

Article

# Genome-Wide Association Study of Body Weight Traits in Chinese Fine-Wool Sheep

Zengkui Lu <sup>1,2,†</sup> , Yaojing Yue <sup>1,2,†</sup>, Chao Yuan <sup>1,2</sup>, Jianbin Liu <sup>1,2</sup>, Zhiqiang Chen <sup>3</sup>, Chune Niu <sup>1,2</sup>, Xiaoping Sun <sup>1,2</sup>, Shaohua Zhu <sup>1,2</sup>, Hongchang Zhao <sup>1,2</sup>, Tingting Guo <sup>1,2,\*</sup> and Bohui Yang <sup>1,2,\*</sup>

<sup>1</sup> Lanzhou Institute of Husbandry and Pharmaceutical Sciences, Chinese Academy of Agricultural Sciences, Lanzhou 730050, China; luzengkui@caas.cn (Z.L.); yueyaojing@126.com (Y.Y.); yuanchao@caas.cn (C.Y.); liujianbin@caas.cn (J.L.); niuchune@caas.cn (C.N.); sunxiaoping@caas.cn (X.S.); zhu87932890@outlook.com (S.Z.); 18837101296@163.com (H.Z.)

<sup>2</sup> Sheep Breeding Engineering Technology Research Center of Chinese Academy of Agricultural Sciences, Lanzhou 730050, China

<sup>3</sup> Novogene Bioinformatics Institute, Beijing 100029, China; chenzhiqiang@novogene.com

\* Correspondence: guotingting@caas.cn (T.G.); yangbohui@caas.cn (B.Y.)

† These authors contributed equally to this work.

Received: 4 December 2019; Accepted: 14 January 2020; Published: 19 January 2020



**Simple Summary:** Body weight traits are economically important in the sheep industry, and it is critical to explore their underlying genetic architecture. Hence, four body weight traits, including birth, weaning, yearling, and adult weights were examined. Through a genome-wide association study on Chinese fine-wool sheep, several candidate single-nucleotide polymorphisms (SNPs) and genes were found potentially associated with the traits of interest. The results of this study may facilitate the potential use of the genes involved in growth and production traits for the genetic improvement of productivity in sheep.

**Abstract:** Body weight is an important economic trait for sheep and it is vital for their successful production and breeding. Therefore, identifying the genomic regions and biological pathways that contribute to understanding variability in body weight traits is significant for selection purposes. In this study, the genome-wide associations of birth, weaning, yearling, and adult weights of 460 fine-wool sheep were determined using resequencing technology. The results showed that 113 single nucleotide polymorphisms (SNPs) reached the genome-wide significance levels for the four body weight traits and 30 genes were annotated effectively, including *AADACL3*, *VGF*, *NPC1*, and *SERPINA12*. The genes annotated by these SNPs significantly enriched 78 gene ontology terms and 25 signaling pathways, and were found to mainly participate in skeletal muscle development and lipid metabolism. These genes can be used as candidate genes for body weight in sheep, and provide useful information for the production and genomic selection of Chinese fine-wool sheep.

**Keywords:** body weight; genome-wide association study; fine-wool sheep; genomic analysis; functional analysis

## 1. Introduction

Genome-wide association studies (GWAS) are used to screen the whole genome for target genes that correlate with phenotypic traits, using single nucleotide polymorphisms (SNPs) as genetic markers. They have become an important method for identifying candidate genes for important economic traits in livestock. Compared with traditional quantitative trait locus (QTL) mapping, GWAS have a greater

capability to detect causal mutations, and can detect mutations in a smaller genetic range [1]. In recent years, GWAS have been included in genetic breeding programs for animals like pigs, cattle, sheep, and chickens, and have identified many genes or molecular markers that could regulate important economic traits in livestock [2–5].

Body weight is the most important growth and development index in sheep production, and it influences meat and wool production, and reproduction of sheep, both directly and indirectly [6]. During production, the birth, weaning, yearling, and adult weights are used to reflect the growth and development of sheep. The growth in body weight is influenced by a combination of the genetic background of the sheep and its feeding and management. Growth in weight is a highly heritable trait and is one of the main indexes of selection [6]. Ghasemi et al. carried out a GWAS on the birth weight of Lori-Bakhtiari sheep and identified three genes (*RAB6B*, Tf serotransferrin and *GIGYF2*) on chromosome one as candidate genes for this trait [4]. In addition, Al-Mamun et al. performed a GWAS on the weaning weight of Australian Merino sheep and identified three genes (*LAP3*, *NCAPG*, and *LCORL*) on chromosome six for use during marker-assisted selection (MAS) for body weight [7]. Other genome scanning and QTL mapping studies of body weight in sheep also found many candidate genes that significantly correlated with body weight traits [8–11]. However, while studies mapping candidate genes that are related to body weight traits in sheep have been performed, the major genes are still unknown. Additionally, many previous studies only examined the candidate genes that significantly correlated with body weight during a single developmental stage, rather than performing a systematic study of body weight traits throughout the different developmental stages.

In this study, the genome-wide associations of the birth, weaning, yearling and adult weights of 460 fine-wool sheep in China were determined using genome resequencing technology. SNPs and candidate genes, related to body weight, were identified, thereby, providing references for the MAS of body weight in fine-wool sheep.

## 2. Materials and Methods

### 2.1. Animals and Sample Collection

All experimental protocols and procedures were approved by the Institutional Animal Care and Use Committee of Lanzhou Institute of Husbandry and Pharmaceutical Science of Chinese Academy of Agricultural Sciences (Approval No. NKMYD201805; Approval Date: 18 October 2018). A total of 460 fine-wool sheep in China were selected for this study, including 220 Alpine Merino sheep (AMS; 75 males and 145 females; Huangcheng, Gansu, China), 120 Chinese Merino sheep (CMS; 60 males and 60 females; Gongnaisi, Xinjiang, China), 60 Aohan fine-wool sheep (AHS; 30 males and 30 females; Chifeng, Inner Mongolia, China), and 60 Qinghai fine-wool sheep (QHS; 30 males and 30 females; Sanjiaocheng, Gangcha, China). All the sheep were selected randomly, without pedigree information. The sheep were farmed on pasture, with appropriate feed being supplemented in winter. Blood samples (5 ml) were collected from the neck veins of all the sheep and placed in EDTA anti-freezing tubes for storage at  $-20\text{ }^{\circ}\text{C}$ .

