Allele-specific expression reveals the phenotypic differences between thin- and fat-tailed sheep

Junjie Shao

Northwest Agriculture and Forestry University

Sangang He

Ministry of agriculture; Institute of biotechnology, Xinjiang academy of animal science

Xiangyu Pan

Northwest Agriculture and Forestry University

Zhirui Yang

Northwest Agriculture and Forestry University

Hojjat Asadollahpour Nanaei

Northwest Agriculture and Forestry University

Lei Chen

Ministry of agriculture; Institute of biotechnology, Xinjiang academy of animal science

Ran Li

Northwest Agriculture and Forestry University

Yu Wang

Northwest Agriculture and Forestry University

Shan Gao

Northwest Agriculture and Forestry University

Han Xu

Northwest Agriculture and Forestry University

Mingjun Liu

Ministry of agriculture; Institute of biotechnology, Xinjiang academy of animal science

Yu Jiang (✉ yu.jiang@nwafu.edu.cn)

Northwest Agriculture and Forestry University

Research

Keywords: allele-specific expression, phenotypic difference, thin- and fat-tailed sheep, whole-genome sequencing, transcriptome

Posted Date: August 13th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-56388/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

1 **Allele-specific expression reveals the phenotypic differences between thin- and fat-**
2 **tailed sheep**

3 Junjie Shao ^{1†}, Sangang He ^{2,3†}, Xiangyu Pan ^{1†}, Zhirui Yang¹, Hojjat Asadollahpour
4 Nanaei¹, Chen Lei ^{2,3}, Ran Li¹, Yu Wang¹, Shan Gao¹, Han Xu¹, Mingjun Liu ^{2,3*} and
5 Yu Jiang ^{1*}

6
7 ¹College of Animal Science and Technology, Northwest A&F University, Yangling,
8 Shaanxi 712100, China, ²Key Laboratory of Ruminant Genetics, Breeding &
9 Reproduction, Ministry of Agriculture, ³Key Laboratory of Animal Biotechnology of
10 Xinjiang, Institute of Biotechnology, Xinjiang Academy of Animal Science, Urumqi,
11 Xinjiang 830026, China.

12
13 † These author contributed equally to this work.

14 * Corresponding author: Yu Jiang yu.jiang@nwafu.edu.cn;

15 MingJun Liu xjlmj2006@126.com.

16

17

18

19

20

21

22

Abstract

Background: The thin-tailed sheep breeds from Europe and the fat-tailed sheep breeds from China exhibit distinct phenotypic differences in fat deposition and meat production traits. However, the molecular mechanisms underlying gene expression related to these phenotypic differences are not well understood. Allele-specific expression (ASE) refers to the significant imbalance of expression levels of two parental alleles. Characterization of such events in F1 hybrid offspring generated from these two groups of sheep breeds can minimize the external factors influencing gene expression and reveal the variants with a *cis*-regulatory effect on gene expression. The aim of the present study was to investigate the genetic factors that influence different fat-deposition and meat production traits between thin- and fat-tailed sheep.

Results: Fifteen F1 hybrids were generated from crosses between Texel and Kazakh sheep as the representative phenotypes of thin- and fat-tailed breeds, respectively. Totally, 33 whole genomes from F1 individuals and their parents were sequenced with an average depth of $\sim 17.21\times$ coverage per sample. ASE analysis results from 70 RNA-seq samples of adipose and skeleton muscle tissues showed 128 ASE candidate genes were related to the function of fat deposition and meat production traits. A genome-wide scan of selective sweeps was also conducted between these two groups of sheep breeds in an effort to identify genomic regions related to fat deposition and meat production, respectively. We detected signatures of selection in ASE genes associated with fat deposition (e.g., *PDGFD*) and meat production traits (e.g., *LRCC2*). Further analysis suggested that *PDGFD* and *LRCC2* genes were speculated to be causative

genes for fat deposition and meat production traits in sheep, respectively. Furthermore, AMPK signaling pathway was significantly enriched in ASE genes related to fatty acid biosynthesis in both adipose and skeleton muscle tissues, while PPAR signaling pathway was significantly enriched in ASE genes related to lipid metabolism in adipose tissue.

Conclusions: Our finding illustrates that the expression of identified ASE genes could potentially lead to the differences in traits of fat deposition and meat production between thin- and fat-tailed sheep.

Keywords: allele-specific expression, phenotypic difference, thin- and fat-tailed sheep, whole-genome sequencing, transcriptome

Background

The sheep (*Ovis aries*) was one of the first grazing animals to be domesticated approximately 11,000 years ago in the Fertile Crescent [1]. Since domestication, a wide variety of sheep breeds with different production traits (e.g. meat, milk and fiber) and morphological characteristics (e.g. coat color and tail shape) have been shaped under the influence of natural and artificial selection [2]. According to the fat content in tail, sheep can be mainly classified into two groups: thin- and fat- tailed breeds. Fat tailed sheep breeds comprise approximately 25% of the world sheep population, which are first documented approximately 5,000 years ago [2]. They are generally well adapted to the dryland environments with seasonal poor nutrition [3, 4]. These sheep breeds are known for their ability to deposit up to 20% of the carcass weight as tail fat [5]. The tail fat plays as a valuable source of energy during migration and winter. However, nowadays with increasing attention to human health, fat is an unpopular constituent of meat for consumers. On the other side, the energy cost for producing fat is higher than an equivalent amount for lean tissue [6]. Therefore, a small size of fat is often favorable for both consumers and producers. Thus, identification of functional candidate genes and selective genomic regions associated with fat deposition and meat production traits is more important in sheep breeding programs.

Previous genome-wide association studies (GWAS) revealed candidate genes related to fat deposition and meat production traits in different sheep breeds. These studies have proposed candidate genes, such as *MSL1*, *PFKFB4*, *TRDN*, *FBF1*, *SETD7* and *GRM1* [7, 8]. On the other side, genomic scan for detecting regions associated with

fat deposition were also performed in several sheep breeds [9-17]. However, the molecular mechanisms underlying gene expression related to such economically important traits remained mostly unexplored. DNA sequence variation can lead to changes in gene expression levels, which is a main cause of phenotype diversity across individuals or populations [18]. Variation in the gene expression can be due to both genetic and non-genetic factors. Allele-specific expression (ASE) refers to two or more alleles at the same loci with imbalanced expression, which is one of the important genetic factors that lead to phenotypic variation in organisms [19-22]. ASE is essential for normal development, cellular programming and many other cellular processes [23]. Therefore, identification of loci involved in this phenomenon is important in developmental biology and genetics. The expression ratios of the parental alleles obtained from RNA-seq data of hybrid offspring can be used as a reliable proxy for ASE. Recently, high-throughput sequencing technologies allowed us to identify ASE genes at a transcriptome-wide level. Several studies in different species such as mouse [24], pig [25], cow [26], goat [27], sheep [28], mule [29] and dzo [29] have shown that the different expression levels of alleles due to ASE may lead to variation in phenotypes.

In this study, to understand the genetic mechanisms underlying fat-deposition and meat production traits differences between thin- and fat-tailed sheep breeds, we chose Texel and Kazakh as the representative of two sheep breeds with thin- and fat-tailed, respectively. By integrating the findings from analysis of selective sweeps and gene expression data, several ASE genes associated with growth and development of fat deposition and meat production traits have been detected. Our results suggested that

identified ASE genes may contribute to phenotypic diversity between these two sheep breeds.

