, *SCD*, *SCD5*, *ELOVL6*, *ACLY*, *SLC27A2* and *LPIN1*) may impair lipolysis or fatty acid oxidation and cause fat accumulation in tail. On the other hand, up-regulated DEGs, especially those are presented in green and pink sub-networks (like *IL6*, *RBP4*, *LEPR*, *PAI-1*, *EPHX1*, *HSD11B1*, and *FMO2*), might contribute to a network controlling fat accumulation in the tail of sheep breed through mediating adipogenesis and fatty acid biosynthesis. Our results highlighted a set of known and novel genes/pathways associated with fat-tail development, which could improve the understanding of molecular mechanisms behind fat deposition in sheep fat-tail.

## Introduction

In recent years, with the growth of world population, there is a considerable demand for the agricultural improvement, especially for meat and milk of animals. In modern agriculture, sheep play an important role in the production of meat, wool and milk. Currently, there are a lot of sheep breeds over a wide geographical range worldwide with a large variation in many phenotypic traits. One of the most important phenotypic traits is the ability to store of fat in the tail. In this regard, generally sheep breeds can be divided into two main groups (thin- and fat-tailed breeds) and five subgroups including fat-long, fat-short, fat-rumped, thin-long and thin-short tailed sheep breeds<sup>1</sup>. It is well known that the fat-tailed sheep breeds are evolved from its thin-tailed ancestor (~5,000 years ago), by the principle of selective breeding (natural or artificial selection) and following initial domestication, in response to hazardous environments<sup>2,3</sup>. Fat-tail is considered as an important energy reservoir for the animal to provide energy as a survival buffer during harsh challenges such as drought and food deprivation periods like winter, when pasture is dormant<sup>4</sup>. Nowadays, fat-tail is losing its importance and lean-tailed breeds are more desirable, which can be explained by providing three reasons: 1) In modern sheep farming systems, intensive or semi-intensive feeding systems are preferred, which cause fat-tail be no longer important as an energy source;

2) fat deposition led to a higher energetic cost than accretion of an equivalent amount of lean tissue and decreases the feed efficiency; and 3) in modern society, consumers prefer low-fat foods to be healthy. Therefore, both customers and producers prefer thin-tailed sheep<sup>5</sup>. It is estimated that the fat-tailed sheep breeds constitute about 26% of the world sheep population<sup>6</sup>. Hence, study of the fat deposition in sheep tail is of great importance to understand the background molecular mechanisms of the fat-tail development as well as to develop the new breeding strategies to produce improved breeding sheep herds.

Up to now, several transcriptome-based studies have been conducted to elucidate the factors and molecular mechanisms responsible for the differences in fat deposition between different sheep breeds<sup>5,7-16</sup>. Earlier studies were focused on single gene to identify the candidate genes involved in regulating the fat-tail formation. In this regard, higher expression of the leptin (*LEP*) gene in different fat tissues of fat genotype than lean genotype Coopworth sheep is reported<sup>17</sup>. In our previous study, fatty acid binding protein 4 (*FABP4*) gene was suggested as a candidate gene relevant to fat deposition in sheep breeds<sup>7</sup>. Moreover, uncoupling protein 1 (*UCP1*)<sup>14</sup> and cell death-inducing DFFA-like effector c (*CIDEc*)<sup>18</sup> genes were reported as other candidate genes involved in regulation of fat deposition and mobilization in tail of sheep breeds.

In recent years, rapid development of next generation sequencing (*NGS*) technologies has lowered the barriers to perform high-throughput gene expression profiling through RNA sequencing (RNA-Seq) approach. This method made it possible to measure the expression of thousands of genes simultaneously, which enables us to better understand the genetic factors associated with phenotypic differences in sheep fat-tail. To date, several studies have been employed this approach to identify the genes linked to fat-tail deposition by conducting a comparative transcriptome study on the fat- and thin-tailed sheep breeds<sup>5,7-14,16,18</sup>. These studies are summarized in Table 1 and more details are provided in Supplementary File S1.

Table 1  
List of the comparative transcriptome studies on the fat- and thin-tailed sheep breeds

Fat (Nu of BR) *	Thin (Nu of BR) *	Up genes **	Down genes **	Type of reads ***	Sex ****	Accession number	Reference
Kazak (1)	Tibetan (1)	280	366	P (100)	M	NA	Wang et al. 2014) <sup>13</sup>
Dorset (1)	Han (1)	266	336	P (100)	M, F	NA	(Miao et al. 2015) <sup>12</sup>
Guangling (1)	Han (1)	3160	2235	P (90)	M	SRP113440	(Li et al. 2018) <sup>10</sup>
Lanzhou (3)	Han (3)	7	3	P (150)	M	PRJNA432669	(Ma et al. 2018) <sup>11</sup>
Lori (3)	Zel (3)	80	184	P (150)	M	PRJNA508203	(Bakhtiarizadeh et al. 2019) <sup>5</sup>
Hulun-Fat (3)	Hulun-Thin (3)	515	1012	P (101)	M, F	PRJNA517348	(Fan et al. 2019) <sup>8</sup>
Han-Fat (3)	Han-Thin (3)	237	284	P (75)	M	PRJNA699984	(Guangli et al. 2020) <sup>19</sup>
DHH (3)	DDH (3)	262	181	P (150)	M	PRJNA745517	(Yuan et al. 2021) <sup>15</sup>
Ghezel (3)	Zel (4)	254	78	P (150)	M	PRJNA598581	(Farhadi et al. 2021) <sup>9</sup>
Bashby (3)	Bashby-Argali (3)	-	-	P (150)	M	NA	(Yuan et al. 2021) <sup>20</sup>
Altay (1)	XFW (1)	1389	6652	P (100)	M	PRJNA627341	(Zhang et al. 2021) <sup>16</sup>

\* Name of the fat- and thin-tailed breeds (Number of biological replications), \*\* Up- and down-regulated genes in fat-tailed sheep breed. \*\*\* P: paired-end and S: single-end reads (length of reads). \*\*\*\* M: male and F: female. The bolded accession numbers are used in this study.

A review of these studies indicates the variability in their results including different set of DEGs or inconsistent gene expression patterns, which can be attributed to different bioinformatics pipeline used, number of biological replicates and other limitations coming from the different nature of the samples. For example, Wang et al., (2014) suggested *NELL1* and *FMO3* as functional candidate genes that may regulate fat mass in the tails of sheep breeds<sup>13</sup>. Fourteen genes (including *FABP4*, *ADIPOQ*, *FABP5*, and

*CD36*) were reported as potential responsible for the difference in fat deposition between fat- and thin-tailed sheep breeds in Li et al., (2018) study, which five of them were ECM related genes including *LAMB3*, *RELN*, *TNXB*, *ITGA8*, and *LAMB4*<sup>10</sup>. A different set of potential related genes to fat-tail development (including *ZFP36*, *JUNB*, *RARRES2*, *ADIRF*, *AFT3*, *ARID5A*, *MYBL2*) were identified in our previous study<sup>5</sup>. In this context, a meta-analysis of several independent studies focused on a specific biological question, offers a useful approach to increase the statistical power. Additionally, combining evidence from several studies increases the reliability and robustness of the results<sup>5,10–13,16,19</sup>. In this study, for the first time a meta-analysis of six independent RNA-Seq studies was performed to provide a basis for the identification of mechanisms underlying fat deposition in sheep tail, which can be exploited in future breeding strategies.

## Material And Methods

### RNA-Seq datasets

A comprehensive review of authoritative papers was conducted to find the studies that performed gene expression profiling by RNA-Seq method to identify molecular mechanisms affecting sheep fat-tails (Table 1, Supplementary File S1). Out of 11 identified studies, four studies<sup>10,12,13,16</sup> had one biological replication per breed and data of one study was not available<sup>16</sup> in the public databases. Hence, six studies that met the requirements for meta-analysis were used in this study (Table 1). Only male samples of Fan et al., (2019)<sup>8</sup> study were used. RNA-Seq reads for the selected studies (Table 1) were retrieved from the NCBI's sequence read archive database (SRA: <https://trace.ncbi.nlm.nih.gov/Traces/sra/>). All samples of each study were processed with the same bioinformatic pipeline, explained as follow (Figure 1).

### Differential expression analysis

The quality of RNA-Seq reads was checked using FastQC (v0.11.5) software<sup>21</sup>. Trimmomatic (v0.38) software<sup>22</sup> was applied to remove low-quality reads/bases and adaptor sequences based on its adaptive trimming algorithm, maximum information (MAXINFO), to balance the benefits of retaining longer reads against the drawback of having low-quality bases. The other trimming criteria were TRAILING: 20, MAXINFO: 80: 0.8, MINLEN: 80. FastQC was used again to reevaluate the quality of the clean reads. The clean reads were then aligned to the ENSEMBL bovine reference genome (rambouillet\_v1.0.104) using HISAT2<sup>23</sup> software (version 2.1.0) based on the default parameters. A list of exon-exon junctions extracted from the Ensembl bovine GTF file (rambouillet\_v1.0.104) was applied to guide the read mapping. To quantify the aligned reads to annotated genes, python script HTSeq-count (version 2.7.3) tool<sup>24</sup> was used (union intersection mode) according to Ensembl bovine GTF file (rambouillet\_v1.0.104) and a gene expression matrix was created for each study. Genes with low read counts (<10 reads across all samples) were filtered out. DESeq2 (v 1.32.0) was used to normalize and perform differential

expression analysis for individual studies using the median-of-ratios method as a normalization factor. Moreover Cox-Reid adjusted profile likelihood approach was used to estimate dispersions.

