





Article

A Study of the Resistance of Hu Sheep Lambs to *Escherichia coli* F17 Based on Whole Genome Sequencing

Yanjun Duan ¹, Pengwei Su ², Yifei Gu ², Xiaoyang Lv ^{3,4}, Xiukai Cao ^{3,4}, Shanhe Wang ^{2,4}, Zehu Yuan ^{3,4} and Wei Sun ^{2,3,4,*}

¹ College of Veterinary Medicine, Yangzhou University, Yangzhou 225009, China; yjduan1996@163.com

² College of Animal Science and Technology, Yangzhou University, Yangzhou 225009, China; supengwei4901@163.com (P.S.); feilin133082@163.com (Y.G.); 007121@yzu.edu.cn (S.W.)

³ Joint International Research Laboratory of Agriculture and Agri-Product Safety of Ministry of Education of China, Yangzhou University, Yangzhou 225009, China; dx120170085@yzu.edu.cn (X.L.); cckai0909@163.com (X.C.); yuanzehu1988@163.com (Z.Y.)

⁴ International Joint Research Laboratory in Universities of Jiangsu Province of China for Domestic Animal Germplasm Resources and Genetic Improvement, Yangzhou University, Yangzhou 225009, China

* Correspondence: dkxmsunwei@163.com

Simple Summary: Acute infectious diseases caused by *Escherichia coli* in lambs or young sheep have severe impacts on the growth and development of sheep, among which *Escherichia coli* F17 (*Escherichia coli* F17, *E. coli* F17) is one of the primary pathogens causing bacterial diarrhoea in lambs. The mechanism of lambs' resistance regulation to it remains to be elucidated.

Abstract: This study aims to analyze the whole genome sequencing of *E. coli* F17 in antagonistic and susceptible Hu sheep lambs. The objective is to investigate the critical mutation loci in sheep and understand the genetic mechanism of sheep resistance to *E. coli* F17 at the genome level. Antagonist and susceptible venous blood samples were collected from Hu sheep lambs for whole genome sequencing and whole genome association analysis. A total of 466 genes with significant SNPs ($p < 1.0 \times 10^{-3}$) were found. GO and KEGG enrichment analysis and protein interaction network analysis were performed on these genes, and preliminary investigations showed that SNPs on CTNNB1, CDH8, APOD, HCLS1, Tet2, MTSS1 and YAP1 genes may be associated with the antagonism and susceptibility of Hu sheep lambs to *E. coli* F17. There are still some shortcomings that have not been explored via in vivo and in vitro functional experiments of the candidate genes, which will be our next research work. This study provides genetic loci and candidate genes for resistance of Hu sheep lambs to *E. coli* F17 infection, and provides a genetic basis for breeding disease-resistant sheep.

Keywords: whole genome sequencing; *Escherichia coli* F17; Hu sheep lambs; antagonistic phenotype; susceptible phenotype



Citation: Duan, Y.; Su, P.; Gu, Y.; Lv, X.; Cao, X.; Wang, S.; Yuan, Z.; Sun, W. A Study of the Resistance of Hu Sheep Lambs to *Escherichia coli* F17 Based on Whole Genome Sequencing. *Animals* **2024**, *14*, 161. <https://doi.org/10.3390/ani14010161>

Academic Editor: Marco Milanese

Received: 20 November 2023

Revised: 27 December 2023

Accepted: 29 December 2023

Published: 3 January 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Sheep, one of the six domesticated animal species, occupy an important position in the livestock industry, and sheep milk, mutton, skin and fleece provide significant economic benefits to the farming industry. *E. coli* disease has a major impact on the entire sheep production cycle, particularly during the lambing and fattening period; acute infectious diseases caused by *E. coli* in lambs or young sheep seriously affect the growth and development of sheep [1].