Methods

Sample collection and DNA Sequencing

Whole blood samples (10 ml) were collected from three Texel rams, fifteen Kazakh ewes and their fifteen F1 hybrids in Yining City of Xinjiang, China. DNA was isolated from whole blood using standard phenol-chloroform method [30]. The quality of genomic DNA was measured by NanoDrop 2000 (Thermo Fisher Scientific Inc., Wilmington, DE, USA). The total amount of genomic DNA from each sample was at least 8 ug and the concentrations of DNA was higher than 50 ng/ul. Paired-end sequencing libraries were conducted using the Illumina HiSeq2000 platform.

SNP genotyping and quality control

High-quality clean reads were obtained after removing adaptors and low-quality reads. Pair-end reads passing the filtering were aligned to the *Ovis aries* v4.0 reference genome using BWA software (v0.7.13, bio-bwa.sourceforge.net) [31, 32]. Heterozygous sites from F1 hybrid individuals were identified by applying the Genome analysis toolkit (GATK, v3.3-2) [33]. Low-quality variants were filtered using the parameter $QUAL < 30$, and the variants were annotated using ANNOVAR (Version: 2013-08-23) [34].

ASE analysis from RNA-seq data

A total of 70 transcriptome data generated from adipose tissue including subcutaneous fat (n=14), tail fat (n=15) and from skeleton muscle tissue including longissimus dorsi (n=13), semimembranosus (n=15), semitendinosus, (n=13) of the F1 hybrid animals were used to identify the SNP with ASE [35]. All BAM files from these samples were then generated by mapping RNA-Seq data to the *Ovis aries* v4.0 reference genome, by using HISAT2 software [36]. Allele-specific reads at the heterozygous sites were obtained from these BAM files using the WASP software [37]. In order to eliminate the error caused by hetero-allele or repeat-sequence of the reference genome, loci of both alleles with < 10 reads were filtered out, and genes with min 2 and max 100 heterozygous loci were selected [28]. Loci with both alleles < 100 reads and each allele in heterozygous sites with < 3 reads were filtered, which can remove loci exhibiting MAE and a relative lower expression [38-40]. Allelic read ratios were counted by a Python Script [27], significant allelic imbalances were identified using the cut-off criteria of allele ratio ≥ 0.65 or ≤ 0.35 [29], and statistical tests were conducted by *Chi* square test and Benjamini-Hochberg FDR correction. Imprinted genes were excluded according to the online database (<http://www.geneimprint.com/site/home>).

Functional enrichment analysis

We then employed KOBAS (v3.0) online tool (<http://kobas.cbi.pku.edu.cn/>) to conduct ASE genes enrichment analysis. Reference genes of *Ovis aries* v4.0 genome were used as background gene list, and corrected *P*-values were calculated by *Chi*-square test and

Benjamini-Hochberg methods [32]. Statistically, over-represented categories were determined by standard false discovery rate (adjust p -value < 0.05).

Differential expression analysis

The transcriptome data from longissimus dorsi of Texel and Kazakh sheep were downloaded and then trimmed using Trimmomatic to remove low quality reads [32, 35, 41]. BAM files were then obtained after mapping the high quality clean reads to the *Ovis aries* v4.0 reference genome [32], and DEGs were identified using R package DESeq 2 (version 1.16.1) [42].

Selective sweeps analysis

We downloaded a total of 52 sheep whole-genome sequencing data including 20 fat-tailed sheep of two breeds from Xinjiang, China, 15 improved thin-tailed sheep of seven breeds from Europe and 17 wild Mouflon sheep (*Ovis orientalis*) from the Middle East (Additional file 1: Table S1) [43, 44]. Selective sweeps between these thin- or fat-tailed sheep breeds and Mouflon sheep populations were scanned using a genome-wide sliding window strategy based on the SNPs with less than 10% missing data [45]. Population differentiation statistic (F_{ST}) between the selected thin- or fat-tailed sheep and wild Mouflon sheep were calculated using a 50 kb sliding window with 10 kb sliding step by vcftools v0.1.14, respectively [46]. To detect genomic regions under selection in thin- or fat-tailed sheep, we calculated the $\ln(\theta\pi_{\text{Mouflon}}/\theta\pi_{\text{thin-tailed sheep}})$ and $\ln(\theta\pi_{\text{Mouflon}}/\theta\pi_{\text{fat-tailed sheep}})$, respectively. Values of $\theta\pi$ in thin-tailed, fat-tailed sheep

and wild Mouflon sheep were calculated using the same parameters as calculating the F_{ST} by vcftools v0.1.14 [46]. In each comparison, the genomic regions in the top 1% F_{ST} values, 1% $\ln(\theta\pi_{\text{Mouflon}}/\theta\pi_{\text{thin-tailed sheep}})$ and top 1% $\ln(\theta\pi_{\text{Mouflon}}/\theta\pi_{\text{fat-tailed sheep}})$ values across the whole genome were considered to be putative selective sweeps. Combining the results of F_{ST} values and $\theta\pi$ ratio, the genes located in selective sweeps which were simply found in thin- or fat-tailed sheep used for subsequent analysis.

Results

The phenotypic differences of growth and meat production traits between Texel and Kazakh sheep breeds

Texel and Kazakh sheep were selected as parental breeds to establish F1 hybrid population. High phenotypic variations were observed between these two breeds: Texel sheep are characterized by well-developed skeleton, high mutton performance, wide and deep chest and relatively higher lean meat ratio, whereas Kazakh sheep have low meat production rate, slower growth rate and relatively higher content of fat in their carcass (Fig. 1a and 1b). In this study, we collected several phenotypic data related to fat deposition and meat production traits. From the above records, nine traits were concerned finally: weights of tail fat, longissimus dorsi, silverside, topside, rump, total fat and total meat; thickness of back fat and back muscle. As shown in Fig. 1 and Additional file 2: Table S2, the weights of tail fat, total fat and back fat thickness in Kazakh sheep are higher than those in Texel sheep. The average of tail fat weight estimated in Kazakh sheep (3252.0 g) was remarkably heavier than that of the Texel

sheep (46.71 g) (Fig. 1a, *t*-test, *P* value < 0.05). However, the average weights of longissimus dorsi, back muscle thickness, silverside and topside from Texel sheep were significantly heavier than that from the Kazakh sheep (Fig. 1f-i, *t*-test, *P* value < 0.05). The detailed differences of phenotype traits between Texel and Kazakh sheep breeds provide a rational framework in which to examine the genetic variation between these two breeds.

Identification of ASE SNPs in the F1 hybrids of Texel and Kazakh sheep

To eliminate the effects of differences in growth and development between Texel and Kazakh sheep, we compared the relative allelic expression in F1 hybrids of these two sheep breeds. Three Texel rams, fifteen Kazakh ewes and their fifteen F1 hybrid offspring were selected for whole genome re-sequencing analysis. Samples were sequenced with an average depth of $\sim 17.21\times$ coverage (Additional file 3: Table S3). We then estimated that each F1 hybrid contained an average of 7.55 Mb heterozygous SNPs, by which the two parental alleles can be distinguished (Additional file 4: Table S4). Among those heterozygous sites, about an average of 1.83% were found in exon regions of genes. In addition, we detected an average of 59.05%, 34.03% and 1.55% heterozygous sites in intergenic, intron and gene up/downstream regions of genome, respectively (Fig. 2a, Additional file 4: Table S4). A total of 810 ASE SNPs were identified from the samples of adipose or skeleton muscle tissue, 60 (7.41%) of the total ASE SNPs were located on upstream and downstream regions of genes (Fig. 2b, Additional file 4: TableS4). Totally, 605 (74.69%) of the total ASE SNPs were found

on exon regions of genes, in which 303 (50.08%) were located on the untranslated regions (UTR) of genes, 155 (25.62%) were synonymous mutation, 73 (12.07%) were non-synonymous mutation and 74 (12.23%) were unknown mutation (Fig. 2b, Additional file 5: TableS5). It has been suggested that these ASE SNPs are located on untranslated regions of genes, which can act as *cis*-regulatory elements (CREs) to regulate the transcription of a neighboring gene [47]. On the other side, the proportion of transition (77.59%) was three times more than that of the transversion (22.41%) (Fig. 2c, Additional file 4: TableS4). And the ratio of transition and transversion was 3:1, which is similar to the previous studies [48].