### **RNA-Seq Meta-analysis**

Raw P-values obtained from previous step were applied to perform meta-analysis using metaRNAseq R package (v1.0.7). There are two p-value combination methods to combine p-values from different studies in this package including Fisher's and inverse normal methods. It is reported that both methods provide similar results. Using these methods enable us to combine results from heterogeneous datasets directly to identify commonly regulated genes among all studies. Here, to minimize the false positive results, both methods were used and differentially expressed genes (DEGs) with a false discovery rate (FDR) < 0.05 found in common between two methods were considered statistically significant. One of the limitations of p-value-based methods is that the combined p-values are estimated regardless of the expression patterns in the different studies. In other words, some genes are significantly up-regulated in some studies, while in the other studies they are significantly down-regulated, which may not reflect the biological reality<sup>25</sup>. This limitation can be overcome by selecting the DEGs that showed the same expression pattern in most of studies. In the present study, out of the DEGs that were common between the two methods, only those genes that showed the same expression trend (up or down regulated) in most of the datasets (>0.50%), were finally considered as DEGs.

### **Jackknife analysis**

One of the concerns in meta-analysis is that the identified DEGs be affected by the properties of a single dataset<sup>26</sup>. Hence, to evaluate the robustness of the meta-analysis results, a Jackknife sensitivity analysis was applied. To do this, meta-analysis was repeated (as described above) six times, each time including all the datasets except one and six sets of DEGs were generated. Finally, number of presence/absence of DEGs in the six sub meta-analysis was compared with the main meta-analysis.

### **Functional enrichment analysis**

EnrichR web-based tool was used to perform the functional enrichment analysis, focusing on gene ontology (GO; molecular process, biological process) and the kyoto gene and genome encyclopaedia (KEGG) pathway. To do this, up- and down-regulated DEGs were submitted separately and adjusted p-value < 0.05 was considered for the selection of the significantly enriched GO terms and pathways.

### **QTL enrichment Analysis**

To determine if the DEGs are located in the QTLs associated with fat metabolism, a co-localization analysis was applied. To do this, all QTLs related to fatness were obtained from sheep AnimalQTLdb database<sup>27</sup>. Then, enrichment of the DEGs within the QTLs were analyzed using GSAQ R package based on Gene Set Validation with QTL (GSVQ) method<sup>28</sup>.

### **Network analysis**

To reveal the protein-protein interactions (PPIs) among the DEGs, STRING<sup>29</sup> database (v11.5) was applied using *Ovis aries* as the reference organism. The PPIs networks were constructed for up- and down-regulated DEGs, separately. Confidence score <0.4 (a commonly used threshold) was considered to discard unreliable PPIs. Moreover, sub-networks within the networks were identified based on the k-means clustering approach (3 clusters) in the database. Cytoscape software (version 3.9) was used to visualize the networks.

## Results

### RNA-Seq data analysis

Six RNA-Seq datasets were analyzed including 18 fat-tailed and 19 thin-tailed sheep breeds (Table 1). In total, more than 952 million paired-end raw reads were obtained for the six datasets, ranging from 7.5 to 50.5 million per sample. Among these, 468 and 484 million reads were belonged to fat- and thin-tailed breeds, respectively. Only 1.6 million reads (~ 0.002%) were removed after quality control and filtering, which shows the quality of the datasets. The clean reads were then aligned to the sheep genome with an average of 87.26% read mapping rate. Detailed information of number of reads and alignment rate per sample are provided in Supplementary File S1.

### Meta analysis

Differential gene expression analysis was performed for each study and the raw p-values of all studies were integrated based on our meta-analysis approach. After multiple testing correction, 761 DEGs were identified. Inspecting the DEGs to have the same log-fold change direction across most of the studies led to identification of 500 DEGs including 221 up- and 279 down-regulated genes in fat- against thin-tailed sheep breeds. Results of the differential expression analysis per study and meta-analysis can be found in Supplementary File S2. To illustrate the common DEGs shared among the individual studies and meta-analysis, a venn diagram was constructed (Fig. 2). Out of 500 DEGs, 330 DEGs were shared with the other studies and 170 DEGs were identified only through the meta-analysis. Moreover, most of the identified DEGs from each individual study were not shared with the other studies. No DEG was found to be common among all the studies and meta-analysis.

### Jackknife sensitivity analysis

To confirm that our meta-analysis results were not affected by any single study, a jackknife sensitivity analysis was conducted by iteratively removing one dataset at a time. Figure 3 shows the common DEGs among the six Jackknife tests and the full meta-analysis results. All the identified DEGs by meta-analysis (except two genes *ASIC3* and *MYBPC1*) were identified in at least one Jackknife test. On the other hand, more than 80% of the DEGs (403 genes including 232 down- and 171 up-regulated) were present in at least 50% of the Jackknife test results (oval circles in Figs. 4 and 5 represent these genes). Sixty-eight

DEGs pass all six jackknife analyses and showed complete consistency over the tests, indicative of their importance as they can be considered as a more robust set of potential candidates.

## Functional enrichment analysis

In up-regulated genes, 145 biological processes and 14 KEGG pathways were significant at  $p$ -value  $< 0.05$  and none of them were significant at  $FDR < 0.05$  (Supplementary File S3). In total, 261 biological processes and 26 KEGG pathways were significantly enriched in down-regulated DEGs (at  $p$ -value  $< 0.05$ ), of which 26 and seven were significant at  $FDR < 0.05$ , respectively (Supplementary File S3). It is worthy to highlight the significant GO terms related to fat metabolism including "unsaturated fatty acid biosynthetic process" and "fatty-acyl-CoA biosynthetic process" in the down-regulated DEGs. Thus, it is reasonable to infer that these genes might be related to fat-tail development.

## QTL enrichment analysis

A total of 98 annotated QTLs related to fatness were obtained from the sheep AnimalQTLdb database<sup>27</sup>. Totally, 113 genes (including 50 up- and 63 down-regulated as represented in bold text in Figs. 4 and 5) were significantly ( $p$ -value = 0.018) located within coordinates of 16 QTLs (Supplementary File S4), which were located in chromosomes 1, 6, 8, 10, 11, 12, 18, 14, 16, 19 and 23. These QTLs pertained to nine different traits including internal fat amount, fat weight in carcass, carcass fat percentage, muscle depth at third lumbar, dressing percentage, subcutaneous fat area, subcutaneous fat thickness, carcass length and lean meat yield percentage. Out of 113 genes, 35 up- and 55 down-regulated DEGs were present in at least 50% of the Jackknife test results ( $\sim 80\%$ ). To check whether the identified DEGs per individual study are significantly enriched in the QTLs, this analysis was performed on the DEGs of each study, separately. Out of six study, only DEGs of the study one was significantly enriched in the QTLs.

## PPI network analysis

To construct an PPI network of DEGs, where the nodes are proteins and the edges represent the predicted functional associations, STRING database<sup>29</sup> was used. ENSEMBL accession numbers of DEGs were first annotated and 361 gene names (135 up and 226 down) were found. Of these, 125 and 218 genes were matched with the database and applied to construct the PPI network in up- and down-regulated DEGs, respectively. A total of 72 nodes with 98 edges (known or predicted interactions) composed the final PPIs network of up-regulated DEGs (PPI enrichment  $p$ -value = 0.002, Fig. 4). Further, the gene network interaction of down-regulated DEGs included 157 nodes with 337 edges (PPI enrichment  $p$ -value =  $< 1.0e-16$ , Fig. 4).

The resulting PPIs networks were subjected to module analysis with kmeans clustering approach and three significant sub-networks were identified in each network. In this regard, three sub-networks including 22 nodes and 20 edges in sub-network orange ( $p$ -value =  $6.51E-09$ ), 24 nodes and 55 edges in yellow sub-network ( $p$ -value =  $5.49E-11$ ) and 20 nodes and 16 edges in light-blue sub-network ( $p$ -value =  $2.89E-06$ ) were found in up-regulated DEGs (Fig. 4). Functional enrichment analysis of the clusters showed only six

significant biological processes in orange sub-network including “gene expression”, “RNA metabolic process” and “translation”. On the other hand, 34 nodes and 41 edges in green sub-network ( $p$ -value  $< 1.0e-16$ ), 64 nodes and 90 edges in pink sub-network ( $p$ -value  $< 1.0e-16$ ) and 51 nodes and 105 edges in purple sub-network ( $p$ -value  $< 1.0e-16$ ) were found in the down-regulated DEGs (Fig. 5). Genes of green sub-network were significantly enriched in “Lipid biosynthetic process” biological process (FDR  $< 0.04$ ) and “Biosynthesis of unsaturated fatty acids” KEGG pathway (FDR  $< 0.01$ ). Moreover, 17 biological processes (such as “Extracellular matrix organization” and “Anatomical structure morphogenesis”) and 11 KEGG pathways (such as “ECM-receptor interaction”, “PI3K-Akt signaling pathway” and “Focal adhesion”) were identified in pink sub-network (Fig. 6). No significant term was found in purple sub-network.

## Discussion

To date, several transcriptome-based studies have been conducted to identify DEGs related to fat-tail formation through comparison of fat- and thin-tailed sheep breeds (Table 1). However, one of the main problems of these studies is a small number of biological replicates due to the cost of sequencing, which can lead to decrease the statistical power to detect DEGs. In this regard, the variation among the identified DEGs by different datasets (Fig. 2) emphasized the importance of meta-analysis to obtain core genes that are consistently DEGs across multiple studies. Some of the lack of agreement from one study to another can be attributed to differences in the experimental conditions such as environmental variables, bioinformatics pipeline, etc.