Intestinal pathogenic *E. coli* types are enteropathogenic *E. coli* (EPEC); Shiga toxin-producing *E. coli* (STEC); enterotoxin-producing *E. coli* (ETEC); enteroaggregative *E. coli* (EAEC); enteroinvasive *E. coli* (EIEC); and diffusely adherent *E. coli* (DAEC) [2]. *E. coli* F17 is mainly detected in diarrhoeal or septic calves or lambs. F17-mediated *E. coli* is mainly composed of the structural subunit F17-A and the adhesin subunit F17-G [3]. Four

structural subunits of F17-A have been identified so far: F17a-A, F17b-A, F17c-A, and F17d-A [4–7]. F17-G has been identified as three adhesin subunits: F17a-G, F17b-G and F17c-G [8]. In a related study, 45 nonenterotoxin-producing *Escherichia coli* strains from sheep and goats that produced F17-associated hyphae were characterized by hyphal structural subunits and adhesin subtypes; the majority of the sheep and goat strains showed phenotypic characteristics of the septic strain [9]. In addition, *E. coli* F17 was detected in the diarrheal feces of lambs in several regions, indicating that *E. coli* F17 is a pathogenic pathogen for lambs [10].

ETEC adapts to its environment in various ways, including by the development of different strains that express multiple colonization factors (CFs) that adhere to the intestinal epithelium and secrete numerous enterotoxins. To avoid ETEC infection, it is necessary to have a deep understanding of its pathogenic mechanism. ETEC secretes colonization factors (CFs) and enterotoxin, two virulence factors which bind to specific receptors of intestinal epithelial cells to induce diarrhoea [11]. In addition, intestinal epithelial cells (IECs) recognize pathogen-associated molecular patterns (PAMPs) through pathogen recognition receptors (PRRs), which secrete a variety of cytokines and chemokines to trigger innate immune defences and promote adaptive immune responses [12,13]. In ETEC infection, the integrity of the intestinal barrier plays a vital role in resisting the attack of various pathogenic microorganisms [14,15]. The physical barrier formed by intestinal epithelial cells and tight intercellular connections is the first line of defence against harmful substances and the basis for maintaining the function and structure of the intestinal barrier [16]. Mucin secreted by goblet cells in intestinal epithelial cells forms a mucus layer, separating intestinal lumen microorganisms from epithelial cells and preventing the invasion of toxins and pathogenic microorganisms [17]. Therefore, the colonization and adhesion of ETEC in intestinal epithelial cells, binding to specific receptors, maintenance of the intestinal barrier function, and activation of the immune system can affect the invasion of ETEC into the body.

To fundamentally solve bacterial diarrhoea caused by *E. coli* F17 in lambs, we should understand the mechanism of diarrhoea and implement disease-resistance breeding accordingly. In this study, we performed whole genome sequencing of antagonistic and susceptible sheep individuals at the genomic level to screen the key SNPs' resistant to *E. coli* F17 infection and provide theoretical support for breeding resistance to the disease.

2. Materials and Methods

2.1. Ethical Statement

All animal experiments were reviewed and approved by the Experimental Animal Welfare and Ethical Institute of Animal Science, Yangzhou University (No: NFNC2020-NFY-6), and were performed in accordance with the Regulations for the Administration of Affairs Concerning Experimental Animals approved by the State Council of the People's Republic of China.

2.2. Sample Collection

Fifty healthy newborn lambs were randomly selected and fed lamb milk without antimicrobial additives and probiotics between 1 and 3 days of age. After 3 days, the lambs were divided into a high-dose group and a low-dose group, and the lambs in each group were challenged by gavage with 50 mL and 1 mL of *E. coli* F17 culture (1×10^9 CFU/mL), respectively, for 4 days. Ten healthy lambs in the high-dose challenge group (antagonistic type) and ten low-dose-challenged lambs with diarrhoea (sensitive type) were selected, according to intestinal pathological sections and the amount of bacterial adherence in the intestinal contents. Six lambs were selected as case and control groups in the antagonistic and sensitive groups [18]. Two mL of venous blood was taken in EDTA-K2 anticoagulant tubes and stored at -80°C . Samples were named as containing A for the antagonistic type and not containing A for the susceptible type.

2.3. Sheep Blood Genomic DNA Extraction and Sequencing

Spectrophotometer, Qubit, and agarose gel electrophoresis were used to identify the concentration and purity of the extracted DNA, and the qualified DNA samples were randomly interrupted using Covaris to form 350–500 bp fragments after quality control. Library construction was then carried out using the TruSeq DNA LT Sample Prep kit, in which the DNA fragments were subjected to end repair, poly(A) tailing, sequencing junction addition, purification, PCR amplification, and other steps to complete the library construction. After constructing the library, a sequencer subjected the DNA fragments to bipartite sequencing. The library construction process included the following: fragment DNA, repair ends and select library size, adenylate 3' ends, ligate adapters, enrich DNA fragments, normalize, and pool libraries. OE Biotech Co., Ltd. (Shanghai, China) performed the sequencing and analysis. The libraries were constructed with TruSeq Nano DNA LT Sample Preparation Kit (Illumina, San Diego, CA, USA).