### 2.2. Phenotypic Measurements

The birth, weaning (3.5 months), yearling (12 months), and adult (30 months) weights of all sheep were measured using an electronic scale, and gender was also recorded. Birth weights were recorded within 0.5 h after birth, and the weaning, yearling, and adult weights were measured after 12 h of fasting. Phenotypic values that aligned absolutely or approximately to a normal distribution were not processed, while extremely deviated values underwent Box-Cox transformation using the *powerTransform* and *bcPower* functions in the *car* extension package in R, in order to obtain an approximately normal distribution [12,13].

### 2.3. DNA Resequencing and Data Preprocessing

Genomic DNA was extracted from each blood sample using the standard phenol-chloroform method [14]. The integrity and purity of the DNA were tested using 1% agarose gel electrophoresis and a Nanodrop 2000 ultraviolet spectrophotometer (Thermo, Waltham, MA, USA). The DNA concentration was measured using a Qubit 2.0 (Invitrogen, Carlsbad, CA, USA). Aliquots (1.5 µL) of DNA were taken from each sample and library construction was performed according to the Truseq Nano DNA HT (Illumina, San Diego, CA, USA) instructions. Based on this library, the whole genomes were re-sequenced with paired-end 150 bp reads using the Illumina HiSeq Xten platform (Illumina), producing the raw data. Index sequences, barcoding sequences, and low-quality sequences were removed by Trimmomatic software (v0.32), with the following parameters being used: MINLEN = 50, LEADING = 20, TRAILING = 20, SLIDINGWINDOW = 5, 20 [15]. The clean reads were compared with the sheep reference genome (Oar\_v4.0, GCF\_000298735.2) using BWA software (v0.7.11), with the following alignment parameters being set: mem -t 4 -k 32 -M [16]. Duplicates were removed using the rmdup command in SAMTools [17]. If multiple read pairs had identical external coordinates, only the pair with the highest mapping quality was retained.

### 2.4. SNP Identification and Annotation

High-quality SNPs were obtained through SNP detection using the mpileup command in SAMTools, with the filtering conditions set as follows: Coverage depth > 3, proportion of mis-assignments < 10%, and minor allele frequency > 5% [17]. The high-quality SNPs obtained using these steps were annotated using ANNOVAR software [18]. The annotation analysis mainly included the genes within which the SNP loci were located, and the variations in the types and positions of the SNP loci.

### 2.5. Genome-wide Association Studies

The correlations between the SNPs and the traits were tested using mixed linear models in GEMMA software [19,20]. In GWAS, individual kinship and population stratification are the main causes of false positive correlations. Therefore, population genetic structure and sex were used as a fixed effect, and individual kinship was used as a random effect to correct for the influences of population structure and individual kinship. The statistical analysis model used in this study was  $y = X\alpha + Z\beta + W\mu + e$ , where  $y$  is the phenotypic trait,  $X$  is a matrix of fixed effects,  $\alpha$  is the estimation parameter of the fixed effects,  $Z$  is a matrix of SNPs,  $\beta$  is the effect of the SNPs,  $W$  is a matrix of random effects,  $\mu$  is the predicted random individuals, and  $e$  is the random error, with the distribution  $e \sim (0, \delta_e^2)$ . The significance threshold for the GWAS was defined using the Bonferroni correction method. The total type I error rate was controlled at 5% and the significance threshold of the genome was  $0.05/N_{\text{snp}}$ , where  $N_{\text{snp}}$  is the number of SNPs remaining after quality control [21].

### 2.6. Bioinformatics Analysis

Genes associated with significantly correlated SNP loci were annotated with the sheep reference genome (Oar\_v4.0, GCF\_000298735.2) and submitted to the DAVID database (<http://david.abcc.ncifcrf.gov/>) for gene ontology (GO) and Kyoto encyclopedia of genes and genomes (KEGG) analysis [22,23]. Significant enrichment in the candidate genes was indicated by a  $p$ -value of  $\leq 0.05$ .

### 2.7. Statistical Analysis

All body weight data reported are expressed as mean  $\pm$  SD. Student's  $t$ -test was carried out using SPSS software for statistical analysis of the data. A  $p$  value of  $< 0.05$  was considered to be statistically significant.

### 3. Results

#### 3.1. Phenotypic Data Analysis of Body Weight Traits

The weight traits of the 460 fine-wool sheep of four breeds measured in this study included the birth, weaning, yearling, and adult weights. The descriptive statistics calculated for the body weight traits, included the mean, median, maximum, and minimum values, as well as the standard deviation. It can be seen from Supplementary Figure S1 that the rams were heavier than the ewes, and that the AHS and AMS were heavier than the CMS and QHS, during all phases of growth. Additionally, a frequency distribution chart of the phenotypic values for these traits was drawn, and an approximately normal distribution was attained after the adjustment of any values that did not align to the normal distribution.