It should be noted that various sheep breeds have been used in the six studies, however these breeds can be used to identify core genes involved in fat deposition in the fat-tail of sheep. While vertebrate species differ in many phenotypic traits, however, organ physiology in mammals are conserved. Furthermore, it is well known that many biological processes and pathways are deeply conserved among the species<sup>30</sup>. Based on this idea that many developmental gene expression patterns are conserved across mammals, rodents can be applied as models of human tissues physiology<sup>31</sup>. Previous studies in this area have shown that inter-study distances between similar tissues of various species were generally less than intra-study distances among various tissues, implying that meta-analysis of different RNA-seq studies can provide insights into the molecular mechanisms behind the mammalian tissues<sup>32</sup>. Here, a meta-analysis was performed to find the common core DEGs across six studies and remove the inconsistencies in these datasets to obtain a deeper biological insight associated with fat deposition in tail of sheep, compared to that achieved through single dataset. These core genes can be considered as putative candidates involved in fat-tail development regardless of the different characteristics of the studies. Therefore, our aim was to shed light into the molecular mechanisms conserved across the sheep breeds instead of identifying specific mechanisms in each breed.

To better integrate the datasets, a standard and similar bioinformatics workflow was applied to analyze the raw RNA-Seq data from the studies, in which different pipelines were used. To further evaluate the consistency of the DEGs across the datasets, Jackknife sensitivity analysis was applied and revealed

that most of the DEGs (> 80%) passed at least 50% of the Jackknife tests, which indicates the robustness of the results and independency of these genes to a single study. It is worth to note that in this study only DEGs with the same pattern of expression over most of the datasets (> 50%) were considered, which make them more robust and constitutes a suitable way for better understanding the regulatory mechanisms of fat-tail development. To assess potential functions of DEGs in fat-tail development, we investigate their relationships with the QTLs related to sheep fatness and found that they were significantly enriched in these QTLs. These findings reinforced that the identified DEGs may play important roles in fat-tail development in sheep.

GO terms as well as KEGG pathways associated with fat metabolism were the most enriched terms in the DEGs, which explains the fact that these genes can be related to fat-tail formation (Supplementary File S3). Some of the DEGs grouped under terms that were directly or indirectly related to fat metabolism, including *LEPDRBP4*, *CSNK2B*, *EPHX1*, *SCD*, *ALOX15*, *SCD5*, *HSPA6*, *ACLY*, *ELOVL6*, *FOXO3*, *PPP5C*, *DAGLA*, *ACSF2*, *DHCR24*, *JUNB*, *FNDC5*, *FMO2*, *FOS*, *FRZB*, *ZBTB16*, *LSS*, *COL5A1*, *SLC27A2*, *SPTLC3*, *PLPP3*, *ZBTB20*, *SMPDL3B*, *LBP*, *LPIN1*, *IL6*, *PDGFRA*, *CCL21*, *SPP1*, *PAI-1*, *HSD11B1*, *MTMR4*, *S100A1*, *COL6A3*, *TNNC2*, *ITGA1*, *SCG5*, *NR4A1*, *DUSP1* and *APOE*<sup>5,7-13,33-36</sup>. Most of these genes passed at least 50% of the Jackknife tests or reside in the QTLs previously found to be linked to sheep fatness (Supplementary File s S2 and S4). Some of these genes require particular attention considering that they have been previously linked to fat deposition.

PPIs network construction by STRING database confirmed that up- and down-regulated DEGs were members of functional interaction networks. It is well known that genes from the same sub-network in a PPI network more probably play similar roles and are implicated in the same biological functions. To investigate this hypothesis, sub-network analysis on the PPI network results was performed and two of the three identified sub-networks in the down-regulated DEGs were significantly enriched in different GO terms and KEGG pathways. Hence, it might be possible to hypothesize that genes of green sub-network may be involved in fat deposition in tail of sheep breeds, as “Biosynthesis of unsaturated fatty acids” and “Lipid biosynthetic process” were significantly enriched (Fig. 5). Some genes of this sub-network that may be of particular interests are *SCD*, *SCD5*, *ELOVL6*, *ACLY*, *DHCR24*, *SPTLC3*, *LSS*, *LPIN1*, *SLC27A2*, *CKAP2*, *FOS*, *JUNB*, *FOXO3*, *EBF2* and *MYBL2*, which are located in the fatness-related QTLs regions or passed at least 50% of the Jackknife tests (Supplementary File s S2 and S4). In our previous study we hypothesized that down-regulation of the related genes to lipid metabolism in fa-tailed breed may be associated with other pathways than fat deposition such as fat composition<sup>5</sup> or fatty acid oxidation. It has been demonstrated that the breed has significant effect on fatty acid composition of tail fat<sup>37</sup>. Here, the similar results were obtained as some members of green sub-network were involved in the pathways associated with fat composition or fatty acid oxidation such as *SCD*, *SCD5*, *ELOVL6*, *ACLY*, *SLC27A2* and *LPIN1*.

Isoforms of stearoyl CoA desaturase (*SCD* and *SCD5*) play important roles in desaturation of saturated fatty acids. Lower expression of these genes has been previously reported in fat-tailed than thin-tailed sheep breeds<sup>10,12</sup>, which can cause lower content of saturated fatty acids in fat-tailed breeds. In chicken,

*SCD5* expression was significantly correlated with levels of stearidonic acid<sup>38</sup>, which might lead to suppress adipocyte differentiation<sup>39</sup>. Very long-chain fatty acids protein 6 (*ELOVL6*) is well known as a major regulator of fatty acid composition in mammals<sup>40</sup>. The first step of de novo lipogenesis, the conversion of citrate into acetyl-CoA, is catalyzed by *ACLY*. This product, acetyl-CoA, can then be used as building block for long chain fatty acids, cholesterol synthesis and/or histone acetylation<sup>41</sup>. It has been shown that *ACLY*-deficient adipocytes accumulate lipid in vivo and display some differences in fatty acid content and synthesis<sup>42</sup>. Mammals express three lipin genes (*LPIN1-3*) that their functions are evolutionarily conserved. One of the important roles of these genes is transcriptional co-regulators of gene expression in the nucleus to promote fatty acid oxidation<sup>43</sup>. In accordance with our results, it was reported that the role of *LPIN1* might be closely associated with fatty acid oxidation in the bovine liver<sup>44</sup>. Solute carrier family 27 member 2 (*SLC27A2*) is a transmembrane transporter coenzyme that participates in the long-chain fatty acid beta-oxidation<sup>45</sup> as well as plays a key role in fatty acid degradation<sup>46</sup>. *SLC27A2* was identified to be closely associated with tail phenotype in Zhang et al. study that investigated the transcriptome profiles of fat deposition in tail of sheep<sup>36</sup>. A recent study investigated the correlation of slow-growing meat type chicken liver gene expression with abdominal fat deposition using a time-course transcriptomic study (from the embryonic to the egg-producing period) and reported *CKAP2* as one of the important candidates involved in fat metabolism<sup>47</sup>. Furthermore, in green sub-network, five genes including *JUNB*, *FOXO3*, *EBF2*, *FOS* and *MYBL2* were transcription factors with known roles in lipid metabolism, which can be considered as regulators of this sub-network. A genome-wide association study showed that *MYBL2* may be involved in abdominal fat deposition in chickens<sup>48</sup>. *JUNB* is a transcription factor whose role in lipid metabolism and fat cell differentiation has been documented<sup>49</sup>. Important roles of *FOXO* protein family in energy homeostasis and lipid metabolism have been highlighted in previous studies<sup>50</sup>. Based on our findings, lower expression of members of the green sub-network in fat-tailed sheep breeds are reasonable and might be potential candidate contributing to shaping fat-tail phenotype through regulating fat composition or fatty acid oxidation.

On the other hand, some pathways related to the fat metabolism including “Extracellular matrix (ECM)-receptor interaction”, “Focal adhesion” and “PI3K-Akt signaling pathway” were significantly enriched in the pink sub-network, which can further contribute to understanding the fat-tail development in sheep (Fig. 6). All the genes associated with these pathways (*SPP1*, *COL1A1*, *ITGA2*, *ITGA1*, *VEGFA*, *COL6A6*, *COL4A6*, *PDGFRA*, *COL6A3*, *SDC1*) were located in the QTLs related to fatness or passed at least 50% of the Jackknife tests (Supplementary File s S2 and S4), which indicated their potential involvement in fat-tail morphogenesis. Focal adhesion is related to ECM and act as mechanical linkages to the extracellular matrix<sup>51</sup>. Hence, the DEGs belonging to these two pathways were very similar (Supplementary File S3). ECM is essential for tissue homeostasis and consists of a complex mixture of functional macromolecules in adipose tissue, such as glycosaminoglycans, collagen, elastin, fibronectin, and laminin<sup>52</sup>. The ECM communicates with cells through cell surface-related elements such as integrin and regulates cell activities such as differentiation, proliferation, migration, adhesion and apoptosis<sup>53</sup>. ECM receptor interaction signaling pathway has been reported as a significant enriched pathway in the DEGs

between fat- and thin-tailed sheep breeds in several previous studies<sup>5,10,15</sup> as well as in the DEGs between lean and obese human<sup>54</sup>. Moreover, interaction between ECM components and transmembrane receptors of fat cells have been demonstrated to be associated with depot-specific adipogenesis in bovine<sup>55</sup>. However, we believed that research studies in animal field have paid no sufficient attention to investigate the role of the ECM receptor interaction pathway in fat metabolism as well as fat-tail formation in sheep. ECM was detected as an inhibited KEGG pathway during differentiation of human mesenchymal stromal-cells into adipocytes. Down-regulated genes belonging to this pathway were in agreement with our results like collagen subunits IV, V and VI (*COL1A1*, *COL6A6*, *COL4A6*, and *COL6A3*), different integrins 1 and 2 (*ITGA1* and *ITGA2*) and proteoglycans like syndecan 1 and 4 (*SDC1*). Down-regulation of these genes are attributed to cytoskeleton reorganization during adipogenesis<sup>56</sup>. Collagen type IV denaturation is reported to be associated with adipogenic differentiation<sup>57</sup>. In a comprehensive assessment to study the role of *COL6A3* in human obesity and diabetes, it was revealed that *COL6A3* expression increased after weight loss and showed a negative correlation with obesity, which is in good agreement with our findings in the current study<sup>58</sup>. Moreover, it has been shown that integrin  $\alpha 5$  is down-regulated during adipogenesis in 3T3-L1 cells. Therefore, up regulation of this gene inhibits cellular differentiation<sup>59</sup>. Accordingly, our findings let hypothesize that ECM might mediate a mechanism involved in the differentiation of fat-tail adipocytes and lipid metabolism, thereby changing the fat-tail morphology of various sheep breeds.