2.4. Sequencing Data Filtering

After the raw data was obtained, the sequencing data quality evaluation was conducted. The sequencing error rate, data volume, comparison rate, coverage, etc., were counted to evaluate whether the sequencing of the database construction met the standard, and follow-up analysis was carried out if the standard was met. In addition, variant information analysis was conducted, and high-quality sequences were compared to the reference genome. The variant information in the sample was detected, and the detected mutation information was analyzed and interpreted. High-throughput sequencing data off the sequencing machine had several sequencing errors, which were caused by the combined action of multiple factors such as the sequencer itself, sequencing reagents, and samples. In general, the error rate of the first few bases and ends of a sequencing fragment can be high, caused by the instability of the sequencing starting instrument and the consumption of chemical reagents during sequencing. To eliminate the effect of sequencing errors on the results, the raw data needed to be quality controlled to obtain clean reads. The preprocessing software was fastp [19]. The quality filtration criteria were as follows: removal of adapter sequences, and reads with N (non-AGCT) bases greater than or equal to 5; sliding window with 4 bases for Windows size, and excise if the average base quality value was less than 20 (see note below for base quality value algorithm); after the above filtering, reads with lengths less than 75 bp or an average base quality value less than 15 were removed.

2.5. Reference Genome Matching

Clean reads were aligned to reference genomes using BWA [20]; the mapping results were converted to the BAM format by SAMtools [21]. The results were de-redundant using Picard, and analyzed using Qualimap software comparison with the reference genome Oar_rambouillet_v1.0.

2.6. SNP Mutation Detection

SNP detection was performed using the Haplotypcaller module of GATK4 [22] software based on the comparison of the samples with the reference genome. To reduce the error rate of SNP detection, the criterion of $QD \geq 2.0$ was chosen for filtering, and only the mutant loci that meet this criterion were retained. $QD:QD$ is the ratio of the quality of a variant divided by the depth of coverage. This is the quality of a variant per unit of depth; most of the false positive variants had a QD value of less than 2. The SNP detection results were annotated using Annovar [23] software.

2.7. Genome-Wide Association Analysis

Case control was performed using PLINK v. 1.09, comparing allele frequencies between cases and controls using the chi-square test, a p -value with a correlation statistic that indicates whether the observed difference is significant or not. After genotyping,

a total of 22988506 SNPs were obtained. Bonferroni correction was used in this study; the suggestive association significance threshold was $p < 0.05$ ($p < 5.87 \times 1.0 \times 10^{-8}$) = ($p < 0.05 / (22988506 / 27)$) at the chromosome-wide level. The significant association indicated that the chromosome-wide level association corresponded to a p -value less than 1.0×10^{-3} [24]. The association result and significant SNPs were visualized in a Manhattan plot with a threshold line. The Manhattan function was plotted with R v. 3.6.0 [25].

2.8. GO and KEGG Enrichment Analysis

Based on the hypergeometric distribution, GO [26] and KEGG [27] pathway enrichment analysis of differential genes were performed to screen the significantly enriched term using R (v 3.2.0). R (v 3.2.0) was used to draw the significant enrichment term's column, chord, and bubble diagram. Between them, GO level2 was the secondary classification of GO.

2.9. Protein Interaction Network Analysis

The interaction relationship in the STRING [28] protein interaction database was used to analyze the differential gene protein interaction network. The differential gene collection was extracted from the database, and the interaction relationship network diagram was constructed by cytoscape [29]. In this study, we performed protein network interactions with proteins related to bacterial infectious diseases, and selected crucial essential node proteins in the entire interaction network.

3. Results

3.1. Quality Control of whole Genome Sequencing Data of Antagonistic and Susceptible Hu Sheep Lambs

Genome sequencing of 12 blood samples from Hu sheep lambs was completed and a total of 435 Gb of clean reads with high sequencing quality ($Q20 \geq 97.13\%$, $Q30 \geq 91.80\%$) were obtained. The comparison rate of all samples was in the range of 97.80–98.30%; the specific quality control results are shown in Table 1. All detected mutations were integrated into a Circos diagram to visualize genomic data, including the density distribution of genomic SNPs on each chromosome (Figure S1), and reference genome alignment data (Figure S2).