#### 3.2. Summary of Sequencing Data

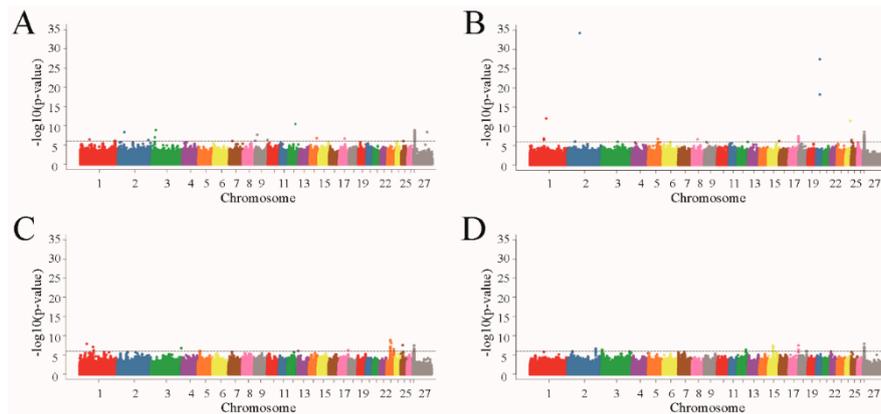
After sequencing was performed using the Illumina HiSeq Xten platform, high-quality next-generation sequencing data for the 460 fine-wool sheep were obtained (Supplementary Table S1). A total of 8,222,158,631,700 raw bases were generated in this study, with 8,189,580,053,700 clean bases being retained after quality control, at an average retention rate of 99.60%. The GC contents of the 460 samples ranged from 41.58% to 47.31%, thus, conforming to the base composition laws. Q20  $\geq$  93.75% and Q30  $\geq$  85.83%. Over 97.44% of the clean reads could be aligned to the sheep reference genome using BWA software. The coverage depth of each sample after genome alignment was about 8.59 times. Following quality control, 12,533,461 high-quality SNPs were obtained.

#### 3.3. Genome-Wide Association Study

Before performing the GWAS analysis, the population structure of the test population had to be analyzed and corrected accordingly [24]. As can be seen from Supplementary Figure S2, with the exception of QHS, the other three breeds are well-separated. Therefore, the effect of population stratification on phenotypic variation needs to be corrected when performing association analysis. The linkage disequilibrium (LD) analysis was performed on the four breeds and the LD decay is shown in Supplementary Figure S2. The results indicate that LD decay tends to be stable when the distance is 100 kb. Therefore, genes located within 100 kb of the significant SNPs are defined as candidate genes.

GEMMA software was used to perform whole-genome correlation analysis based on a mixed linear model for the birth, weaning, yearling, and adult weights. The GWAS results showed that a total of 113 SNPs in the genome correlated significantly with the four body weight traits (Figure 1).

For birth weight, 29 significantly correlated SNPs were detected on the chromosomes Chr1, Chr2, Chr3, Chr7, Chr9, Chr12, Chr14, Chr17, Chr25, and Chr27, and were effectively annotated to eight genes (Figure 1A; Table 1). For weaning weight, 38 significantly correlated SNPs were detected on Chr1, Chr2, Chr3, Chr5, Chr8, Chr16, Chr17, Chr20, Chr24, Chr25, and Chr27, and effectively annotated to 10 genes (Figure 1B; Table 1). For yearling weight, 31 significantly correlated SNPs were detected on Chr1, Chr3, Chr5, Chr13, Chr17, Chr23, Chr25, and Chr27, and effectively annotated to 12 genes (Figure 1C; Table 1). For adult weight, 15 significantly correlated SNPs were detected on Chr2, Chr3, Chr12, Chr15, Chr17, Chr18, and Chr27, and effectively annotated to six genes (Figure 1D; Table 1). As can be seen on the Q-Q chart, the observed and expected  $p$ -values for the four traits were clearly separated by  $10^{-4}$ , indicating that the correlation results were good (Figure 2).

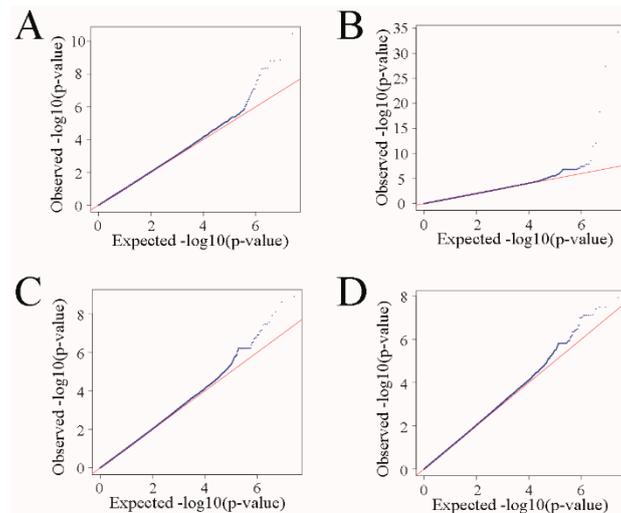


**Figure 1.** Manhattan plots of body weight traits for four fine-wool sheep breeds. The grey horizontal lines in the Manhattan plots indicate the suggestive significance ( $10^{-6}$ ) thresholds. (A) birth weight; (B) weaning weight; (C) yearling weight; (D) adult weight.

**Table 1.** Genome-wide association studies (GWAS)-identified significant single-nucleotide polymorphisms (SNPs), associated traits, and nearest candidate genes.