Phosphatidylinositol 3-kinase/protein kinase B (PI3K/AKT) signaling pathway, as the other enriched pathway in pink sub-network, is a key pathway that specifically phosphorylates phosphatidylinositol, generate an intracellular second messengers and mediates glucose and lipid metabolism<sup>60</sup>. It is reported that PI3K/AKT signaling pathway has different functions in various adipocytes, as promote and inhibit the proliferation/differentiation in human adipocytes and 3T3-L1, respectively<sup>61,62</sup>. It has been shown that inhibition of this pathway in children with simple obesity participates in the occurrence and progression of obesity<sup>63</sup>. In a recent study, PI3K/AKT signaling pathways was enriched in the target genes of differentially expressed miRNAs between fat- and thin-tailed sheep breeds<sup>64</sup> and suggested to be involved in biological processes related to fat deposition in fat-tail tissue. In the current study, enrichment of this pathway in the down-regulated genes in fat-tailed breeds suggesting that it may inhibits the proliferation and differentiation of lipid metabolism in fat-tail tissue and leads to differences in fat deposition between the different sheep breeds. Some of the DEGs belonging to PI3K/AKT signaling pathway in this study were included *SPP1*, *PDGFRA*, *VEGFA* and *TNC*. Positive and negative effects of *SPP1* in fat deposition had been demonstrated in previous reports. Studies have established that this gene is synthesized by adipocytes and its higher expression is related to fat deposition<sup>65</sup>. In contrast, it is reported that interaction of *SPP1* with integrin  $\alpha v/\beta 1$  inhibits the adipogenesis of mesenchymal stem cells<sup>66</sup>. In agreement with our results, a recent study found that *SPP1* negatively regulated adipogenic differentiation of peripheral blood-derived mesenchymal stem cells and interaction between novel-miR-659 and *SPP1* coregulate fat deposition in castrated male pigs<sup>34</sup>. All these findings support the potential function of PI3K/AKT signaling pathway and its related genes in mediating lipolysis and energy expenditure, which can be led to lower fat deposition.

Some of the up-regulated DEGs in this study have been previously reported as candidate genes involved in fat-tail/fatness development or as DEG associated with fat metabolism in animal. Moreover, they were found in the fatness-related QTLs or passed at least 50% of the Jackknife tests (Supplementary File s S2 and S4). Some of the most important genes in light blue sub-network (Fig. 4) were *IL6*, *RBP4*, *LEPR*, *PAI-1*, *CSNK2B* and *MTMR4*. Recent studies highlighted the role of interleukins in lipid metabolism<sup>44</sup>. Interleukin-6 (IL-6) is known as a key regulator of adipose homeostasis in obesity. It is worth to note that in Farhadi et al., study<sup>9</sup> this gene was down-regulated in fat-tailed breed and they suggested it as a potential candidate gene in fat-tail formation. However, our analysis revealed that this gene was up-regulated in fat-tailed breed in all studies, except Farhadi et al., study<sup>9</sup>. In this regard, results of a study showed that expression of *IL-6* in lymphedema (a morbid disease characterized by adipose deposition<sup>67</sup>) is associated with adipose deposition. Since, *IL-6* was identified as a highly connected gene in light blue sub-network (Fig. 4), the other genes in this cluster are expected to be potential candidate genes in fat-tail formation. Interestingly, most of the connected genes with *IL-6* were reported to be involved in fat deposition processes, which reinforce the importance of light blue sub-network in fat deposition. For example, retinol binding protein 4 (*RBP4*), which is a novel adipokine, is mainly secreted by adipocytes and is related to obesity. This gene contributes to systemic insulin resistance that can lead to fat deposition<sup>68</sup>. Leptin receptor (*LEPR*), as other gene in light blue sub-network that is also located in fatness-related QTLs regions, was reported as candidate gene affecting fat deposition in pig<sup>69</sup>. Higher levels of plasminogen activator inhibitor-1 (*PAI-1*) in plasma was reported to be a biochemical marker of obesity<sup>70</sup>. The association between *PAI-1* and fat deposition had been well established in animal studies and the location of this gene in the QTLs linked to sheep fatness make it interesting for further functional investigations<sup>71</sup>.

*EPHX1*, *HSD11B1*, *FMO2*, *S100A1* and *HSPA6* were some of the important up-regulated genes in orange sub-network (Fig. 4). Association between *EPHX1* and adipogenesis in mesenchymal stem cells through the activation of cryoprotective lipid mediators had been explained previously<sup>72</sup>. Rosu-Myles et al., reported that *EPHX1* expressing cells in human stromal cultures can be led to increased numbers of cells that have committed to the adipocyte lineage<sup>73</sup>. *FMO2* is a member of the *FMO* gene family and catalyze the *NADPH*-dependent oxidative metabolism of a wide array of foreign chemicals as well as is involved in fat deposition, adipogenesis and fatty acid biosynthesis<sup>11</sup>. *FMO3*, as other member of this family, was reported as important candidate in fat metabolism of sheep through inhibiting fatty acid oxidation<sup>13</sup>. Consistent with these observations, *FMOs* 1, 2, and 4 knockout mice exhibited a lean phenotype and stored less triglycerides in their white adipose tissue compared to wild-type mice, despite similar food intake<sup>74</sup>. *HSD11B1* is well known to be closely related to the accumulation of abdominal fat<sup>75</sup>. Recently, an RNA-Seq study was performed to explore the effects of castration on fat deposition in different parts of pigs and *HSD11B1* was reported as a factor affects glucose uptake by adipocytes and leads to obesity<sup>76</sup>. In Arora et al., study *HSPA6* was reported as an important regulator of fatty acid metabolism in the skeletal muscles of sheep<sup>33</sup>. Adipogenesis regulatory factor (*ADIRF*), as other important up-regulated gene, promotes adipocyte differentiation by enhancing the expression of peroxisome proliferator-

activated receptor gamma (*PPARG*) and CCAAT-enhancer-binding protein alpha (*CEBPA*) in 3T3 L1 cells and play an important role in fat cell development. Higher expression of this gene in obese individuals, suggesting a role for *ADIRF* in the development of obesity<sup>77</sup>. Angiopoietin-like 8 (*ANGPTL8*) exhibits its effects by inducing insulin receptor to inhibits lipolysis and controls post-prandial fat storage in white adipose tissue. This gene directs fatty acids to adipose tissue for storage during the fed state. Serpin family E member 1 (*SERPINE1*) was highly expressed in obese individuals and demonstrated as a key gene associated with the network pathway analysis of obesity<sup>78</sup>. Altogether, these results support the potential role of up-regulated genes, especially light blue and orange sub-networks in fat deposition in tail of sheep.

Totally, 170 unique DEGs (77 up- and 93 down-regulated) were found in meta-analysis that were not identified as DEGs in individual studies (Fig. 2), which can be attributed to the more statistical power of meta-analysis than individual studies for identifying new candidate genes associated with fat-tail formation. Of which, 38 DEGs (17 up and 21 down-regulated) were located in QTLs regions related to fatness of sheep, which further support that their functions might be relevant. Some of these genes have been previously reported to be related to lipid metabolism including *NR4A1*, *ACSF2*, *MYC*, *SPP1*, *PLPP3*, *PHOSPHO1* and *ACP6*. For example, *NR4A1* encodes a nuclear receptor (transcription factor) that is involved in regulation of lipid metabolism and modulating lipolysis in muscle<sup>35</sup>. It is reported that female *NR4A1* deficient mice exhibited higher fat mass compared to wild-type mice, under the same high-fat diet<sup>79</sup>. This gene plays a vital role in the regulation of liver fat content<sup>35</sup>. Expression of *NR4A1* is reported to be negatively correlated with body-fat content and insulin sensitivity, as its expression was significantly lower in the muscle of obese men in comparison to lean men<sup>80</sup>. *MYC*, which is a transcription factor, plays vital roles in lipid metabolism<sup>81</sup>. Furthermore, *ACSF2* is involved in the acyl-CoA metabolic process and malonyl-CoA metabolic process in mammals as well as reported to be associated with avian lipid metabolism<sup>82,83</sup>. These findings suggested a link between differentially expressed of these genes and fat-tail development in sheep.