Table 1. Summary of QC results.

Samples	R-R	C-R	C-R-P	R-B	C-B	C-B-P	G-C	>Q20	>Q30
s2S-250	328124474	323466836	98.58%	49.22G	48.38G	98.30%	43.32%	97.56%	93.05%
s2S-936	316593058	311927296	98.53%	47.49G	46.66G	98.26%	43.47%	97.52%	92.96%
s2S-088	280099462	275608400	98.40%	42.01G	41.22G	98.10%	43.24%	97.08%	91.80%
s2S-037	291558274	286959356	98.42%	43.73G	42.89G	98.06%	43.41%	97.30%	92.28%
s2S-341	298812648	293506672	98.22%	44.82G	43.84G	97.80%	43.63%	97.13%	91.87%
s2S-940	292608542	288114728	98.46%	43.89G	43.05G	98.08%	43.58%	97.24%	92.08%
s2A-926	229463262	225940978	98.46%	34.42G	33.80G	98.21%	43.48%	97.55%	93.06%
s2A-965	251932032	248203718	98.52%	37.79G	37.10G	98.17%	43.40%	97.69%	93.38%
s2A-349	253327108	249116262	98.34%	38.00G	37.22G	97.94%	43.76%	97.19%	92.02%
s2A-187	266199676	261823920	98.36%	39.93G	39.13G	98.01%	43.37%	97.14%	91.90%
s2A-982	247628212	243590354	98.37%	37.14G	36.40G	97.99%	43.60%	97.23%	92.22%
s2A-937	229237666	225756938	98.48%	34.39G	33.69G	97.98%	43.09%	97.45%	92.68%

Note: R-R (raw reads), C-R (clean reads), C-R-P (clean reads percent), R-B (raw base), C-B (clean base), C-B-P (clean base percent), G-C (GC content).

3.2. Genome-Wide Association Analysis

An association analysis was performed on 22988506 SNPs distributed in 27 chromosomes via whole-genome sequencing analysis of antagonistic and susceptible Hu sheep lamb blood DNA (Figure 1). There were a total of 6868 significant SNPs ($p < 1.0 \times 10^{-3}$) that are annotated to 466 genes (Table S1).

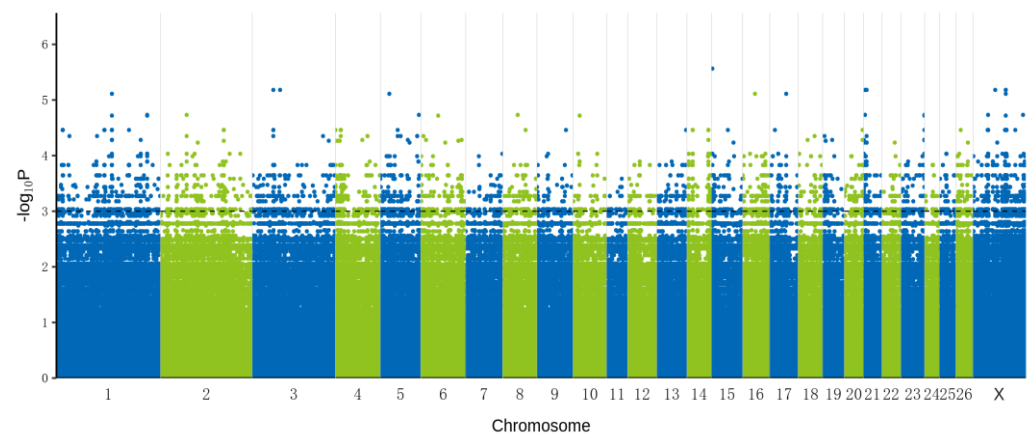


Figure 1. Manhattan plot showing SNPs on sheep chromosomes associated with Hu sheep lambs' ability to resist *E. coli* F17. The black dashed line corresponds to a suggestive chromosome-wide threshold of 1.0×10^{-3} .