Analysis of ASE genes in adipose tissue

A total number of 106 ASE genes were identified in the samples of adipose tissues, in which 48 (45.28%) common ASE genes were found in both tail fat and subcutaneous fat samples. We identified 42 (39.62%) and 16 (15.09%) unique ASE genes from tail fat and subcutaneous fat samples, respectively (Fig. 3a, Additional file 5: Table S5). Additionally, we found 89 (83.96%) of 106 ASE genes have less than 10 ASE SNPs, and 17 (16.04%) of these ASE genes have more than 10 but less than 51 ASE SNPs (Fig. 3b, Additional file 5: Table S5). According to the previous report, in which the authors identified differentially expressed genes (DEGs) in samples of tail fat between thin- and fat-tailed sheep breeds, here we also found two up-regulated (*PDGFD* and *IRF2BP2*) and one down-regulated ASE genes (*TEN1*) in samples of tail fat, by comparing these two breeds [15]. *IRF2BP2* gene is located in the quantitative trait loci

(QTL) which is related to tail fat deposition, and *TEN1* gene was considered to be associated with tail fat weight trait in Hulun Buir sheep [49]. A total of two and three ASE SNPs were found in *IRF2BP2* and *TEN1* genes, respectively. The expression of ASE SNPs identified in *IRF2BP2* showing consistent in both tail fat and subcutaneous fat samples, and also for *TEN1* gene, we found the SNPs showing consistent ASE in tail fat samples (Additional file 5: Table S5).

In order to further investigate the potential functional and metabolic process of these identified ASE genes and DEGs, we performed Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment pathway analysis and found Chemokine signaling, AMPK signaling and PPAR signaling pathways were significantly enriched (Fig. 3c). Four ASE genes (*PPARG*, *PCK1*, *LPL* and *SCD* genes) and two DEGs (*LOC101115115* and *FABP5*) were significantly annotated in PPAR signaling pathway (Fig. 3d). In addition, six ASE genes (*PCK1*, *PFKFB1*, *PFKFB4*, *PPARG*, *SCD* and *SLC2A4* genes) and two DEGs (*RAB2A* and *CREB3L1*) were significantly annotated in AMPK signaling pathway. In terms of directions of ASE in both tail fat and subcutaneous fat samples, we found *PFKFB1* gene showing consistent in samples from both tail fat and subcutaneous fat, *PCK1* and *SCD* genes showing consistent in subcutaneous fat samples, while *PPARG* and *SLC2A4* genes showing inconsistent in tail fat samples. Besides this, *PFKFB4* and *LPL* genes were also found showing consistent in some specific subcutaneous fat and tail fat samples (Additional file 5: Table S5). The above results elucidated that DEGs and ASE genes which enriched in PPAR signaling and AMPK pathways may contribute to the difference of lipid metabolism in adipose tissue

between thin- and fat-tailed sheep.

Analysis of ASE genes in skeleton muscle tissue

The analysis identified a total of 47 ASE genes, in which 15 (31.91%) of them were shared by longissimus dorsi, semitendinosus and semimembranosus samples. Also, 13 (27.66%), 4 (8.51%) and 4 (8.51%) ASE genes were found in semimembranosus, semitendinosus and longissimus dorsi samples, respectively (Fig. 4a, Additional file 4: Table S4). We found that 40 (85.11%) of 45 ASE genes have less than 10 SNPs and the other 7 (14.89%) ASE genes have more than 10 but less than 30 SNPs (Fig. 4b, Additional file 5: Table S5). Compared with ASE genes identified in adipose tissue, 21 (17.07%) ASE genes were private to skeleton muscle tissue and 24 ASE genes were common in both of these two tissues (Fig. 4c, Additional file 5: Table S5). It can be seen that majority of ASE genes were more inclined towards tissue-specific. Among the identified ASE genes in skeleton muscle tissue, *EPM2A*, *LRRC2*, *PFKFB4* genes were thought to be associated with meat production traits in sheep [7]. *EPM2A* gene showing consistent in samples from semitendinosus and longissimus dorsi. *LRRC2* and *PFKFB4* genes showing consistent in some specific samples of semimembranosus, while *PFKFB4* genes showing inconsistent in any semitendinosus samples. The above analysis illustrated that some identified ASE genes showed inconsistent ASE in different tissues or different samples from the same tissue. It is suggested that one of the alleles may function better in one specific tissue or one sample type. The F1 hybrids of Texel × Kazakh sheep may able to use the better allele in growth and development,

eventually leading to higher performance in their meat production traits.

Identification of ASE genes for directly comparison between Texel and Kazakh sheep

We also downloaded the transcriptome data of longissimus dorsi of Texel and Kazakh sheep and identified DEGs between these two sheep breeds. As we expected, we found a total of 4,446 genes showing significant differential expression between skeleton muscle tissues from Texel and Kazakh breeds; including 1,988 down-regulated and 2,458 up-regulated (adjust p-value ≤ 0.05 , $\log_2$ RPKM ≥ 1 and $\log_2$ Fold change ≥ 1) (Additional file 6: Table S6). In order to investigate pathways associated with meat production traits between thin- and fat-tailed sheep breeds, we then conducted an integrated functional analysis of ASE genes identified in skeleton muscle tissues of F1 hybrids and DEGs as mentioned above. Several signaling pathways such as MAPK signaling, Rap1 signaling, AMPK signaling and Notch signaling were significantly enriched for those ASE genes and DEGs (Fig. 4d). *FASN* and *PFKFB4* ASE genes were annotated in 5'-adenosine monophosphate-activated protein kinase (AMPK) pathway, which plays a key role in the regulation of muscle regeneration and size [50]. *FASN* gene could promote the synthesis of long chain saturated fatty acids as the rate-limiting enzyme, SNPs in this gene are associated with fatty acids content in meat and fat tissues [51]. Two ASE SNP were found in *FASN* gene, which were located in the exonic region of the genes and the directions of these two ASE SNPs were consistent (Additional file 5: Table S5). Moreover, we found *FASN* gene and other eight ASE genes were down-

regulated in longissimus dorsi samples of Texel sheep, as compared to those of in Kazakh sheep (Fig. 4e). These results implied that AMPK signaling pathway is an important factor in regulation of different energy metabolism in skeleton muscle tissue between thin- and fat-tailed sheep breeds.

Investigation of ASE genes located in the selected regions

In order to investigate whether the identified ASE genes were under selection in thin- or fat-tailed sheep, we performed whole genome sequences analysis between 17 wild Mouflons sheep and 20 fat-tailed sheep from Xinjiang and 15 improved thin-tailed sheep from Europe, respectively. By combining the results from population-differentiation statistic (F_{ST}) and $\ln(\theta\pi_{\text{Mouflon}} / \theta\pi_{\text{fat-tailed sheep}})$ tests, a total of 176 putative selective genomic regions with an average size of 100.97 kb and comprised with 254 genes were found in the fat-tailed sheep (Additional file 7: Table S7). We further found six ASE genes including *PDGFD*, *PPDPF*, *PM20D1*, *TUBA8*, *PPMIK* and *PPA2* genes were selected in the fat tailed sheep breeds (Fig. 5a). As a member of the platelet derived growth factor (PDGF), *PDGFD* gene plays an important role in regulation of adipocyte function and its expression is closely related to the fat deposition in sheep tail [15, 52]. From eight samples of adipose tissue, seven ASE SNPs were found in *PDGFD* gene. Five SNPs were located in the 3'-UTR region and the remaining two SNPs were located in the exon regions. In addition, the third and fourth ASE SNP were sample-specific, but the other ones were at least with two samples (Fig. 5b, Additional file 4: Table S4). Notably, the genotype patterns located in the upstream

region of *PDGFD* gene were quite different between the wild Mouflon sheep and fat-tailed sheep (Fig. 5b).