## Conclusion

In this study, a meta-analysis of six RNA-Seq datasets that compared transcriptome profiles of fat- and thin-tailed sheep breeds, were performed. By considering the expression pattern of DEGs across the different studies and performing Jackknife analysis, a list of robust DEGs were obtained, which were enriched in the QTLs related to sheep fatness. Moreover, functional interactions among the DEGs were confirmed using PPIs network analysis. Functional enrichment analysis showed that the DEGs were enriched in the GO terms/KEGG pathways related to fat metabolism such as “fatty-acyl-CoA biosynthetic process”, “ECM-receptor interaction” and “PI3K-Akt signaling pathway”. Cluster analysis of the PPIs network led to identification several sub-networks that were directly or indirectly involved in fat deposition. Some down-regulated DEGs in green and pink sub-networks were *SCD*, *SCD5*, *ELOVL6*, *ACLY*, *SLC27A2*, *LPIN1*, *COL1A1*, *COL6A6*, *COL4A6*, *COL6A3*, *ITGA1*, *ITGA2*, *SDC1*, *SPP1*, which probably promote the development of fat-tail through regulating lipolysis or fatty acid oxidation. In contrast, up-

regulated DEGs such as *IL6*, *RBP4*, *LEPR*, *PAI-1*, *CSNK2B*, *MTMR4*, *EPHX1*, *HSD11B1*, *FM02*, *S100A1* and *HSPA6* were mainly enriched in green and pink sub-networks and may contribute to a network controlling fat accumulation in tail of sheep breed through mediating adipogenesis and fatty acid biosynthesis. Overall, our meta-analysis successfully identified a core set of DEGs associated with lipid metabolism including well-known genes related to fat deposition as well as newly identified genes such as *NR4A1* and *ACSF2*. Therefore, it is reasonable to infer that the suggested sub-networks and their gene members might be potential candidate contributing to shaping fat-tail phenotype. Although our findings were in good agreement with previous studies, however they are in transcriptome level and follow-up functional studies are required to investigate the mechanisms by which these genes contribute to the fat deposition in tail of sheep breeds.

## Declarations

### Acknowledgements

We would like to thank the University of Tehran for the support of this work.

### Conflict of Interests

The authors declare no conflict of interests regarding the publication of this paper.

### Contributions

M.R.B conceived the ideas. M.R.B and SF.H designed study and analyzed the data. M.R.B, SF.H and A.S interpreted the data and wrote the main manuscript text. All authors read and approved the final manuscript.

### Data availability

The datasets generated and/or analysed during the current study are available in the NCBI's sequence read archive database as follow:

Datasets 1: <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA508203/>

Dataset 2: <https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA699984>

Dataset 3: <https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA432669>

Dataset 4: <https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA745517>

Dataset 5: <https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA517348>

Dataset 6: <https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA598581>

# References

1. Kalds, P. *et al.* Trends towards revealing the genetic architecture of sheep tail patterning: Promising genes and investigatory pathways. *Animal Genetics* **52**, 799-812 (2021).
2. Beynon, S. E. *et al.* Population structure and history of the Welsh sheep breeds determined by whole genome genotyping. *BMC genetics* **16**, 1-15 (2015).
3. Muigai, A. W. & Hanotte, O. The origin of African sheep: archaeological and genetic perspectives. *African Archaeological Review* **30**, 39-50 (2013).
4. Atti, N. & Hamouda, M. B. Relationships among carcass composition and tail measurements in fat-tailed Barbarine sheep. *Small Ruminant Research* **53**, 151-155 (2004).
5. Bakhtiarizadeh, M. R., Salehi, A., Alamouti, A. A., Abdollahi-Arpanahi, R. & Salami, S. A. Deep transcriptome analysis using RNA-Seq suggests novel insights into molecular aspects of fat-tail metabolism in sheep. *Scientific reports* **9**, 1-14 (2019).
6. Eryavuz, A., Avci, G., Kuculkurt, I. & Fidan, A. Comparison of plasma leptin, insulin and thyroid hormone concentrations and some biochemical parameters between fat-tailed and thin-tailed sheep breeds. *Revue de médecine vétérinaire* **158**, 244-249 (2007).
7. Bakhtiarizadeh, M. R., Moradi-Shahrbabak, M. & Ebrahimie, E. Underlying functional genomics of fat deposition in adipose tissue. *Gene* **521**, 122-128 (2013).
8. Fan, H. *et al.* A transcriptomic study of the tail fat deposition in two types of Hulun Buir sheep according to tail size and sex. *Animals* **9**, 655 (2019).
9. Farhadi, S., Shodja Ghias, J., Hasanpur, K., Mohammadi, S. A. & Ebrahimie, E. Molecular mechanisms of fat deposition: IL-6 is a hub gene in fat lipolysis, comparing thin-tailed with fat-tailed sheep breeds. *Archives Animal Breeding* **64**, 53-68 (2021).
10. Li, B. *et al.* Transcriptome analysis of adipose tissues from two fat-tailed sheep breeds reveals key genes involved in fat deposition. *BMC genomics* **19**, 1-13 (2018).
11. Ma, L. *et al.* Comparative transcriptome profiling of mRNA and lncRNA related to tail adipose tissues of sheep. *Frontiers in genetics* **9**, 365 (2018).
12. Miao, X., Luo, Q., Qin, X., Guo, Y. & Zhao, H. Genome-wide mRNA-seq profiling reveals predominant down-regulation of lipid metabolic processes in adipose tissues of Small Tail Han than Dorset sheep. *Biochemical and biophysical research communications* **467**, 413-420 (2015).
13. Wang, X. *et al.* Transcriptome profile analysis of adipose tissues from fat and short-tailed sheep. *Gene* **549**, 252-257 (2014).
14. Yuan, Y.-N., Liu, W.-Z., Liu, J.-H., Qiao, L.-Y. & Wu, J.-L. Cloning and ontogenetic expression of the uncoupling protein 1 gene UCP1 in sheep. *Journal of applied genetics* **53**, 203-212 (2012).
15. Yuan, Z. *et al.* Integrative analysis of Iso-Seq and RNA-seq data reveals transcriptome complexity and differentially expressed transcripts in sheep tail fat. *PeerJ* **9**, e12454 (2021).
16. Zhang, W. *et al.* Comparative Transcriptome Analysis of Key Genes and Pathways Activated in Response to Fat Deposition in Two Sheep Breeds With Distinct Tail Phenotype. *Frontiers in Genetics*

- 12, 639030 (2021).
17. Kumar, B., Francis, S. M., Suttie, J. M. & Thompson, M. P. Expression of obese mRNA in genetically lean and fat selection lines of sheep. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* **120**, 543-548 (1998).
  18. Gao, L. *et al.* Cloning of sheep cell death-inducing DFFA-like effector c (CIDEc) cDNA and its differential expression in tail fat tissue of Altay sheep (*Ovis aries*) in persistent starvation. *J. Agricul. Biotechol* **23**, 227-235 (2015).
  19. Guangli, Y. *et al.* Identification of the genetic basis for the large-tailed phenotypic trait in Han sheep through integrated mRNA and miRNA analysis of tail fat tissue samples. (2020).
  20. Su, X., He, H., Fang, C., Liu, L. & Liu, W. Screening and Identification of The Key LncRNAs Associated With Fat Deposition in Ovine Tail. (2021).
  21. Brown, J., Pirrung, M. & McCue, L. A. FQC Dashboard: integrates FastQC results into a web-based, interactive, and extensible FASTQ quality control tool. *Bioinformatics* **33**, 3137-3139 (2017).
  22. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120 (2014).
  23. Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature biotechnology* **37**, 907-915 (2019).
  24. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *bioinformatics* **31**, 166-169 (2015).
  25. Rau, A., Marot, G. & Jaffrézic, F. Differential meta-analysis of RNA-seq data from multiple studies. *BMC bioinformatics* **15**, 1-10 (2014).
  26. Rogic, S., Wong, A. & Pavlidis, P. Meta-Analysis of Gene Expression Patterns in Animal Models of Prenatal Alcohol Exposure Suggests Role for Protein Synthesis Inhibition and Chromatin Remodeling. *Alcoholism: Clinical and Experimental Research* **40**, 717-727 (2016).
  27. Hu, Z.-L., Park, C. A., Wu, X.-L. & Reecy, J. M. Animal QTLdb: an improved database tool for livestock animal QTL/association data dissemination in the post-genome era. *Nucleic acids research* **41**, D871-D879 (2013).
  28. Das, S., Rai, A., Mishra, D. C. & Rai, S. N. Statistical approach for gene set analysis with trait specific quantitative trait loci. *Scientific reports* **8**, 1-12 (2018).
  29. Wagner, R. A. & Fischer, M. J. The string-to-string correction problem. *Journal of the ACM (JACM)* **21**, 168-173 (1974).
  30. Merkin, J., Russell, C., Chen, P. & Burge, C. B. Evolutionary dynamics of gene and isoform regulation in Mammalian tissues. *Science* **338**, 1593-1599 (2012).
  31. Chan, E. T. *et al.* Conservation of core gene expression in vertebrate tissues. *Journal of biology* **8**, 1-17 (2009).
  32. Skinnider, M. A. & Foster, L. J. Meta-analysis defines principles for the design and analysis of co-fractionation mass spectrometry experiments. *Nature Methods* **18**, 806-815 (2021).