3.3. GO and KEGG Enrichment Analysis

Genome-wide association analysis was used to select SNPs associated at the chromosome level ($p < 1.0 \times 10^{-3}$); there were 466 genes with significant SNPs for GO and KEGG enrichment analysis. The GO level2 (Figure 2 and Table S2) enrichment results show a total of 45 GO level2 entries, of which GO entries classified as biological processes had the most significant number of GO entries with 21 entries, followed by cell composition with 15 entries. Molecular function ontology had the fewest entries with 9 enrichment items. Biological processes associated with bacterial infections included biological adhesion, immune system process, cell junction, and receptor regulator activity. The KEGG enrichment results (Figure 3 and Table S3) listed the top 20 items ($p < 0.05$) with significant enrichment, among them circadian entrainment, insulin secretion, the Wnt signaling pathway, the thyroid hormone signaling pathway, and the Hippo signaling pathway. Multiple species may be associated with bacterial infectious diseases.

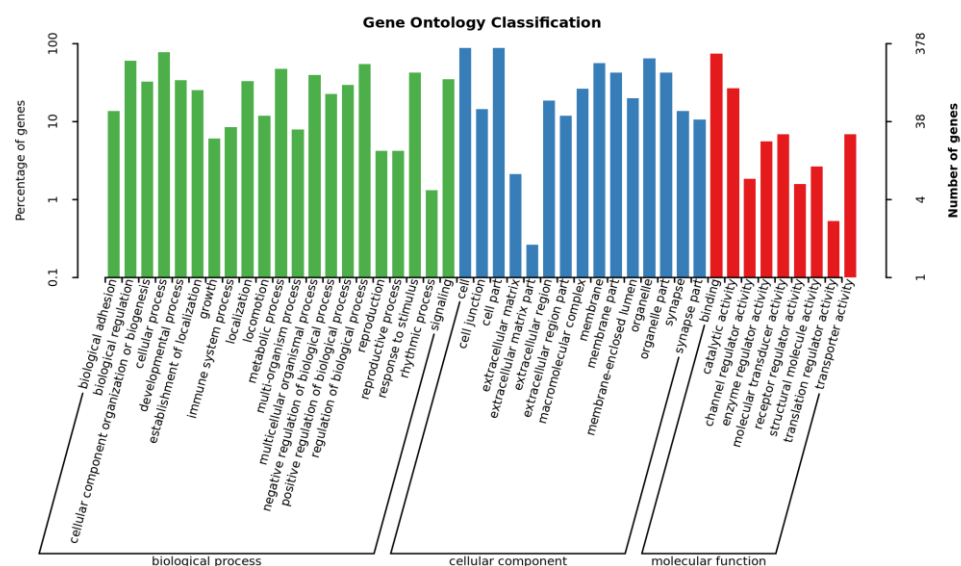


Figure 2. GO terms (gene ontology categories identified in genes with significant SNPs). Note: horizontal coordinates—enriched GO entries; left vertical coordinates—percentage of genes; right vertical coordinates—number of genes.

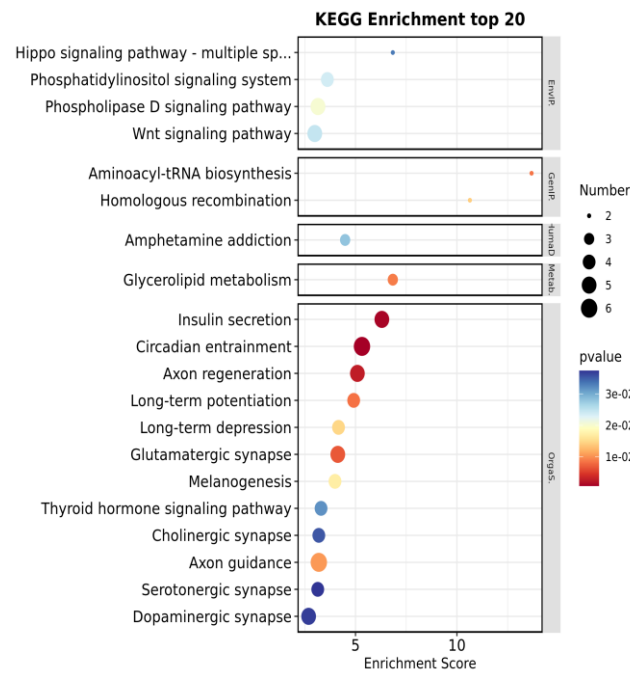


Figure 3. KEGG terms top 20. Note: horizontal scale—enriched scores; vertical scale—enriched entries.