In addition, a total number of 194 putative selective sweep regions which with an average size of 86.80 kb and comprised 256 genes were identified in the thin-tailed sheep according to the values of F_{ST} and $\ln(\theta\pi_{\text{Mouflon}}/\theta\pi_{\text{thin-tailed sheep}})$ (Additional file 8: Table S8). Out of these selected genes, we found seven genes (*LRRC2*, *BIN1*, *LOC101118164*, *IRF2BP2*, *SFRP1*, *LOC1011045430* and *BCO2*) showing ASE (Fig. 6a). *LRRC2* gene is mainly expressed in samples of longissimus dorsi, semitendinosus and semimembranosus from sheep and is located in the QTL that is associated with muscle density [49, 53]. A total of nine ASE SNPs were identified in the 3'-UTR region of *LRRC2* gene, four of these ASE SNPs were shared by at least two samples and the remaining ASE SNPs were found with sample-specific (Fig. 6b, Additional file 4: Table S4). The genotype pattern of selection sweep region located in *LRRC2* gene were different between wild Mouflon sheep and thin-tailed sheep (Fig. 6c) These results suggested that those ASE genes which were selected in the thin- or fat-tailed sheep may contribute the different fat deposition and meat production traits between thin- and fat-tailed sheep.

Discussion

The great variations in different phenotypic traits of two selected sheep breeds provide an opportunity for studying molecular mechanisms underlying gene expression related to phenotypic difference between thin- and fat-tailed sheep breeds. In this study, the

candidate causative genes were detected by ASE analysis, which avoid the effect of *trans*-acting and has become research focus in recent years [24-28]. We divided all sheep individuals into 15 family trios. Each family consists of one Texel ram, one Kazakh ewe and their F1 hybrid offspring. All these individual genomes were then re-sequenced and used to find heterozygous loci. The unbiased mapping BAM files from allele-specific reads of RNA-seq data were produced by the aids of WASP software, which can map the reads to the correct location of reference genome with a low error rate [37]. A total of 128 ASE genes were identified from the collected samples of adipose and skeleton muscle tissues, the expression profile of majority ASE genes in F1 hybrid tissues were tended to be tissue-specific. In a recent study on analysis of ASE in immune-related tissues, the authors reported individual-specific and tissue-specific gene expression for several genes in a crossbred sheep breed (Texel $\times$ Scottish Blackface) [28]. In mouse, Pinter et al. (2015) showed that 95% and 82% of the identified ASE genes in F1 hybrids and in a reciprocal cross were also with tissue specificity, respectively [54]. Two alleles in rice hybrids might play different role in different environmental conditions and developmental stages, a higher proportion of ASE genes in those hybrids showed inconsistent patterns [55]. In the current study, 22, 81 and 25 ASE genes were identified in samples from skeleton muscle, adipose and both sharing tissues, respectively. Many of these genes showed inconsistent patterns in terms of directions of ASE bias across different samples. It can be speculated that one allele may prefer to one tissue, while the other one prefer to other tissues in one case. And in the other case, one allele might be inclined to one sample, while the other one

might be inclined to other samples in the same tissue.

Comparative transcriptome analysis is another efficient way used for identifying candidate causative genes that associated with distinct phenotypic differences between thin- and fat-tailed sheep breeds [56-58]. However, it presented little power to identify genetic factors that contribute to these phenotypic differences because the identified DEGs always reflect both *cis*- and *trans*-acting regulatory variations [59]. In addition, the different growing conditions between these two groups of sheep breeds can also influence gene expression. Among the identified DEGs in the skeleton muscle tissue between thin- and fat-tailed sheep breeds, we found that three (*TEN1*, *PDGFD* and *IRF2BP2*) and nine (*FASN*, *CD68*, *CIRI*, *EEF1A1*, *EHD1*, *GAA*, *RGMB*, *SMTNL1* and *GLUL*) DEGs showing ASE in selected samples of adipose and skeleton muscle tissues between thin- and fat-tailed sheep breeds, respectively. All of the nine ASE genes were down-regulated in longissimus dorsi samples of Texel sheep, in which the expression level of *FASN* gene was negatively correlated with intramuscular fat (IMF) content in the longissimus of Kazakh sheep [60]. *TEN1* gene showing ASE in samples collected from both adipose and skeleton muscle tissues, and all of the ASE SNPs identified in this gene were located in 5'-UTR region. Promoters are typically located in 5'-UTR region of the gene and mutations in promoters often cause disease in human [61]. Overexpression of *PDGFD* gene in sheep pre-adipocytes could significantly increase the expression levels of *PPARG* and *LPL* genes [52]. Three ASE SNPs were found in *PPARG* gene, which can control the expression of genes related to adipocyte differentiation and lipid metabolism as an important transcription factors [62]. It was

also involved in the intramuscular adipocyte differentiation in the early growing stage of cattle [62]. A total of 15 ASE SNPs were identified in *LPL* gene from eight samples of adipose tissues, and its expression in the longissimus dorsi of sucking lambs could be affected by fatty acid indicators including polyunsaturated fatty acid (PUFA), saturated fatty acid (SFA), conjugated linoleic acid (CLA) and n-6/n-3 according to the difference of forage type and lamb sex [63]. Results of KEGG pathway enrichment analysis from identified ASE genes and DEGs in adipose and skeleton muscle tissues revealed enriched categories related with AMPK signaling pathway. Activating AMPK could inhibit fatty acid synthesis while promote fatty acid oxidation and regulate brown adipogenesis [64, 65]. It also plays a crucial role in the regulation of skeletal muscle development [50]. Besides this, PPAR signaling pathway was significantly enriched for ASE genes identified in adipose tissue of F1 hybrids and DEGs that was found in tail fat samples of between thin- and fat-tailed sheep. Genetic variants of genes in the PPAR signaling pathway were associated with the traits of porcine meat quality (PMQ) [66]. It is illustrated that PPAR signaling and AMPK signaling pathways may be the key pathways regulating differential fatty acid profile and lipid metabolism between thin- and fat-tailed sheep breeds.