Traits	Chr	Position (bp)	MAF	$-\log_{10}(p\text{-Value})$	Distance (bp)	Gene	Variant Effect
birth weight	ch1	253212365	0.10	6.111011551	−19101	<i>SLCO2A1</i>	intergenic
	ch9	14382214	0.49	7.664224268	−2084	<i>LY6K</i>	intergenic
	ch9	90030866	0.15	6.323004957	−26215	<i>RALYL</i>	intergenic
	ch12	52839429	0.50	10.45675318	0	<i>AADACL3</i>	intronic
	ch17	39807635	0.06	6.65587487	0	<i>C17H4orf45</i>	intronic
	ch25	13791395	0.48	6.006388368	−17811	<i>BICC1</i>	intergenic
	ch27	6893599	0.16	8.797304666	10577	<i>GPR143</i>	intergenic
	ch27	7049963	0.15	8.355957873	0	<i>SHROOM2</i>	intronic
weaning weight	ch1	101447761	0.21	6.814810126	7370	<i>C1H1orf68</i>	intergenic
	ch1	118783826	0.50	12.06478576	−16239	<i>CLIC6</i>	intergenic
	ch3	118519704	0.07	6.007008521	0	<i>TMTC2</i>	intronic
	ch20	26410829	0.50	27.43388012	−16968	<i>STK19</i>	intergenic
	ch20	26410829	0.50	27.43388012	−25619	<i>DXO</i>	intergenic
	ch20	26410829	0.50	27.43388012	−27940	<i>SKIV2L</i>	intergenic
	ch24	35415932	0.49	11.41651058	−19192	<i>NAT16</i>	intergenic
	ch24	35415932	0.49	11.41651058	−28975	<i>VEGF</i>	intergenic
	ch27	6892015	0.21	8.574556919	8993	<i>GPR143</i>	intergenic
	ch27	6892015	0.21	8.574556919	−87231	<i>SHROOM2</i>	intergenic
yearling weight	ch1	51280492	0.50	7.90882916	−29110	<i>RABG7B</i>	intergenic
	ch1	96929132	0.50	7.121270413	29522	<i>TRNAQ-CUG</i>	intronic
	ch1	96929132	0.50	7.121270413	27565	<i>TRNAN-GUU</i>	intronic
	ch1	101447761	0.21	6.206583675	7370	<i>C1H1orf68</i>	intergenic
	ch5	13636592	0.11	6.080914687	17967	<i>ZNF557</i>	intergenic
	ch13	1884594	0.35	6.049430267	0	<i>PLCB4</i>	intronic
	ch23	33324859	0.50	8.093591468	0	<i>NPC1</i>	intronic
	ch23	33324859	0.50	8.093591468	−16422	<i>C23H18orf8</i>	intronic
	ch23	52260021	0.12	6.009205597	0	<i>DCC</i>	intronic
	ch25	12802212	0.50	7.554054255	−22550	<i>ZNF25</i>	intergenic
	ch27	7053187	0.19	7.464649666	170165	<i>GPR143</i>	intronic
	ch27	7053187	0.19	7.464649666	0	<i>SHROOM2</i>	intronic
adult weight	ch2	206466290	0.06	6.642091631	0	<i>PARD3B</i>	intronic
	ch3	4796131	0.09	6.34382137	0	<i>MED27</i>	intronic
	ch18	57956306	0.11	6.049410169	24508	<i>SERPINA14</i>	intergenic
	ch18	57956306	0.11	6.049410169	−4231	<i>SERPINA12</i>	intergenic
	ch27	6743376	0.05	7.929222061	0	<i>TBLIX</i>	intronic
	ch27	6997506	0.19	6.124895773	0	<i>SHROOM2</i>	intronic

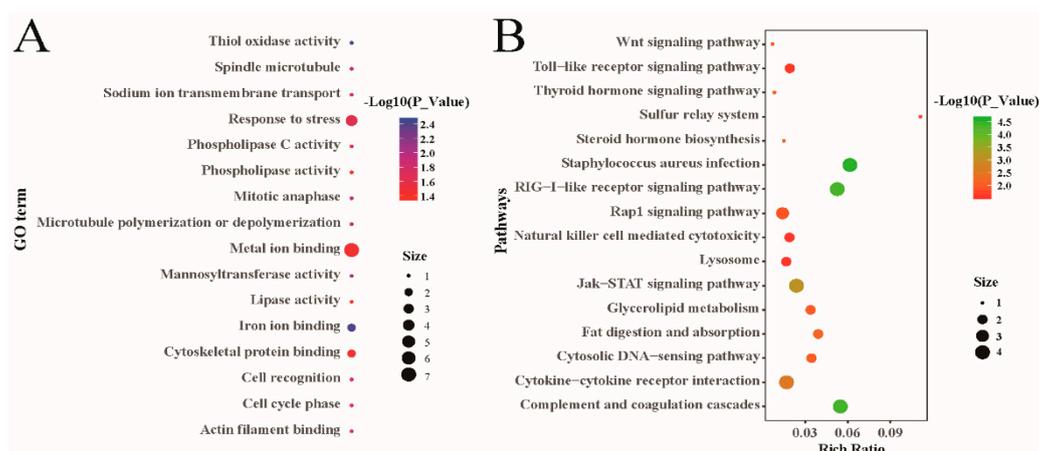
Chr: chromosome; MAF: minor allele frequency; Variant effect: SNP effect on the gene.



**Figure 2.** Quantile-quantile plots of body weight traits using mixed linear model approach. Blue dots represent the  $-\log_{10}(p\text{-value})$  of the entire study and the red line represents the expected values under the null hypothesis of no association. (A) birth weight; (B) weaning weight; (C) yearling weight; (D) adult weight.

### 3.4. Bioinformatic Analysis

To further understand the functions of the significantly-correlated SNPs, functional enrichment analysis of the genes to which the SNPs were annotated was carried out. A total of 78 GO terms were significantly enriched ( $p < 0.05$ ), including 25 biological processes, 13 cellular components, and 40 molecular functions (Supplementary Table S2). These GO terms were mainly related to microtubular movement, the cytoskeleton and the stress response, among others (Figure 3A). The KEGG pathway analysis revealed that the genes associated with the significantly correlated SNPs were highly enriched in 25 pathways ( $p < 0.05$ , Supplementary Table S3). Further analysis indicated that these pathways, which included the toll-like receptor signaling pathway, the wnt signaling pathway, and fat digestion, were associated with immunological functions, the development of skeletal muscle, and lipid metabolism (Figure 3B).



**Figure 3.** Gene ontology (GO) (A) and Kyoto encyclopedia of genes and genomes (KEGG) (B) analysis for the regional candidate genes with genome-wide significant association.

## 4. Discussion

Body weight is an important economic trait in sheep production. Studies have shown that the heritability of sheep body weight is moderately high, with the heritability of birth weight and weaning

weight ranging from 0.30 to 0.35, the heritability of yearling weight being 0.40–0.45, and that of adult weight being 0.39 [25]. A trait with a higher heritability can be more accurately selected. It is, therefore, vitally important for sheep production and breeding that the complex molecular mechanisms behind sheep body weight and the important functional genes that influence this trait are identified.