3.4. Protein Interaction Network Analysis

The GO-enriched protein interaction network (Figure 4) showed that a total of 62 node proteins interacted with biological adhesion, immune system processes, cell junctions, and receptor regulator activity related to bacterial infections. Candidate proteins CTNNB1, CDH8, APOD, HCLS1, Tet2, MTSS1 were screened. The KEGG protein interaction network (Figure 5) showed multiple species signaling pathways are associated with bacterial infectious diseases: circadian entrainment, insulin secretion, the Wnt signaling pathway, the thyroid hormone signaling pathway, and the Hippo signaling pathway. There were a total of 9 node proteins in the protein interaction network, the candidate gene YAP1 was screened out.

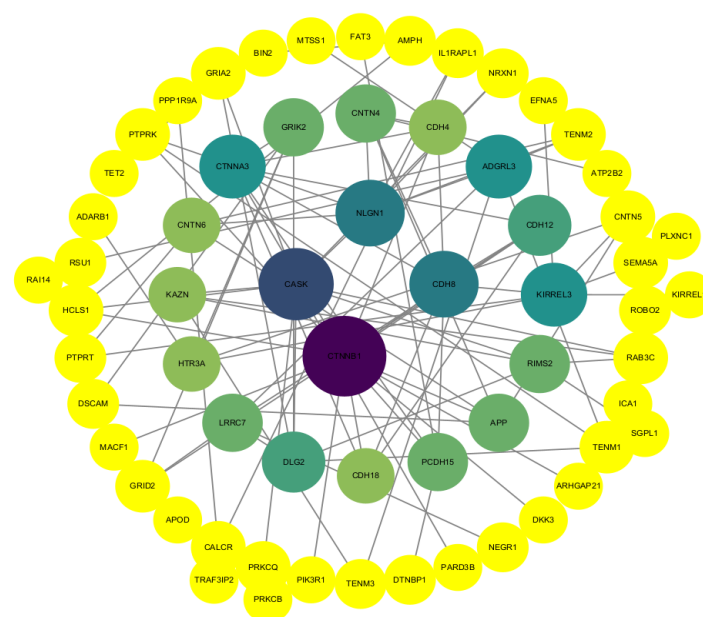


Figure 4. GO-enriched protein interaction network. Nodal protein weights are expressed in colour shades and shape sizes, with darker colours and larger bodies indicating higher weights.

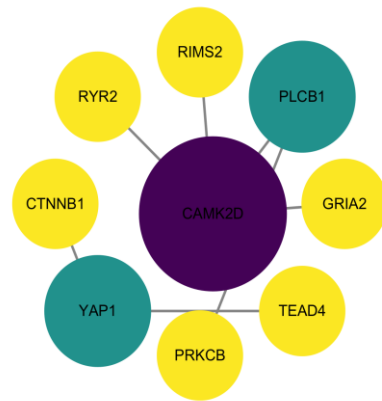


Figure 5. KEGG-enriched protein interaction network. Nodal protein weights are expressed in colour shades and shape sizes, with darker colours and larger bodies indicating higher weights.

4. Discussion

In this study, an association analysis was performed on 22988506 SNPs distributed in 27 chromosomes. There were 466 genes with significant SNPs. GO and KEGG enrichment analysis of genes with differential SNPs was performed. The biological processes associated with bacterial infections in GO enrichment analysis included biological adhesion, immune system processes, cell junctions, and receptor regulator activity. Candidate proteins CTNNB1, CDH8, APOD, HCLS1, Tet2, MTSS1 were screened through the protein interaction network. In the KEGG enrichment analysis, five distinct signaling pathways were involved in the regulation of bacterial infectious diseases, namely circadian entrainment, insulin secretion, the Wnt signaling pathway, the thyroid hormone signaling pathway, and the Hippo signaling pathway—multiple species signaling pathways. The candidate gene YAP1 was screened through the protein interaction network. Preliminary investigations showed that SNPs on the CTNNB1, CDH8, APOD, HCLS1, Tet2, MTSS1 and YAP1 genes may be associated with *E. coli* F17 antagonistic and susceptible Hu sheep lambs.