Candidate causative genes that contribute to phenotypic differences could also be identified by selection signals or GWAS, and genomic regions associated with tail fat deposition and meat production traits have been widely studied in thin- and fat-tailed sheep populations [7-10, 12, 14, 15, 67]. However, it is hard to discriminate causative genes as there are many other genes located in genomic region for these traits. Beside

404 this, the identified selection signals often comprise a higher proportion of false positive
405 signals. We then chose the overlapping genes between the ASE genes and those that
406 were selected in the improved thin-tailed sheep from Europe or the fat-tailed sheep from
407 Xinjiang in order to understand the molecular mechanisms underlying gene expression
408 associated with different fat deposition and meat production traits between thin- and
409 fat-tailed sheep. Seven (*BCO2*, *BINI*, *IRF2BP2*, *LRRC2*, *LOC101118164*,
410 *LOC101104530* and *SFRP1* genes) and six (*PDGFD*, *PPDPF*, *PM20D1*, *PPMIK*,
411 *PPA2* and *TUBA8* genes) ASE genes were selected in thin- and fat-tailed sheep,
412 respectively. Among these ASE genes, *LRRC2* gene was identified as the important
413 gene which were thought to be associated with meat production traits at chromosome-
414 wise significance level in sheep [49], while *PDGFD* gene was considered to be a
415 primary causal gene for fat deposition in the tails of sheep [15]. A total of nine and
416 seven ASE SNPs were found in *LRRC2* and *PDGFD* genes, respectively, majority of
417 these ASE SNPs located in 3'-UTR region of their respective genes. It is widely
418 accepted that miRNA binds to 3'-UTR region of genes and then regulate gene
419 expression. The small RNA derived from dominant allele can repress recessive allele
420 via *trans* action, and the target-site SNPs of miRNA is associated with ASE SNP of
421 miRNA target genes [68-70]. Both of these two genes exhibiting ASE may be the result
422 of miRNA binds to those ASE SNPs in 3'-UTR region. Enhancers were typically
423 located on downstream (3') and upstream (5') of the genes, promoters and enhancers
424 belong to CREs, and mutation of CREs can generate many phenotypic variants [71-73].
425 The distinct genotype pattern in the upstream region of *LRRC2* and *PDGFD* genes

maybe another important cause for the gene showing ASE. In addition, we found *BCO2* and *SFRP1* genes showing ASE in the samples from adipose tissue, while *BIN*, *PM20D1*, *PPMIK* genes showing ASE in the samples from both skeleton muscle and adipose tissues. *BCO2* gene plays an important role in the process of β -carotene metabolism [74, 75]. Seven ASE SNPs were identified in *BCO2* gene from the samples of subcutaneous fat and tail fat, six of these ASE SNPs were located in exon region of the gene and the remaining one ASE SNP was located in its 3'-UTR region. Knocking out the biallelic *BCO2* gene could lead to yellow fat in Tan sheep and a nonsense mutation in this gene is strongly related to the yellow fat phenotype in Norwegian sheep [76, 77]. A total of seven ASE SNPs were found in *PM20D1* gene, which involved in the cellular lipid process. Six of these ASE SNPs were located in 3'-UTR region of *PM20D1* gene, and the other one was located in exon region of the gene. The miR-324-5p could target the 3'-UTR region of *PM20D1* gene, overexpression of miR-324-5p could suppressed the expression of it [78]. *PPMIK* gene showing ASE in the samples of adipose and skeleton muscle tissues, and the gene located in the QTL that is related to fat and muscle density [49]. Collectively, our findings will expand our knowledge about the molecular mechanisms of gene expression related to fat deposition and meat production variations between thin- and fat-tailed sheep breeds.

Conclusions

The results from this study fill research gap about the genetic basis underlying the different fat deposition and meat production traits between thin- and fat-tailed sheep

breeds. The combination results of ASE and DEGs provide several candidate genes putatively involved in economically important traits. Overall, the identification of ASE genes will contribute to explain the genetic mechanisms of phenotypic difference between thin- and fat-tailed sheep, especially for tail fat deposition and meat production traits, and lay the research foundation for sheep breeding.

Supplementary information

Additional file 1: Table S1 Sample information used for identifying signals of selection in thin- or fat-tailed sheep breeds populations.

Additional file 2: Table S2 The data of fat deposition and meat production traits about Texel sheep and Kazakh sheep.

Additional file 3: Table S3 Samples information used for in this study.

Additional file 4: Table S4 Heterozygous loci identified in each F1 hybrids.

Additional file 5: Table S5 ASE genes identified in this study.

Additional file 6: Table S6 Differentially expressed genes in longissimus dorsi between Texel sheep and Kazakh sheep.

Additional file 7: Table S7 Selection signals identified in the fat-tailed sheep.

Additional file 8: Table S8 Selection signals identified in the thin-tailed sheep.

Abbreviations

ASE: allele-specific expression; SNP: Single nucleotide polymorphism; BAM: Binary Alignment MAP; BWA: Burrows-Wheeler Aligner program; GATK: Genome Analysis

Toolkit; GO: Gene ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes;
GWAS: Genome-wide association studies; DEGs: Differential expressed genes; RPKM:
Reads Per Kilobase per Million mapped reads.

Acknowledgements

We are grateful to the High-Performance Computing Center (HPC) of Northwest A&F
University (NWAFU) for providing bioinformatics analysis resources.

Authors' contributions

Y. J. and M. J. L. designed the experiment. S. G. H. and X. Y. P. collected sample and
conducted DNA extraction. J. J. S., L. C., Z. R. Y., X. H. and S. G. conducted
bioinformatics analysis. J. J. S., X. Y. P., H. A. N., R. L. and Y. W. wrote the manuscript.
Y. J. and M. J. L. revised the manuscript. All authors reviewed and approved the final
manuscript.

Funding

This work was supported by the research grants from the National Natural Science
Foundation of China (31822052), the National Thousand Youth Talents Plan, the Key
Laboratory Research Grant of Xinjiang Province (2017D04028), Tianshan Innovation
Team program (2018D14004), the startup fund from Northwest A&F University
(Z1090219002).

Availability of data and materials

The whole genome sequence data has been deposited in NCBI short read archive under study PRJNA 623062.

Ethics approval and consent to participate

All animal experiments in this study were carried out in accordance with the regulations of Instructive Notions with Respect to Caring for Experimental Animals, Ministry of Science and Technology of China. All protocols of animal handling and sampling in this study was approved by Experimental Animal Management Committee of the Northwest A&F University (No. 2014ZX08008002). All efforts were made to minimize animal suffering.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no conflict of interests.

Author details

¹College of Animal Science and Technology, Northwest A&F University, Yangling, Shaanxi 712100, China, ²Key Laboratory of Ruminant Genetics, Breeding & Reproduction, Ministry of Agriculture, ³Key Laboratory of Animal Biotechnology of

Xinjiang, Institute of Biotechnology, Xinjiang Academy of Animal Science, Urumqi,
Xinjiang 830026, China.