Population stratification is an important factor causing false positive GWAS results [26]. Studies have shown that using a mixed linear model that simulates population structure, kinship, and family structure is the most effective method of reducing the effects of population stratification currently available [24]. Therefore, in this study, a population genetic structure was included as a fixed effect and individual kinship as a random effect, to correct for the effects of population structure and individual kinship relationships. It can be seen from the Q-Q plots of the four body weight traits that no population stratification phenomena were apparent, based on the observed and expected  $p$ -values. However, this phenomenon was apparent for the SNPs, which were highly correlated with the body weight traits. The  $p$ -values of the significantly correlated SNPs in this study were smaller than  $10^{-6}$ , so it is probable that these SNP correlations do exist.

The genome-wide association analysis of the birth, weaning, yearling, and adult weights of 460 fine-wool sheep was performed using a mixed linear model in this study, and a total of 113 SNPs reached genome-wide significance after detection. Interestingly, except for sex chromosomes, significantly correlated genes on autosomes in the four body weight traits were all different. We, therefore, speculated that the sheep body weight was controlled by different genes during different phases, and that the body weight within a single phase was regulated by many genes. Sheep body weight is a quantitative trait that is influenced by both the genotype and the environment. The polygene hypothesis for quantitative traits, therefore, supports this speculation that body weight is regulated by micro-effect genes. It has also been found in a GWAS study of live weights in chickens that weight traits are regulated by micro-effect genes [27]. Gene expression is selective and affected by time, space, and environmental factors. Except for a very small number of genes that maintain stable expression in any external environment (e.g.,  $\beta$ -actin genes), whether most genes are expressed and the level of expression, are regulated based on changes in the external environment. Environmental factors, include material aspects, such as nutritional status. However, psychological factors, such as stress are also relevant. These environmental factors can change the endocrine system of the body, which in turn changes the expression of genes. The fine-wool sheep that we chose are mainly grazing animals and the nutritional level of their pastures varies among the seasons, which affects their gene expression. It has also been found in other studies that different levels of feed nutrition can alter gene expression [28,29]. Secondly, grazing sheep, stimulated by hot summers and cold winters, can also change gene expression [30,31].

In addition, we found that the candidate body weight genes, identified in this study, had poor repeatability, compared with previously reported GWAS results for sheep body weight [4,7,10,32]. In previous sheep body weight GWAS, variation detection was performed using BeadChips, so only a very limited number of mutations were detected. The number of mutations that can be detected using the resequencing technique, used in this study, is several orders of magnitude higher than can be identified using BeadChips. Furthermore, the four sheep breeds selected for this study were established by long-term artificial selection and the linkage disequilibrium level of their genomes was thus very high. As a result, some SNPs unrelated to the measured traits displayed rising signal values, caused by linkage with target SNPs, which may have influenced the signal values of the target SNPs.

Body weight growth is closely related to the growth of muscle, fat, and bone tissues; candidate genes that correlated with the growth and development of these tissues were identified in this study. The *AADACL3* (arylacetamide deacetylase like 3) gene located on chromosome 12 that correlated with birth weight participates in fat metabolism [33]. The *VGF* (nerve growth factor inducible) gene on chromosome 24, which correlates with weaning weight, is involved in regulating the food intake and body weight of animals, with the knockout of the *VGF* gene in mice leading to weight reduction, in vivo fat reduction, and excessive energy consumption [34–37]. The *VGF* gene may also regulate fat synthesis and decomposition [38,39]. The *NPC1* (niemann-Pick Type C) gene on chromosome 23 that

correlated with yearling weight participates in controlling the steady state of animal energy metabolism and plays an important regulatory role in body weight and fat metabolism [40–42]. The *SERPINA12* (serpin family A member 12) gene on chromosome 18 that correlated with adult weight is closely associated with animal body composition and lipid distribution, and could serve as a marker for lipid metabolism [43,44].

Many of the genes that were found to be associated with significantly correlated SNPs in this study were enriched in GO terms, such as actin filament binding ( $p < 0.05$ ). The long-term and complex process of animal muscle development mainly relies on the proliferation and hypertrophy of muscle fibers [45]. Actin plays a vital role in the transformation of myotubes into myofibers [46]. Many of the other significantly-enriched GO terms, such as microtubule and cytoskeletal protein, are also essential for muscle development.

The KEGG pathway analysis revealed that the wnt signaling pathway was remarkably enriched for the adult weight trait. Belonging to the secretory glycoprotein family, wnt participates in the proliferation and differentiation of multiple precursor cells [47]. Previous research has found that the wnt signaling pathway is indispensable for embryonic and postnatal skeletal muscle homeostasis [48]. During embryonic skeletal muscle development, the wnt signaling pathway induces myogenesis, mainly by regulating *MRFs* (myogenic regulatory factors) [49,50]. During skeletal muscle development, after birth, the classical wnt signaling pathway mainly regulates the differentiation of skeletal muscle satellite cells, while the nonclassical wnt signaling pathway mainly mediates the self-renewal of skeletal muscle satellite cells and the growth of muscle fibers [51,52]. Bone growth and development are closely related to body weight, with bone development preceding skeletal muscle development, and the wnt signaling pathway also exerts an important effect on the bone development process [53,54]. However, in addition to the wnt signaling pathway, the role of the Jak-STAT (janus kinase- signal transducer and activator of transcription) signaling pathway in satellite cells has also gradually been revealed in recent years. The Jak1/STAT1/STAT3 pathway promotes the proliferation of activated satellite cells and prevents their premature differentiation into myotubes. The Jak2/STAT2/STAT3 pathway also mediates the positive regulation of satellite cell differentiation by *MyoD* (myoblast determination protein) and *MEF2* (myocyte enhancer factor 2) [55,56]. Some GO terms related to lipid metabolism were also enriched and these GO terms were also impacted by the environment in which the fine-wool sheep selected for this study were located. The four fine-wool breeds selected for this study were mainly grazed in areas that had severe fodder grass shortages during spring and winter. As they lived in a harsh environment, lipid metabolism was critically important for their production and reproduction. Furthermore, this hostile environment also impacted the sheep's health, and we found that the significantly enriched pathways included some signaling pathways related to body immunity, such as the toll-like receptor signaling pathway.