33. Arora, R. *et al.* Transcriptome profiling of longissimus thoracis muscles identifies highly connected differentially expressed genes in meat type sheep of India. *PloS one* **14**, e0217461 (2019).
34. Xiao, L. *et al.* The Novel-miR-659/SPP1 Interaction Regulates Fat Deposition in Castrated Male Pigs. *Animals* **12**, 944 (2022).
35. Zhang, L. *et al.* The orphan nuclear receptor 4A1: a potential new therapeutic target for metabolic diseases. *Journal of Diabetes Research* **2018** (2018).
36. Zhang, W. *et al.* Comparative transcriptome analysis of key genes and pathways activated in response to fat deposition in two sheep breeds with distinct tail phenotype. *Frontiers in genetics* **12** (2021).
37. Moharrery, A. Effect of docking and energy of diet on carcass fat characteristics in fat-tailed Badghisian sheep. *Small Ruminant Research* **69**, 208-216 (2007).
38. Mihelic, R. *et al.* Genes controlling polyunsaturated fatty acid synthesis are developmentally regulated in broiler chicks. *British poultry science* **61**, 508-517 (2020).
39. Li, Y. *et al.* Suppression of adipocyte differentiation and lipid accumulation by stearidonic acid (SDA) in 3T3-L1 cells. *Lipids in health and disease* **16**, 1-10 (2017).
40. Shi, H. *et al.* Fatty acid elongase 6 plays a role in the synthesis of long-chain fatty acids in goat mammary epithelial cells. *Journal of dairy science* **100**, 4987-4995 (2017).
41. Chypre, M., Zaidi, N. & Smans, K. ATP-citrate lyase: a mini-review. *Biochemical and biophysical research communications* **422**, 1-4 (2012).
42. Zhao, S. *et al.* ATP-citrate lyase controls a glucose-to-acetate metabolic switch. *Cell reports* **17**, 1037-1052 (2016).
43. Péterfy, M., Harris, T. E., Fujita, N. & Reue, K. Insulin-stimulated interaction with 14-3-3 promotes cytoplasmic localization of lipin-1 in adipocytes. *Journal of Biological Chemistry* **285**, 3857-3864 (2010).
44. Baumgard, L., Hausman, G. & Fernandez, M. S. Insulin: pancreatic secretion and adipocyte regulation. *Domestic animal endocrinology* **54**, 76-84 (2016).
45. Feng, Y. *et al.* FOXM1 as a prognostic biomarker promotes endometrial cancer progression via transactivation of SLC27A2 expression. *International Journal of Clinical and Experimental Pathology* **11**, 3846 (2018).
46. Jiang, Z. *et al.* Discovery of novel genetic networks associated with 19 economically important traits in beef cattle. *International journal of biological sciences* **5**, 528 (2009).
47. Xing, S. *et al.* Time Course Transcriptomic Study Reveals the Gene Regulation During Liver Development and the Correlation With Abdominal Fat Weight in Chicken. *Frontiers in Genetics*, 1593 (2021).
48. Wu, X. *et al.* Differential expression of six chicken genes associated with fatness traits in a divergently selected broiler population. *Molecular and cellular probes* **30**, 1-5 (2016).

49. Luther, J. *et al.* Fra-2/AP-1 controls adipocyte differentiation and survival by regulating PPAR $\gamma$  and hypoxia. *Cell Death & Differentiation* **21**, 655-664 (2014).
50. Binh, T. D., Pham, T. L., Men, T. T., Dang, T. T. & Kamei, K. LSD-2 dysfunction induces dFoxO-dependent cell death in the wing of *Drosophila melanogaster*. *Biochemical and biophysical research communications* **509**, 491-497 (2019).
51. Romer, L. H., Birukov, K. G. & Garcia, J. G. Focal adhesions: paradigm for a signaling nexus. *Circulation research* **98**, 606-616 (2006).
52. Khan, T. *et al.* Metabolic dysregulation and adipose tissue fibrosis: role of collagen VI. *Molecular and cellular biology* **29**, 1575-1591 (2009).
53. Thomas, K., Engler, A. J. & Meyer, G. A. Extracellular matrix regulation in the muscle satellite cell niche. *Connective tissue research* **56**, 1-8 (2015).
54. Henegar, C. *et al.* Adipose tissue transcriptomic signature highlights the pathological relevance of extracellular matrix in human obesity. *Genome biology* **9**, 1-32 (2008).
55. Lee, H.-J. *et al.* Comparative transcriptome analysis of adipose tissues reveals that ECM-receptor interaction is involved in the depot-specific adipogenesis in cattle. *PloS one* **8**, e66267 (2013).
56. Casado-Díaz, A. *et al.* Transcriptomic analyses of adipocyte differentiation from human mesenchymal stromal-cells (MSC). *Journal of cellular physiology* **232**, 771-784 (2017).
57. Mauney, J. & Volloch, V. Human bone marrow-derived stromal cells show highly efficient stress-resistant adipogenesis on denatured collagen IV matrix but not on its native counterpart: implications for obesity. *Matrix Biology* **29**, 9-14 (2010).
58. McCulloch, L. J. *et al.* COL6A3 is regulated by leptin in human adipose tissue and reduced in obesity. *Endocrinology* **156**, 134-146 (2015).
59. Liu, J. *et al.* Changes in integrin expression during adipocyte differentiation. *Cell metabolism* **2**, 165-177 (2005).
60. Huang, W. *et al.* Global transcriptome analysis identifies differentially expressed genes related to lipid metabolism in Wagyu and Holstein cattle. *Scientific Reports* **7**, 1-11 (2017).
61. Choi, E. O. *et al.* Zanthoxylum schinifolium leaf ethanol extract inhibits adipocyte differentiation through inactivation of the extracellular signal regulated kinase and phosphoinositide 3-kinase/Akt signaling pathways in 3T3-L1 pre-adipocytes. *Molecular Medicine Reports* **12**, 1314-1320 (2015).
62. Xu, Y. *et al.* Panax notoginseng saponins modulate the gut microbiota to promote thermogenesis and beige adipocyte reconstruction via leptin-mediated AMPK $\alpha$ /STAT3 signaling in diet-induced obesity. *Theranostics* **10**, 11302 (2020).
63. Su, X. *et al.* PI3K/Akt pathway expression in children with different obesity degrees and its relationship with glucolipid metabolism and insulin resistance. *American Journal of Translational Research* **13**, 6592 (2021).
64. Fei, X. *et al.* Transcriptome reveals key microRNAs involved in fat deposition between different tail sheep breeds. *PloS one* **17**, e0264804 (2022).

65. Chapman, J. *et al.* Osteopontin is required for the early onset of high fat diet-induced insulin resistance in mice. *PloS one* **5**, e13959 (2010).
66. Chen, Q. *et al.* An Osteopontin-Integrin Interaction Plays a Critical Role in Directing Adipogenesis and Osteogenesis by Mesenchymal Stem Cells. *Stem Cells* **32**, 327-337, doi:10.1002/stem.1567 (2014).
67. Cuzzone, D. A. *et al.* IL-6 regulates adipose deposition and homeostasis in lymphedema. *American Journal of Physiology-Heart and Circulatory Physiology* **306**, H1426-H1434 (2014).
68. Cheng, J. *et al.* Ectopic expression of RBP4 impairs the insulin pathway and inguinal fat deposition in mice. *Journal of physiology and biochemistry* **70**, 479-486 (2014).
69. Xing, K. *et al.* Identification of key genes affecting porcine fat deposition based on co-expression network analysis of weighted genes. *Journal of animal science and biotechnology* **12**, 1-16 (2021).
70. Barnard, S. A., Pieters, M. & De Lange, Z. The contribution of different adipose tissue depots to plasma plasminogen activator inhibitor-1 (PAI-1) levels. *Blood reviews* **30**, 421-429 (2016).
71. Ekström, M., Liska, J., Eriksson, P., Sverremark-Ekström, E. & Tornvall, P. Stimulated in vivo synthesis of plasminogen activator inhibitor-1 in human adipose tissue. *Thrombosis and haemostasis* **108**, 485-492 (2012).
72. Tomita, Y. *et al.* Cardiac neural crest cells contribute to the dormant multipotent stem cell in the mammalian heart. *The Journal of cell biology* **170**, 1135-1146 (2005).
73. Rosu-Myles, M. *et al.* Identification of a candidate proteomic signature to discriminate multipotent and non-multipotent stromal cells. *PLoS One* **7**, e38954 (2012).
74. Veeravalli, S. *et al.* The phenotype of a flavin-containing monoxygenase knockout mouse implicates the drug-metabolizing enzyme FMO1 as a novel regulator of energy balance. *Biochemical pharmacology* **90**, 88-95 (2014).
75. Baudrand, R. *et al.* Overexpression of 11 $\beta$ -hydroxysteroid dehydrogenase type 1 in hepatic and visceral adipose tissue is associated with metabolic disorders in morbidly obese patients. *Obesity surgery* **20**, 77-83 (2010).
76. Liu, X. *et al.* Integrated Analysis of the ceRNA Network and M-7474 Function in Testosterone-Mediated Fat Deposition in Pigs. *Genes* **13**, 668 (2022).
77. Ni, Y. *et al.* A Novel pro-adipogenesis factor abundant in adipose tissues and over-expressed in obesity acts upstream of PPAR $\gamma$  and C/EBP $\alpha$ . *Journal of bioenergetics and biomembranes* **45**, 219-228 (2013).
78. Kaur, P. *et al.* SERPINE 1 links obesity and diabetes: a pilot study. *Journal of proteomics & bioinformatics* **3**, 191 (2010).
79. Perez-Sieira, S. *et al.* Female Nur77-deficient mice show increased susceptibility to diet-induced obesity. *PLoS One* **8**, e53836 (2013).
80. Kanzleiter, T. *et al.* Overexpression of the orphan receptor Nur77 alters glucose metabolism in rat muscle cells and rat muscle in vivo. *Diabetologia* **53**, 1174-1183 (2010).

81. Wahlström, T. & Henriksson, M. A. Impact of MYC in regulation of tumor cell metabolism. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms* **1849**, 563-569 (2015).
82. Tian, W. *et al.* Dynamic expression profile, regulatory mechanism and correlation with egg-laying performance of ACSF gene family in chicken (*Gallus gallus*). *Scientific reports* **8**, 1-10 (2018).
83. Watkins, P. A., Maiguel, D., Jia, Z. & Pevsner, J. Evidence for 26 distinct acyl-coenzyme A synthetase genes in the human genomes. *Journal of lipid research* **48**, 2736-2750 (2007).