References

1. Chessa B, Pereira F, Arnaud F, Amorim A, Goyache F, Palmarini M, et al. Revealing the history of sheep domestication using retrovirus integrations. *Science*. 2009;324:532-6.
2. Ryder ML. *Sheep and Man*. London: Duckworth press; 1983. P. 17-92.
3. Xu SS, Ren X, Yang GL, Xie XL, Zhao YX, Li MH, et al. Genome-wide association analysis identifies the genetic basis of fat deposition in the tails of sheep (*Ovis aries*). *Anim Genet*. 2017;48:560-9.
4. Atti N, Bocquier F, Khaldi G. Performance of the fat-tailed Barbarine sheep in its environment: adaptive capacity to alternation of underfeeding and re-feeding periods. *Animal Research*. 2004;53:76.
5. Yousef AR, Kohram H, Shahneh AZ, Nik-Khah A, Campbell AW. Comparison of the meat quality and fatty acid composition of traditional fat-tailed (Chall) and tailed (Zel) Iranian sheep breeds. *Meat Science*. 2012;92:417-22.
6. Farahani AH, Shahrabak HM, Shahrabak MM, Yeganeh MH. Relationship of fat-tail and body measurements with some economic important traits in fat-tail Makoei breed of Iranian sheep. *Afr J Biotechnol*. 2010;9:5989-92.
7. Zhang L, Liu JS, Zhao FP, Ren HX, Xu LY, Du LX, et al. Genome-wide association studies for growth and meat production traits in sheep. *PLoS One*. 2013;8:e66569.
8. Zhang TY, Gao HD, Sahana G, Zan YJ, Fan HY, Zhao FP, et al. Genome-wide association studies revealed candidate genes for tail fat deposition and body size in the Hulun Buir sheep. *J Anim Breed Genet*. 2019;136:362-70.
9. Yuan Z, Liu E, Liu Z, Kijas JW, Zhu C, Wei C, et al. Selection signature analysis reveals genes associated with tail type in Chinese indigenous sheep. *Anim Genet*. 2016;48:55-66.
10. Moradi MH, Javaremi AN, Shahrabak MM, Dodds KG, McEwan JC. Genomic scan of selective sweeps in thin and fat tail sheep breeds for identifying of candidate regions associated with fat deposition. *BMC Genetics*. 2012;13:10.
11. Moioli B, Pilla F, Ciani E. Signatures of selection identify loci associated with fat tail in sheep. *J Anim Sci*. 2015;93:4660-9.
12. Ahbara A, Bahbahani H, Almathen F, Abri MA, Agoub MO, Mwacharo JM, et al. Genome-wide variation, candidate regions and genes associated with fat deposition and tail morphology in Ethiopian indigenous sheep. *Front Genet*. 2018;9:699.
13. Zhu CY, Fan HY, Yuan ZH, Hu SJ, Ma XM, Du LX, et al. Genome-wide detection of CNVs in Chinese indigenous sheep with different types of tails using ovine high-density 600K SNP arrays. *Sci Rep*. 2016;6:27822.
14. Wei CH, Wang HH, Liu G, Wu MM, Cao JX, Du LX, et al. Genome-wide analysis reveals population structure and selection in Chinese indigenous sheep breeds. *BMC Genomics*. 2015;16:194.

- 542 15. Li X, Yang J, Shen M, Xie XL, Liu GJ, Li MH, et al. Whole-genome resequencing of wild
543 and domestic sheep identifies genes associated with morphological and agronomic traits. *Nat*
544 *comun.* 2020;11:2815.
- 545 16. Mwacharo JM, Kim ES, Eibeltagy AR, Naga AM, Rischkowsky BA, Rothschild MF. Genomic
546 footprints of dryland stress adaptation in Egyptian fat-tail sheep and their divergence from
547 East African and western Asia cohorts. *Sci Rep.* 2017;7:17647.
- 548 17. Masrangelo S, Moioli B, Ahbara A, Latairish S, Portolano B, Ciani E, et al. Genome-wide
549 scan of fat-tail sheep identifies signals of selection for fat deposition and adaptation. *Anim*
550 *Prod Sci.* 2018. doi: 10.1071/AN17753.
- 551 18. King MW, Wilson AC. Evolution at two levels in humans and chimpanzees. *Science.* 1975;
552 188:107-16.
- 553 19. Kwaepila N, Burns G, Leong SY. Immunohistological localisation of human FAT1 (hFAT)
554 protein in 326 breast cancers. Does this adhesion molecule have a role in pathogenesis?
555 *Pathology.* 2006;38:125-31.
- 556 20. Ge B, Pokholok DK, Kwan T, Grundberg E, Morcos L, Pastinen T, et al. Global patterns of
557 *cis* variation in human cells revealed by high-density allelic expression analysis. *Nat Genet.*
558 2009;41:1216-22.
- 559 21. Heap GA, Yang JH, Downes K, Healy BC, Hunt KA, Plagnol V, et al. Genome-wide analysis
560 of allelic expression imbalance in human primary cells by high-throughput transcriptome
561 resequencing. *Hum Mol Genet.* 2010;19:122-34.
- 562 22. Tung J, Akinyi MY, Mutura S, Altmann J, Wray GA, Alberts SC. Allele-specific gene
563 expression in a wild nonhuman primate population. *Mol Ecol.* 2011;20:725-39.
- 564 23. Wittkopp PJ, Kalay G. *cis*-regulatory elements: molecular mechanisms and evolutionary
565 processes underlying divergence. *Nat Rev Genet.* 2012;13:59-69.
- 566 24. Hasin-Brumshtein Y, Hormozdiari F, Martin L, Eskin E, Drake TA, et al. Allele-specific
567 expression and eQTL analysis in mouse adipose tissue. *BMC Genomics.* 2014;15:471.
- 568 25. Maroille T, Lemonnier G, Lecardonnel J, Esquerré D, Ramayo-Caldas Y, Estellé J, et al.
569 Deciphering the genetic regulation of peripheral blood transcriptome in pigs through
570 expression genome-wide association study and allele-specific expression analysis. *BMC*
571 *Genomics.* 2017;18:967.
- 572 26. Chamberlain AJ, Jagt CJ, Hayes BJ, Khansefid M, Marett LC, Goddard ME, et al. Extensive
573 variation between tissues in allele specific expression in an outbred mammal. *BMC Genomics.*
574 2015;16:993.
- 575 27. Cao YH, Xu H, Li R, Gao S, Luo J, Jiang Y, et al. Genetic basis of phenotypic differences
576 between Chinese Yunling black goats and Nubian goats revealed by allele-specific expression
577 in their F1 hybrids. *Front Genet.* 2019;10:145.
- 578 28. Salavati M, Bush SJ, Vera SP, Hume DA, Clark EL, et al. Elimination of reference mapping
579 bias reveals robust immune related allele-specific expression in crossbred sheep. *Front Genet.*
580 2019;10:863.
- 581 29. Wang Y, Gao S, Zhao Y, Chen WH, Shao JJ, Jiang Y, et al.: Allele-specific expression and
582 alternative splicing in horse×donkey and cattle×yak hybrids. *Zool Res.* 2019; 40:293-304.
- 583 30. Miller SA, Dykes, DD, Polesky HF. A simple salting out procedure for extracting DNA from
584 human nucleated cells. *Nucleic Acids Res.* 1988;16:1215.
- 585 31. Aceto GM, Lellis LD, Catalano T, Veschi S, Radice P, Curia C, et al., Mariani-Costantini R,

586 Cama A, Curia MC: Nonfluorescent denaturing HPLC-based primer-extension method for
587 allele-specific expression: application to analysis of mismatch repair genes. *Clin Chem.*
588 2009;55:1711-8

589 32. Jiang Y, Xie M, Chen WB, Xu X, Wang W, Dalrymple BP, et al. The Sheep Genome
590 Illuminates Biology of the Rumen and Lipid Metabolism. *Science.* 2014;344:1168-73.

591 33. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, DePristo M, et al. The Genome
592 Analysis Toolkit : A MapReduce framework for analyzing next-generation DNA sequencing
593 data. *Genome Res.* 2010;20:1297-1303.

594 34. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from
595 high-throughput sequencing data. *Nucleic Acids Res.* 2010;38:e164.

596 35. Pan XY, Wang Y, Li ZJ, Chen XQ, Heller R, Jiang Y, et al. Tracing the origin of a new organ
597 by inferring the genetic basis of rumen evolution. *bioRxiv.* 2020. 10.1101/2020.02.19.955872

598 36. Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL. Transcript-level expression analysis of
599 RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat Protoc.* 2016;11:1650-67.

600 37. Geijn B, McVicker G, Gilad Y, Pritchard JK. WASP: allele-specific software for robust
601 molecular quantitative trait locus discovery. *Nat Methods.* 2015;12:1061-3.

602 38. Degner JF, Marioni JC, Pai AA, Pickrell JK, Nkadori E, Pritchard JK, et al. Effect of read-
603 mapping biases on detecting allele-specific expression from RNA-sequencing data.
604 *Bioinformatics.* 2009;25:3207-12.

605 39. Stevenson KR, Coolon JD, Wittkopp PJ. Sources of bias in measures of allele-specific
606 expression derived from RNA-seq data aligned to a single reference genome. *BMC Genomics.*
607 2013;14:536.