## 5. Conclusions

In conclusion, this study identified genomic regions, and gene-associated biological pathways, related to the body weight traits of Chinese fine-wool sheep. The identified candidate genes were found to mainly participate in skeletal muscle development and lipid metabolism. The genes identified are good candidates for further functional validation to uncover the biological mechanisms underlying body weight variation in Chinese fine-wool sheep.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2076-2615/10/1/170/s1>, Figure S1: Violin plots showing distribution of the body weight in each group; Figure S2: Population structure and distribution of LD for four fine-wool sheep breeds; Table S1: Sequencing reads, alignment statistics and mean genome-wide coverage of each sample. Table S2: GO analysis for the regional candidate genes with genome-wide significant association. Table S3: KEGG analysis for the regional candidate genes with genome-wide significant association.

**Author Contributions:** Conceptualization, B.Y.; formal analysis, Z.L., Z.C., and T.G.; investigation, Z.L., Y.Y., C.Y., J.L., C.N., and X.S.; resources, H.Z. and S.Z.; writing—original draft preparation, Z.L.; writing—review and

editing, Z.L.; project administration, B.Y. and Z.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Modern China Wool Cashmere Technology Research System (CARS-39-02), the Agricultural Science and Technology Innovation Program of Chinese Academy of Agricultural Sciences (CAAS-ASTIP-2015-LIHPS), and the Major Output Research Topic of Chinese Academy of Agricultural Sciences (CAAS-ZDXT2018006).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Hirschhorn, J.N.; Daly, M.J. Genome-wide association studies for common diseases and complex traits. *Nat. Rev. Genet.* **2005**, *6*, 95–108. [[CrossRef](#)] [[PubMed](#)]
- Gao, N.; Chen, Y.L.; Liu, X.H.; Zhao, Y.X.; Zhu, L.; Liu, A.L.; Jiang, W.; Peng, X.; Zhang, C.L.; Tang, Z.S.; et al. Weighted single-step GWAS identified candidate genes associated with semen traits in a Duroc boar population. *BMC Genomics* **2019**, *20*, 797. [[CrossRef](#)] [[PubMed](#)]
- Li, B.; Fang, L.; Null, D.J.; Hutchison, J.L.; Connor, E.E.; VanRaden, P.M.; VandeHaar, M.J.; Tempelman, R.J.; Weigel, K.A.; Cole, J.B. High-density genome-wide association study for residual feed intake in Holstein dairy cattle. *J. Dairy Sci.* **2019**, *102*, 11067–11080. [[CrossRef](#)] [[PubMed](#)]
- Ghasemi, M.; Zamani, P.; Vatankhah, M.; Abdoli, R. Genome-wide association study of birth weight in sheep. *Animals* **2019**, *13*, 1797–1803. [[CrossRef](#)]
- Liu, Z.; Yang, N.; Yan, Y.Y.; Li, G.Q.; Liu, A.Q.; Wu, G.Q.; Sun, C.J. Genome-wide association analysis of egg production performance in chickens across the whole laying period. *BMC Genet.* **2019**, *20*, 67. [[CrossRef](#)]
- Wei, C.H.; Liu, C.S. *Encyclopedia of Modern Sheep Production Technology*; China Agriculture Press: Beijing, China, 2014; pp. 70–77.
- Al-Mamun, H.A.; Kwan, P.; Clark, S.A.; Ferdosi, M.H.; Tellam, R.; Gondro, C. Genome-wide association study of body weight in Australian Merino sheep reveals an orthologous region on OAR6 to human and bovine genomic regions affecting height and weight. *Genet. Sel. Evol.* **2015**, *47*, 66. [[CrossRef](#)]
- Matika, O.; Riggio, V.; Anselme-Moizan, M.; Law, A.S.; Pong-Wong, R.; Archibald, A.L.; Bishop, S.C. Genome-wide association reveals QTL for growth, bone and in vivo carcass traits as assessed by computed tomography in Scottish Blackface lambs. *Genet. Sel. Evol.* **2016**, *48*, 11. [[CrossRef](#)]
- Rupp, R.; Senin, P.; Sarry, J.; Allain, C.; Tasca, C.; Ligat, L.; Portes, D.; Woloszyn, F.; Bouchez, O.; Tabouret, G.; et al. A point mutation in suppressor of cytokine signalling 2 (Socs2) increases the susceptibility to inflammation of the mammary gland while associated with higher body weight and size and higher milk production in a sheep model. *PLoS Genet.* **2015**, *11*, e1005629. [[CrossRef](#)]
- Zhang, L.; Liu, J.S.; Zhao, F.P.; Ren, H.X.; Xu, L.Y.; Lu, J.; Zhang, S.F.; Zhang, X.N.; Wei, C.H.; Lu, G.B.; et al. Genome-wide association studies for growth and meat production traits in sheep. *PLoS ONE* **2013**, *8*, e66569. [[CrossRef](#)]
- Riggio, V.; Matika, O.; Pong-Wong, R.; Stear, M.J.; Bishop, S.C. Genome-wide association and regional heritability mapping to identify loci underlying variation in nematode resistance and body weight in Scottish Blackface lambs. *Heredity* **2013**, *110*, 420–429. [[CrossRef](#)]
- Fox, J.; Weisberg, S. *An R Companion to Applied Regression*, 2nd ed.; SAGE Press: Thousand Oaks, CA, USA, 2011; pp. 285–287.
- R Core Team. R: A Language and Environment for Statistical Computing. Available online: <https://www.R-project.org/> (accessed on 15 October 2019).
- Listed, N. References to molecular biology techniques. *Curr. Protoc. Hum. Genet.* **2001**, *19*, A.3A.1–A.3A.2.
- Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)] [[PubMed](#)]
- Li, H.; Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **2010**, *26*, 589–595. [[CrossRef](#)]
- Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R. The sequence alignment/map format and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079. [[CrossRef](#)]
- Wang, K.; Li, M.Y.; Hakonarson, H. ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **2010**, *38*, e164. [[CrossRef](#)]