## Figures

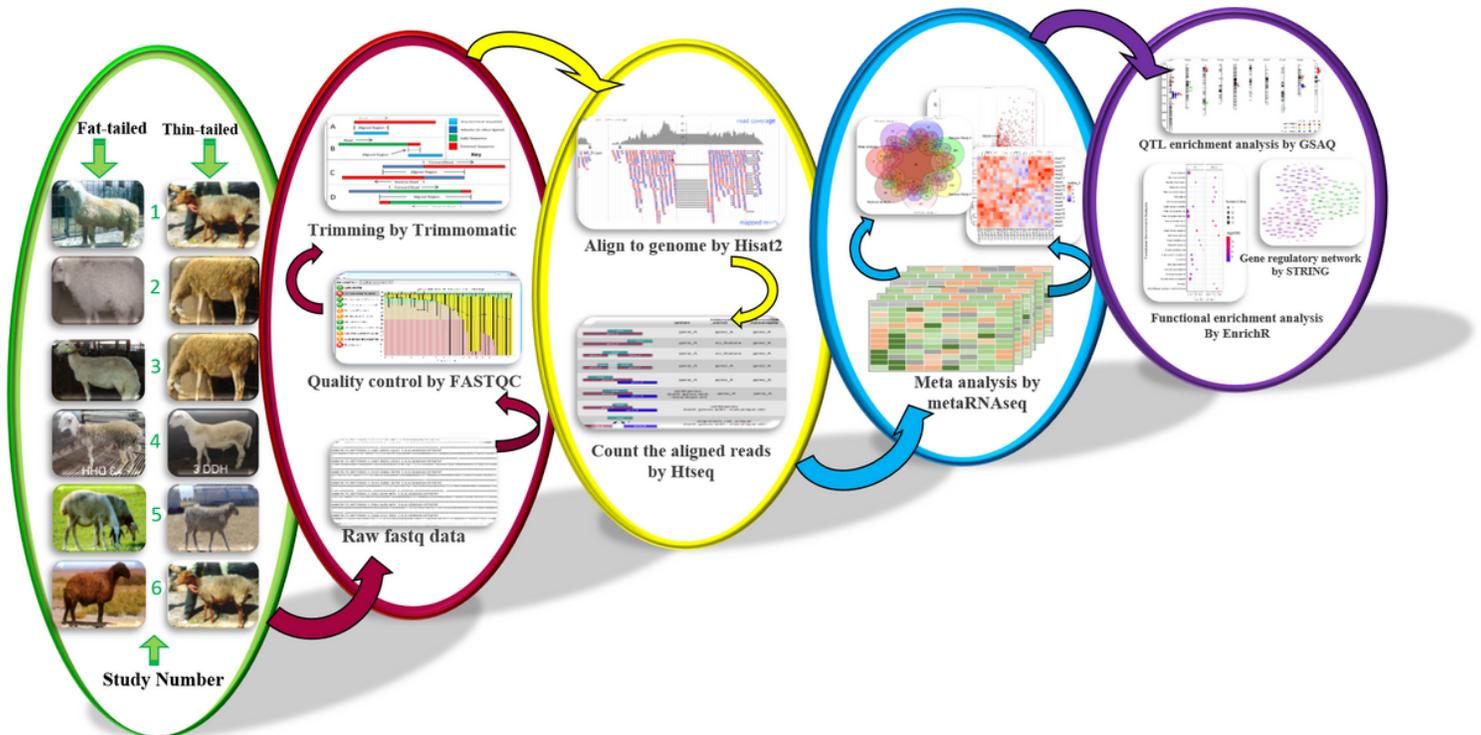
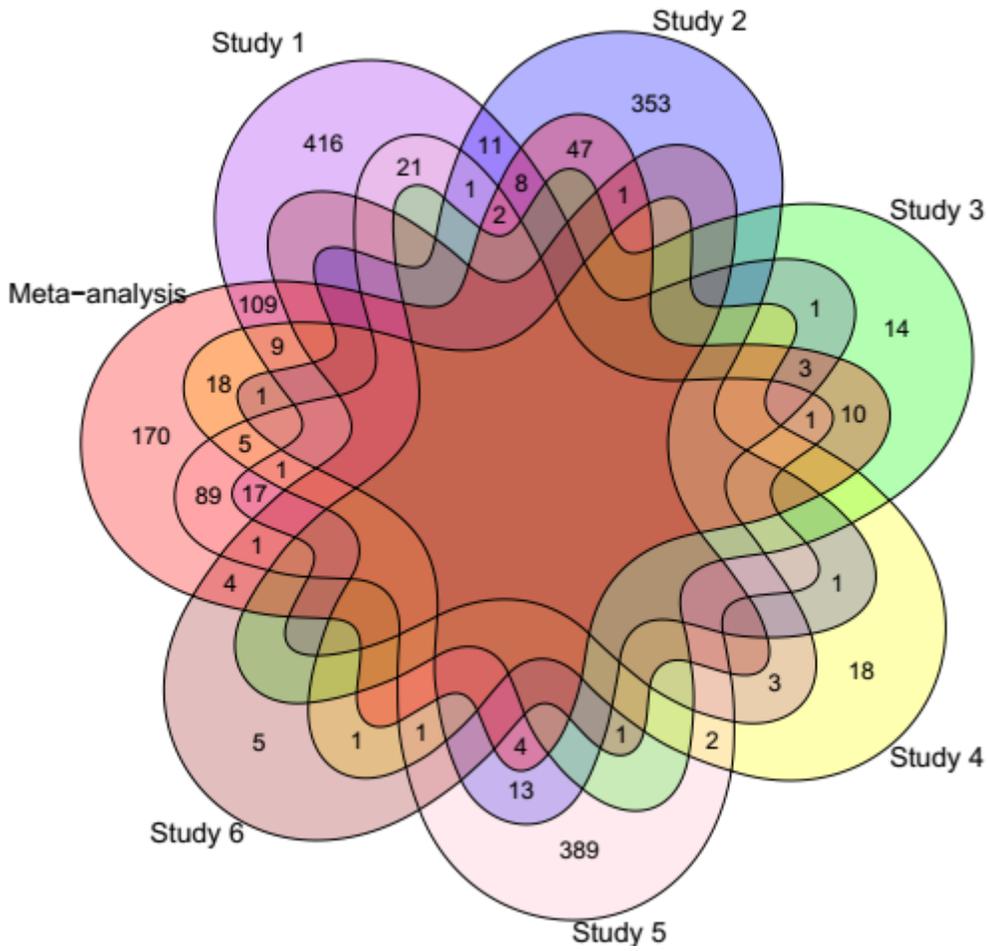


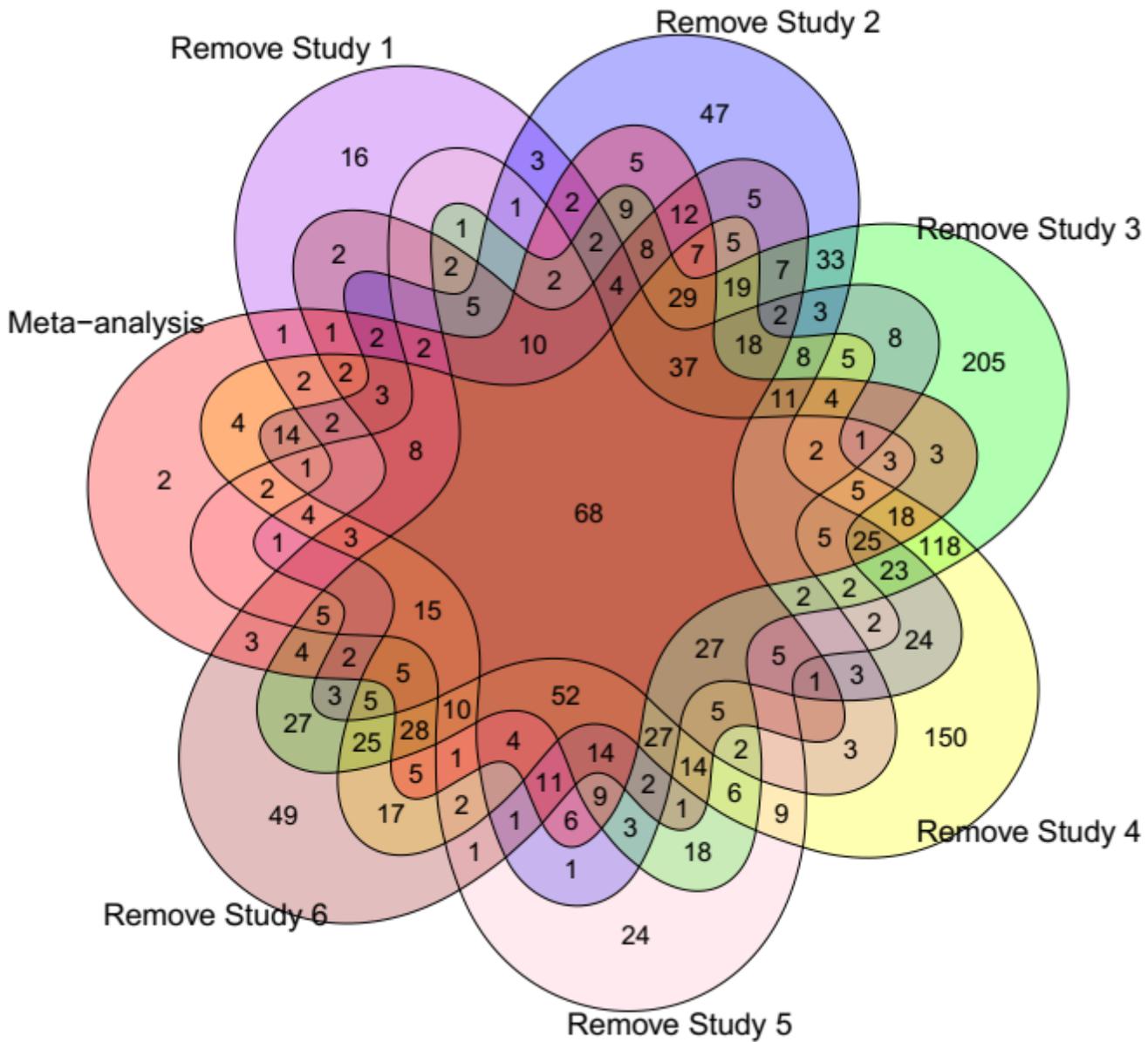
Figure 1

The bioinformatics pipeline for RNA-Seq meta-analysis of fat-tail tissue in sheep breeds