608 40. Mayba O, Gilbert HN, Liu J, Haverty PM, Jiang ZS, Zhang Z. MBASED: allele-specific
609 expression detection in cancer tissues and cell lines. *Genome Biol.* 2014;15:405.

610 41. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data.
611 *Bioinformatics.* 2014;30:2114-20.

612 42. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-
613 seq data with DESeq2. *Genome Biol.* 2014;15:550.

614 43. Pan ZY, Li SD, Liu QY, Wang Z. Chu MX, Li YX, et al. Whole-genome sequences of 89
615 Chinese sheep suggest role of *RFXR2* in the development of unique horn phenotype as
616 response to semi-feralization. *GigaScience.* 2018;7:giy019.

617 44. Sanchez MN, Nguyen Q, McWilliam S, Tellam R, Reverter A, Kijas J, et al. Sheep genome
618 functional annotation reveals proximal regulatory elements contribute to the evolution of
619 modern breeds. *Nat Commun.* 2018;9:859.

620 45. Zhao YX, Yang J, Lv FH, Hu JX, Xie XL, Li MH, et al. Genomic reconstruction of the history
621 of native sheep reveals the peopling patterns of nomads and the expansion of early pastoralism
622 in East Asia. *Mol Biol Evol.* 2017;34:2380-95.

623 46. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, Durbin R, et al. The variant call format
624 and VCFtools. *Bioinformatics.* 2011; 27:2156-58.

625 47. Kleinjan DA, Heyningen VV. Long-range control of gene expression: emerging mechanisms
626 and disruption in disease. *Am J Hum Genet.* 2005;76:8-32

627 48. Guan DL, Luo NJ, Tan XS, Zhao ZQ, Huang YF, Zhao YJ, et al. Scanning of selection
628 signature provides a glimpse into important economic traits in goats (*Capra hircus*). *Sci Rep.*
629 2016;6:36372

630 49. Zlobin AS, Volkova NA, Borodin PM, Aksenovich TI, Tsepilov YA. Recent advances in
631 understanding genetic variants associated with growth, carcass and meat productivity traits in
632 sheep (*Ovis aries*): an update. Arch Anim Breed. 2019;62:579-83.

633 50. Thomson DM. The role of AMPK in the regulation of skeletal muscle size, hypertrophy, and
634 regeneration. Int J Mol Sci. 2018;19:3125.

635 51. Cui HX, Zhang MQ, Liu QQ, Zhao GP, Chen JL, Wen J. Liver dominant expression of fatty
636 acid synthase (FAS) gene in two chicken breeds during intramuscular-fat development. Mol
637 Biol Rep. 2012;39:3479-84.

638 52. Li Q, Lu ZK, Jin ML, Fei XJ, Quan K, Wei CH, et al. Verification and analysis of sheep tail
639 type-associated PDGFD gene polymorphisms. Animals. 2020;10:89.

640 53. Chen L, Qiu Q, Jiang Y, Wang K, Lin ZS, Wang W, et al. Large-scale ruminant genome
641 sequencing provides insights into their evolution and distinct traits. Science. 2019;
642 364:eaav6202

643 54. Pinter SF, Colognori D, Beliveau BJ, Sadreyev RI, Payer B, Lee JT, et al. Allelic imbalance
644 is a prevalent and tissue-specific feature of the mouse transcriptome. genetics. 2015;200:537-
645 49.

646 55. Shao L, Xing F, Xu CH, Zhang QH, Che J, Zhang QF, et al. Patterns of genome-wide allele-
647 specific expression in hybrid rice and the implications on the genetic basis of heterosis. Proc
648 Natl Acad Sci. 2019;116:5653-8.

649 56. Bakhtiarzadeh MR, Shahrabak MM, Ebrahimie E. Underlying functional genomics of fat
650 deposition in adipose tissue. Gene. 2013;521:122-8.

651 57. Kang DJ, Zhou GX, Zhou SW, Zeng J, Wang XL, Chen YL, et al. Comparative transcriptome
652 analysis reveals potentially novel roles of Homeobox genes in adipose deposition in fat-tailed
653 sheep. Sci Rep. 2017;7:14491.

654 58. Bakhtiarzadeh MR, Salehi A, Alamouti AA, Arpanahi RA, Salami SA. Deep transcriptome
655 analysis using RNA-seq suggests novel insights into molecular aspects of fat-tail metabolism
656 in sheep. Sci Rep. 2019;9:9203.

657 59. McManus JC, Coolon JD, Duff MO, Main JE, Graveley BR, Wittkopp PJ. Regulatory
658 divergence in *Drosophila* revealed by mRNA-seq. Genome Res. 2010;20:816-25.

659 60. Qiao Y, Huang ZG, Li QF, Liu ZS, Hao CL, Xie Z, et al. Developmental changes of the FAS
660 and HSL mRNA expression and their effects on the content of intramuscular fat in Kazak and
661 Xinjiang Sheep. J Genet Genomics. 2007;34:909-17.

662 61. Savinkova LK, Ponomarenko MP, Ponomarenko PM, Drachkova IA, Lysova MV, Kolchanov
663 NA, et al. TATA box polymorphisms in human gene promoters and associated hereditary
664 pathologies. Biochemistry. 2009;74:117-29.

665 62. Graugnard DE, Piantoni P, Bionaz M, Berger LL, Faulkner DB, Looor JJ. Adipogenic and
666 energy metabolism gene networks in longissimus lumborum during rapid post-weaning
667 growth in Angus and Angus × Simmental cattle fed high-starch or low-starch diets. BMC
668 Genomics. 2009;10:142.

669 63. Dervishi E, Joy M, Serrano M, Calvo JH, et al. The forage type (grazing versus hay pasture)
670 fed to ewes and the lamb sex affect fatty acid profile and lipogenic gene expression in the
671 longissimus muscle of suckling lambs. J Animal Sci. 2012;90:54-66.

672 64. O'Neill HM, Holloway GP, Steinberg GR. AMPK regulation of fatty acid metabolism and
673 mitochondrial biogenesis: implications for obesity. Mol Cell Endocrinol. 2013;366:135-51.

65. Day EA, Ford RJ, Steinberg GR. AMPK as a therapeutic target for treating metabolic diseases. *Trends Endocrin Met.* 2017;28:545-60.
66. He K, Wang Q, Wang Z, Pan YC. Association study between gene polymorphisms in PPAR signaling pathway and porcine meat quality traits. *Mamm Genome.* 2013;24:322-31.
67. Moradi MH, Javaremi AN, Shahrababak MM, Dodds KG, McEwan JC, et al. Genomic scan of selective sweeps in thin and fat tail sheep breeds for identifying of candidate regions associated with fat deposition. *BMC Genetics.* 2012;13:10.
68. Kim J, Bartel DP. Allelic imbalance sequencing reveals that single-nucleotide polymorphisms frequently alter microRNA directed repression. *Nat Biotechnol.* 2009;27:472-7.
69. Tarutani Y, Shiba H, Iwano M, Watanabe M, Isogai A, Takayama S, et al. Trans-acting small RNA determines dominance relationships in Brassica self-incompatibility. *Nature.* 2010;466:983-6.
70. Zhang WS, Edwards A, Zhu DX, Flemington EK, Deininger P, Zhang K. miRNA-mediated relationships between cis-SNP genotypes and transcript intensities in lymphocyte cell lines. *PLoS One.* 2012;7:e31429.
71. Wray GA. The evolutionary significance of cis-regulatory mutations. *Nat rev genet.* 2007;8:206-16.
72. Carroll SB. Evo-Devo and an expanding evolutionary synthesis: A genetic theory of morphological evolution. *Cell.* 2008;134:26-36.
73. Orgogozo V, Stern DL. The loci of evolution: How predictable is genetic evolution? *Evolution.* 2008;62:2155-77.
74. Lintig JV. Colors with functions: elucidating the biochemical and molecular basis of carotenoid metabolism. *Annu Rev Nutr.* 2010;30:35-56.
75. Lintig JV. Provitamin A metabolism and functions in mammalian biology. *Am J Clin Nutr.* 2012;96:1234s-44s.
76. Vage DI, Boman IA. A nonsense mutation in the betacarotene oxygenase 2 (*BCO2*) gene is tightly associated with accumulation of carotenoids in adipose tissue in sheep (*Ovis aries*). *BMC Genetics.* 2010;11:10..
77. Niu Y, Jin M, Li Y, Li P, Zhou J, Chen Y, et al. Biallelic -carotene oxygenase 2 knockout results in yellow fat in sheep via CRISPR/Cas9. *Anim Genet.* 2017;242-4..
78. Li DD, Liu Y, Gao W, Han JK, Zhang MD, Pang WY, et al. Inhibition of miR-324-5p increases PM20D1-mediated white and brown adipose loss and reduces body weight in juvenile mice. *Eur J Pharmacol.* 2019;86.