19. Yu, J.M.; Pressoir, G.; Briggs, W.H.; Vroh Bi, I.; Yamasaki, M.; Doebley, J.F.; McMullen, M.D.; Gaut, B.S.; Nielsen, D.M.; Holland, J.B.; et al. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* **2006**, *38*, 203–208. [[CrossRef](#)]
20. Zhou, X.; Stephens, M. Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.* **2012**, *44*, 821–824. [[CrossRef](#)]
21. Nicodemus, K.K.; Liu, W.; Chase, G.A.; Tsai, Y.Y.; Fallin, M.D. Comparison of type I error for multiple test corrections in large single-nucleotide polymorphism studies using principal components versus haplotype blocking algorithms. *BMC Genet.* **2005**, *6*, S78. [[CrossRef](#)]
22. Huang da, W.; Sherman, B.T.; Lempicki, R.A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **2009**, *4*, 44–57. [[CrossRef](#)]
23. Huang da, W.; Sherman, B.T.; Lempicki, R.A. Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **2009**, *37*, 1–13. [[CrossRef](#)]
24. Price, A.L.; Zaitlen, N.A.; Reich, D.; Patterson, N. New approaches to population stratification in genome-wide association studies. *Nat. Rev. Genet.* **2010**, *11*, 459–463. [[CrossRef](#)] [[PubMed](#)]
25. Zhao, Y.Z. *Sheep Production*; China Agriculture Press: Beijing, China, 2013; pp. 594–595.
26. Price, A.L.; Patterson, N.J.; Plenge, R.M.; Weinblatt, M.E.; Shadick, N.A.; Reich, D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **2006**, *38*, 904–909. [[CrossRef](#)] [[PubMed](#)]
27. Yuan, Y.Y.; Peng, D.Z.; Gu, X.R.; Gong, Y.Z.; Sheng, Z.Y.; Hu, X.X. Polygenic basis and variable genetic architectures contribute to the complex nature of body weight -a genome-wide study in four Chinese indigenous chicken breeds. *Front. Genet.* **2018**, *9*, 229. [[CrossRef](#)]
28. Dervishi, E.; Joy, M.; Alvarez-Rodriguez, J.; Serrano, M.; Calvo, J.H. The forage type (grazing versus hay pasture) fed to ewes and the lamb sex affect fatty acid profile and lipogenic gene expression in the longissimus muscle of suckling lambs. *J. Anim. Sci.* **2012**, *90*, 54–66. [[CrossRef](#)]
29. Metzler-Zebeli, B.U.; Hollmann, M.; Sabitzer, S.; Podstatzky-Lichtenstein, L.; Klein, D.; Zebeli, Q. Epithelial response to high-grain diets involves alteration in nutrient transporters and Na<sup>+</sup>/K<sup>+</sup>-ATPase mRNA expression in rumen and colon of goats. *J. Anim. Sci.* **2013**, *91*, 4256–4266. [[CrossRef](#)]
30. Lu, Z.K.; Chu, M.X.; Li, Q.; Jin, M.L.; Fei, X.J.; Ma, L.; Zhang, L.P.; Wei, C.H. Transcriptomic analysis provides novel insights into heat stress responses in sheep. *Animals* **2019**, *9*, 387. [[CrossRef](#)]
31. Jing, X.P.; Peng, Q.H.; Hu, R.; Zou, H.W.; Wang, H.Z.; Yu, X.Q.; Zhou, J.W.; Degen, A.; Wang, Z.S. Dietary supplements during the cold season increase rumen microbial abundance and improve rumen epithelium development in Tibetan sheep. *J. Anim. Sci.* **2018**, *96*, 293–305. [[CrossRef](#)]
32. Gholizadeh, M.; Rahimi-Mianji, G.; Nejati-Javaremi, A. Genomewide association study of body weight traits in Baluchi sheep. *J. Genet.* **2015**, *94*, 143–146. [[CrossRef](#)]
33. Zhang, W.C.; Bin, Y.; Zhang, J.J.; Cui, L.L.; Ma, J.W.; Chen, C.Y.; Ai, H.S.; Xiao, S.J.; Ren, J.; Huang, L.S. Genome-wide association studies for fatty acid metabolic traits in five divergent pig populations. *Sci. Rep.* **2016**, *6*, 24718. [[CrossRef](#)]
34. Altshuler, D.; Hirschhorn, J.N. Upsetting the balance: VGF and the regulation of body weight. *Neuron* **1999**, *23*, 415–417. [[CrossRef](#)]
35. Jiang, C.; Lin, W.J.; Sadahiro, M.; Shin, A.C.; Buettner, C.; Salton, S.R. Embryonic ablation of neuronal VGF increases energy expenditure and reduces body weight. *Neuropeptides* **2017**, *64*, 75–83. [[CrossRef](#)] [[PubMed](#)]
36. Lewis, J.E.; Brameld, J.M.; Hill, P.; Cocco, C.; Noli, B.; Ferri, G.L.; Barrett, P.; Ebling, F.J.; Jethwa, P.H. Hypothalamic over-expression of VGF in the Siberian hamster increases energy expenditure and reduces body weight gain. *PLoS ONE* **2017**, *12*, e0172724.
37. Foglesong, G.D.; Huang, W.; Liu, X.L.; Slater, A.M.; Siu, J.; Yildiz, V.; Salton, S.R.; Cao, L. Role of hypothalamic VGF in energy balance and metabolic adaptation to environmental enrichment in mice. *Endocrinology* **2016**, *157*, 983–996. [[CrossRef](#)] [[PubMed](#)]
38. Fargali, S.; Scherer, T.; Shin, A.C.; Sadahiro, M.; Buettner, C.; Salton, S.R. Germline ablation of VGF increases lipolysis in white adipose tissue. *J. Endocrinol.* **2012**, *215*, 313–322. [[CrossRef](#)] [[PubMed](#)]
39. Watson, E.; Hahm, S.; Mizuno, T.M.; Windsor, J.; Montgomery, C.; Scherer, P.E.; Mobbs, C.V.; Salton, S.R. VGF ablation blocks the development of hyperinsulinemia and hyperglycemia in several mouse models of obesity. *Endocrinology* **2005**, *146*, 5151–5163. [[CrossRef](#)]