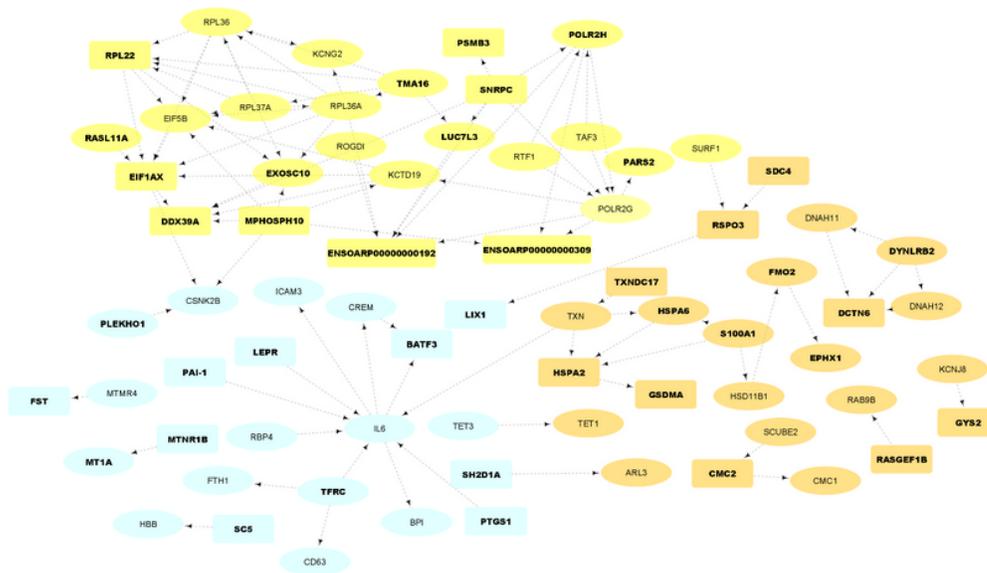
**Figure 2**

Venn diagram showing the number of common DEGs shared among the individual studies and meta-analysis.



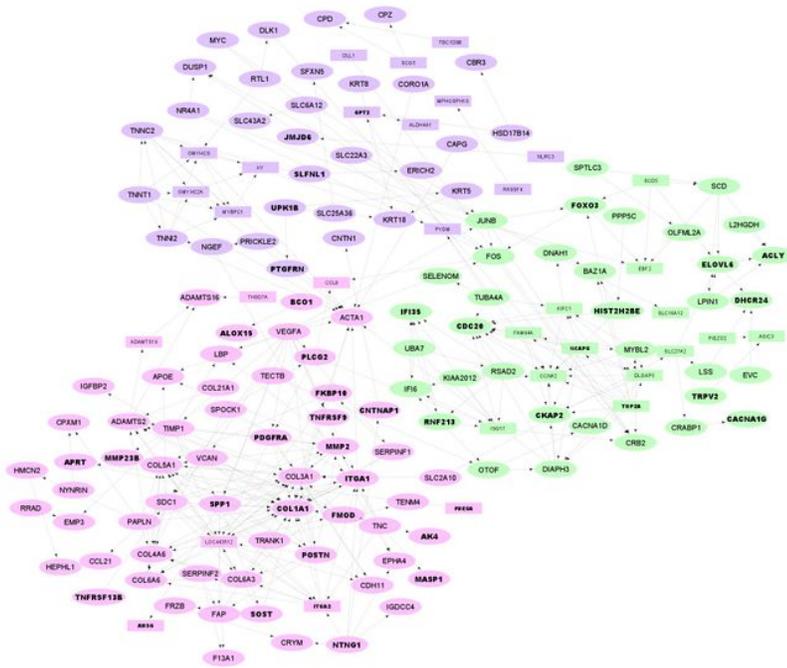
**Figure 3**

Venn diagram showing the number of common DEGs among different Jackknife tests and meta-analysis results.



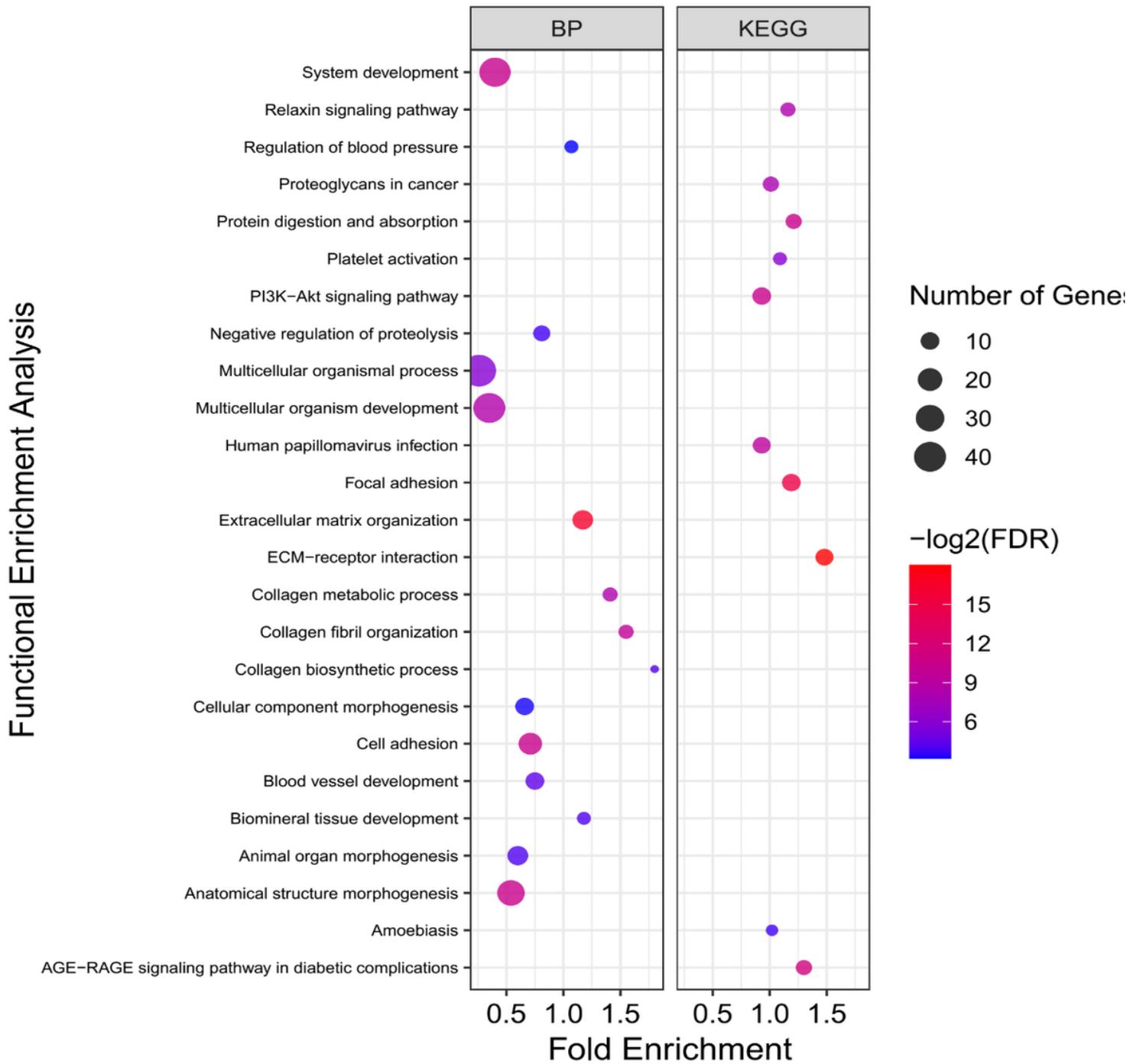
**Figure 4**

PPI network and functional clusters of the up-regulated DEGs. Orange, yellow and light blue nodes represent orange, yellow and light blue sub-networks, respectively. Oval circles indicate the DEGs that were presented in at least 50% of the Jackknife test results. The DEGs that were located in QTL regions related to fatness are represented as bold texts.



**Figure 5**

PPI network and functional clusters of the down-regulated DEGs. Green, pink and purple nodes represent green, pink and purple clusters, respectively. Oval circles indicate the DEGs that were presented in at least 50% of the Jackknife test results. The DEGs that were located in QTL regions related to fatness are represented as bold texts.



**Figure 6**

Functional enrichment analysis results of pink sub-network. Size and color of points represent  $-\log_2$  of FDR and number of genes associated with each term, respectively. BP= Biological processes.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryFileS1.xlsx](#)
- [SupplementaryFileS2.xlsx](#)
- [SupplementaryFileS3.xlsx](#)
- [SupplementaryFileS4.xlsx](#)