Figure legend and Tables

Fig. 1 Phenotypic difference between Texel and Kazakh sheep. a-b, Photos from Texel and Kazakh sheep, respectively. c-k, Comparison of fat deposition and meat production traits between Texel (n = 5) and Kazakh sheep (n = 29). Groups with

significant differences were analyzed by two-tailed unpaired *t*-test, * $p < 0.05$, ** $p < 0.01$.

Fig. 2 Heterozygous sites identified from F1 hybrids of Texel × Kazakh sheep. a, Annotation of the whole genome SNPs; b, Annotation of ASE SNPs identified from the adipose and skeleton muscle tissues; c, Percentage of different substitution types in the ASE SNP.

Fig. 3 ASE genes identified in adipose tissue. a, Venn diagram showing the number of ASE genes identified in between subcutaneous fat and tai fat samples; b, Distribution of ASE SNPs in per ASE gene identified in adipose tissue; c, Results of KEGG pathway enrichment for ASE genes identified in adipose tissue of F1 hybrids and DEGs identified in skeleton muscle tissue between thin- and Kazakh sheep breeds d, ASE genes were annotated in PPAR signaling pathway.

Fig. 4 ASE genes identified in skeleton muscle tissue. a, Venn diagram showing the number of ASE genes identified in samples from longissimus dorsi, semitendinosus and semimembranosus; b, Distribution of ASE SNPs in ASE genes identified in adipose tissue; c, Comparison of ASE genes identified between skeleton muscle and adipose tissues; d, Results of KEGG pathway enrichment for ASE genes identified in skeleton muscle tissue of F1 hybrids and DEGs identified in skeleton muscle tissue between Texel and Kazakh sheep breeds; e, ASE genes were differential expressed in skeleton

muscle tissue between thin- and fat-tailed sheep, * $p < 0.05$, ** $p < 0.01$. Data are showed as mean $\pm$ standard error.

Fig.5 Genomic regions with selection sweep signals identified between Mouflon sheep and fat-tailed sheep populations. a, ASE genes were significantly selected by F_{ST} & $\ln(\theta\pi_{\text{Mouflon}} / \theta\pi_{\text{fat-tailed sheep}})$; b, Genotype patterns for the upstream region of *PDGFD* gene between Mouflon sheep (n=17) and fat-tailed sheep (n=20). Detailed information of ASE SNPs identified in *PDGFD* gene (Samples 1-4 from subcutaneous fat, Samples 5-8 from tail fat).

Fig.6 Genomic regions with selection sweep signals identified between Mouflon sheep and thin-tailed sheep populations. a, ASE genes were significantly selected by F_{ST} & $\ln(\theta\pi_{\text{Mouflon}} / \theta\pi_{\text{thin-tailed sheep}})$; b, Detailed information of ASE SNPs identified in *LRCC2* gene (Sample 1 from longissimus dorsi, Samples 2-3 from semimembranosus); c, Genotype patterns of the *LRRC2* gene region between Mouflon sheep (n=17) and thin-tailed sheep (n=15).

Figures

Figure 1

Phenotypic difference between Texel and Kazakh sheep. a-b, Photos from Texel and Kazakh sheep, respectively. c-k, Comparison of fat deposition and meat production traits between Texel (n = 5) and

Kazakh sheep (n = 29). Groups with significant differences were analyzed by two-tailed unpaired t-test, * p < 0.05, ** p < 0.01.

Figure 2

Heterozygous sites identified from F1 hybrids of Texel × Kazakh sheep. a, Annotation of the whole genome SNPs; b, Annotation of ASE SNPs identified from the adipose and skeleton muscle tissues; c, Percentage of different substitution types in the ASE SNP.

Figure 3

ASE genes identified in adipose tissue. a, Venn diagram showing the number of ASE genes identified in between subcutaneous fat and tai fat samples; b, Distribution of ASE SNPs in per ASE gene identified in adipose tissue; c, Results of KEGG pathway enrichment for ASE genes identified in adipose tissue of F1 hybrids and DEGs identified in skeleton muscle tissue between thin- and Kazakh sheep breeds d, ASE genes were annotated in PPAR signaling pathway.

Figure 4

ASE genes identified in skeleton muscle tissue. a, Venn diagram showing the number of ASE genes identified in samples from longissimus dorsi, semitendinosus and semimembranosus; b, Distribution of ASE SNPs in ASE genes identified in adipose tissue; c, Comparison of ASE genes identified between skeleton muscle and adipose tissues; d, Results of KEGG pathway enrichment for ASE genes identified in skeleton muscle tissue of F1 hybrids and DEGs identified in skeleton muscle tissue between Texel and

Kazakh sheep breeds; e, ASE genes were differential expressed in skeleton muscle tissue between thin- and fat-tailed sheep, * $p < 0.05$, ** $p < 0.01$. Data are showed as mean $\pm$ standard error.

Figure 5

Genomic regions with selection sweep signals identified between Mouflon sheep and fat-tailed sheep populations. a, ASE genes were significantly selected by F_{ST} & $\ln(\theta\pi_{\text{Mouflon}} / \theta\pi_{\text{fat-tailed sheep}})$; b, Genotype patterns for the upstream region of PDGFD gene between Mouflon sheep ($n=17$) and fat-tailed sheep ($n=20$). Detailed information of ASE SNPs identified in PDGFD gene (Samples 1-4 from subcutaneous fat, Samples 5-8 from tail fat).

Figure 6

Genomic regions with selection sweep signals identified between Mouflon sheep and thin-tailed sheep populations. a, ASE genes were significantly selected by F_{ST} & $\ln(\theta\pi\theta\pi\text{-Mouflon} / \theta\pi\theta\pi\text{-thin-tailed sheep})$; b, Detailed information of ASE SNPs identified in LRRC2 gene (Sample 1 from longissimus dorsi, Samples 2-3 from semimembranosus); c, Genotype patterns of the LRRC2 gene region between Mouflon sheep (n=17) and thin-tailed sheep (n=15).

Supplementary Files

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

- [TableS10805.xlsx](#)
- [TableS20805.xlsx](#)
- [TableS30805.xlsx](#)
- [TableS40805.xlsx](#)
- [TableS50805.xlsx](#)
- [TableS60805.xlsx](#)
- [TableS70805.xlsx](#)
- [TableS80805.xlsx](#)