In GO enrichment analysis and protein–protein interaction network analysis, CTNNB1 (catenin beta 1), an important node protein, is a proto-oncogene encoding β -Catenin protein and a key effector molecule downstream of the Wnt signaling pathway. The function of the Wnt/ β -catenin signaling pathway is more about affecting cell proliferation, polarity and tissue homeostasis, which have a variety of phenotypic regulatory functions in signal transduction. The organism must thus maintain normal Wnt/ β -catenin signaling, due to the imbalance caused by different external or internal stimuli which change cell proliferation, cause apoptosis and inflammation, etc. In the process of infection with *Salmonella typhi*, *Salmonella enteritidis*, and *Escherichia coli*, for example, Wnt/ β -catenin is an important target of several virulence factors produced by such bacteria [30]. *Clostridioides difficile* infection is an important pathogenic microorganism that causes diarrhea in elderly people hospitalized for a long time. TcdB is an exotoxin of *Clostridioides difficile*, which induces macrophage inflammation and epithelial damage, and destroys the normal function of lysosomes in macrophages. TcdB inhibits CTNNB1/ β -catenin activity and lysosomal acidification in macrophages; the resulting lysosomal dysfunction affects the inflammatory response and causes tissue damage through autophagy [31]. It has been reported that mutations at the upper site of CTNNB1 are associated with tuberculosis caused by a conjugated divergent bacterium; this study also showed that CTNNB1 can affect host susceptibility to a conjugated divergent bacterium and this can be used as a marker gene to identify tuberculosis [32]. CDH8 is a classic cadherin which has been studied in bacterial infectious diseases. For example, an endemic wild boar population was divided into a *Mycobacterium tuberculosis*-infected group and an uninfected group, and genome-wide association analysis screened CDH8 as a candidate gene for host genetic susceptibility to tuberculosis caused by divergent bacteria [33]. More research on CDH8 has shown that mutations in the CDH8 gene lead to cerebellar hypoplasia and motor dysfunction, also

an essential factor in autism [34–36]. In order to clarify the genetic background of pig susceptibility to ETEC *E. coli* F4ab/ac and establish a haplotype map of candidate regions, experimental materials were taken from European wild boar and Swedish Large White pigs, and the susceptibility of pigs to F4ab/ac was classified by adhesion experiments. The candidate regions were also screened for polymorphisms, including the presence of the APOD gene, which may be related to the susceptibility of *E. coli* F4ab/ac [37]. Brucellosis is a major zoonotic disease worldwide and poses a major threat to human safety. The literature reports that there is a genetic association between PTPRT gene polymorphisms and goats' susceptibility to *Brucella* [38]. Kostmann's disease is a rare autosomal recessive form of severe congenital neutropenia, characterized by the cessation of promyelocyte/myelocyte maturation in the bone marrow caused by homozygous mutations in the X1 gene encoding the mitochondrial protein HCLS1 [39], resulting in severe recurrent bacterial infections in early infancy. The Tet2 gene is involved in the immune response of bone marrow cells and lymphocytes to infections of respiratory epithelial cells by *Pseudomonas aeruginosa*, affecting the regulation of host defence against it [40]. In addition, the TET2 gene also affects the host's resistance to tuberculosis, the specific mechanism being that the binding of Tet2 to NF- κ B induces the demethylation of TNF, where NF- κ B is an important factor affecting the phenotypes of cell proliferation, apoptosis and inflammation. It is also an indispensable signal transduction pathway in the body; it affects the release of downstream TNF and then changes the host's immune response. It can be used as a new host immunization strategy against tuberculosis [41]. *Staphylococcus aureus* EDIN toxin can induce endothelial cell torn macrophylum (TEM) tunneling, thereby destroying the host endothelial barrier and promoting bacterial transmission. In the loss of MIM (MTSS1) in the transfer of I-BAR domain proteins, the toxin can sense the occurrence of TEM, and finally induce TEM, so that MTSS1 can change the integrity of the host endothelial barrier and affect the invasion of toxins [42]. These studies revealed that CTNBNB1, CDH8, APOD, PTPRT, HCLS1, Tet2 and MTSS1 genes can change the host's resistance to bacterial infections. In this study, these genes are also annotated by whole genome sequencing and association analysis in antagonistic and susceptible Hu lambs; the susceptibility of Hu lambs to *E. coli* F17 may be related to these genes.