40. Jelinek, D.; Heidenreich, R.A.; Erickson, R.P.; Garver, W.S. Decreased *Npc1* gene dosage in mice is associated with weight gain. *Obesity* **2010**, *18*, 1457–1459. [[CrossRef](#)] [[PubMed](#)]
41. Jelinek, D.; Millward, V.; Birdi, A.; Trouard, T.P.; Heidenreich, R.A.; Garver, W.S. *Npc1* haploinsufficiency promotes weight gain and metabolic features associated with insulin resistance. *Hum. Mol. Genet.* **2011**, *20*, 312–321. [[CrossRef](#)]
42. Bambace, C.; Dahlman, I.; Arner, P.; Kulyte, A. NPC1 in human white adipose tissue and obesity. *BMC Endocr. Disord.* **2013**, *13*, 5. [[CrossRef](#)]
43. Gajewska, J.; Kurylowicz, A.; Mierzejewska, E.; Ambroszkiewicz, J.; Chelchowska, M.; Weker, H.; Pietrzykowska, M.; Puzianowska-Kuznicka, M. Are omentin rs2274907 and vaspin rs2236242 gene polymorphisms related to body composition, lipid profile and other adipokines in prepubertal healthy children? *Endocr. Res.* **2019**, *15*, 1–8. [[CrossRef](#)]
44. Breitfeld, J.; Wiele, N.; Gutschmann, B.; Stumvoll, M.; Bluher, M.; Scholz, M.; Kovacs, P.; Tonjes, A. Circulating adipokine VASPIN is associated with serum lipid profiles in Humans. *Lipids* **2019**, *54*, 203–210. [[CrossRef](#)]
45. Wigmore, P.M.; Stickland, N.C. Muscle development in large and small pig fetuses. *J. Anat.* **1983**, *137*, 235–245. [[PubMed](#)]
46. Picard, B.; Berri, C.; Lefaucheur, L.; Molette, C.; Sayd, T.; Terlouw, C. Skeletal muscle proteomics in livestock production. *Brief. Funct. Genom.* **2010**, *9*, 259–278. [[CrossRef](#)] [[PubMed](#)]
47. Ng, L.F.; Kaur, P.; Bunnag, N.; Suresh, J.; Sung, I.C.H.; Tan, Q.H.; Gruber, J.; Tolwinski, N.S. WNT signaling in disease. *Cells* **2019**, *8*, 826. [[CrossRef](#)]
48. Abu-Elmagd, M.; Robson, L.; Sweetman, D.; Hadley, J.; Francis-West, P.; Munsterberg, A. Wnt/Lef1 signaling acts via *Pitx2* to regulate somite myogenesis. *Dev. Biol.* **2010**, *337*, 211–219. [[CrossRef](#)]
49. Jones, A.E.; Price, F.D.; Le Grand, F.; Soleimani, V.D.; Dick, S.A.; Megeney, L.A.; Rudnicki, M.A. Wnt/beta-catenin controls follistatin signalling to regulate satellite cell myogenic potential. *Skelet. Muscle* **2015**, *5*, 14. [[CrossRef](#)]
50. Murphy, M.M.; Keefe, A.C.; Lawson, J.A.; Flygare, S.D.; Yandell, M.; Kardon, G. Transiently active Wnt/beta-catenin signaling is not required but must be silenced for stem cell function during muscle regeneration. *Stem Cell Rep.* **2014**, *3*, 475–488. [[CrossRef](#)]
51. Rudnicki, M.A.; Williams, B.O. Wnt signaling in bone and muscle. *Bone* **2015**, *80*, 60–66. [[CrossRef](#)] [[PubMed](#)]
52. Habib, S.J.; Chen, B.C.; Tsai, F.C.; Anastassiadis, K.; Meyer, T.; Betzig, E.; Nusse, R. A localized Wnt signal orients asymmetric stem cell division in vitro. *Science* **2013**, *339*, 1445–1448. [[CrossRef](#)]
53. Karner, C.M.; Long, F. Wnt signaling and cellular metabolism in osteoblasts. *Cell. Mol. Life Sci.* **2017**, *74*, 1649–1657. [[CrossRef](#)]
54. Maeda, K.; Kobayashi, Y.; Koide, M.; Uehara, S.; Okamoto, M.; Ishihara, A.; Kayama, T.; Saito, M.; Marumo, K. The regulation of bone metabolism and disorders by wnt signaling. *Int. J. Mol. Sci.* **2019**, *20*, 5525. [[CrossRef](#)]
55. Trenerry, M.K.; Della Gatta, P.A.; Cameron-Smith, D. JAK/STAT signaling and human in vitro myogenesis. *BMC Physiol.* **2011**, *11*, 6. [[CrossRef](#)] [[PubMed](#)]
56. Price, F.D.; von Maltzahn, J.; Bentzinger, C.F.; Dumont, N.A.; Yin, H.; Chang, N.C.; Wilson, D.H.; Frenette, J.; Rudnicki, M.A. Inhibition of JAK-STAT signaling stimulates adult satellite cell function. *Nat. Med.* **2014**, *20*, 1174–1181. [[CrossRef](#)] [[PubMed](#)]