In the KEGG enrichment analysis, circadian entrainment is associated with innate immunity. The gut microbiota produces circadian rhythms and the rhythmic expression of antimicrobial proteins is driven by the daily rhythm of epithelial attachment by segmented filamentous bacteria (SFB) (a member of the mouse gut microbiota), activating immune circuits involving group 3 innate lymphoid cells. This circuit triggers oscillations of STAT3 expression and activation in epithelial cells, producing rhythmic antimicrobial protein expression that affects resistance to *Salmonella typhimurium* infection [43]. Perception of microbial ligands by macrophages triggers an increase in glucose metabolism, which provides energy to support antimicrobial inflammation and cytokine production and results in insulin secretion which regulates glucose homeostasis [44]. The Wnt signaling pathway can regulate intestinal epithelial cell proliferation, maintain the intestinal mucosal barrier's integrity, induce Paneth cell differentiation, and increase the expression of antimicrobial peptides [45]. In addition, Wnt signaling pathway involvement plays a key role in the tug-of-war between hosts and pathogens [46]. The thyroid hormone signaling pathway is associated with *Neisseria meningitidis*. Thyroxine supplementation in mice infected with *Neisseria meningitidis* enhances bacterial clearance, reduces inflammation, and promotes survival [47]. In a study of mastitis caused by *Staphylococcus aureus* infection, the Hippo signaling pathway was activated in a transcriptome analysis of *Staphylococcus aureus* infection and uninfected MAC-T cells [48]. In the interaction of five gene protein networks that regulate bacterial infectious disease signaling, the transcription factor YAP-1/YAP is activated when *Pseudomonas aeruginosa* infects the intestinal barrier of worms; the deletion of YAP-1 significantly reduces the survival rate of worms and increases the accumulation of bacteria in the host, which deploys pathogen defenses when it recognizes the loss of the intestinal epithelial barrier [49]. Another study also pointed out that YAP affects pathogen

infection in the pulmonary epithelial barrier, and that in order to maintain this barrier, the regeneration of alveolar epithelium plays an essential role in lung damage caused by infectious pathogens. In the lung inflammation of *Streptococcus pneumoniae* infection, mice lacking Yap/Taz in alveolar epithelial type II cells exhibited a prolonged lung inflammatory response and delayed alveolar epithelial regeneration, affecting the reconstruction of the lung epithelial barrier [50]. In KEGG enrichment analysis and protein–protein interaction network analysis, this paper describes the functions of circadian entrainment, insulin secretion, the Wnt signaling pathway, the thyroid hormone signaling pathway, and the Hippo signaling pathway—multiple species signaling pathways. The candidate gene YAP1 was screened out.

Despite the results of whole genome association analysis, via GO and KEGG enrichment analysis, and protein interaction network analysis, we have predicted the genetic loci and candidate genes related to the resistance of Hu sheep to *E. coli* F17. However, there are still some shortcomings that need to be verified by in vivo and in vitro functional experiments to verify the correlation between candidate gene functions and genetic loci. This is therefore our next line of research.

In recent years, as the demand for sheep has increased and the scale of sheep production has expanded, the likelihood of sheep diseases has also increased. The rapid spread of sheep diseases has led to an increase in the culling of sheep and losses to breeding farms. *Escherichia coli* infections are a major threat to young animals and can be fatal in severe cases [51,52]. Therefore, the breeding of high quality disease-resistant sheep is an important issue, and the use of key genes or candidate sites that influence the occurrence of sheep diseases can improve the economic benefits of breeding enterprises and promote the healthy development of the sheep industry. However, there are few such studies in sheep. There are more studies on sheep economic, growth and development traits, such as the genome-wide association analysis to study SNPs related to body size in Hu sheep, and to screen 5 SNPs related to body height and 4 SNPs related to chest circumference at the chromosomal significance level [53]. Genome-wide association analysis and the phenotype of sheep wool and skin wrinkles have identified variants in several candidate genes [54], as has a genome-wide association analysis to study sheep production and reproductive performance [55]. At present, some genes have been verified by experiments that elucidate their genetic mechanisms, but there are still genes to be discovered and studied in the future. In the last few decades, genome-wide association analysis has provided a large body of reference materials for sheep genetics and breeding, which can identify sites of variation, and key genes and pathways associated with sheep phenotypes, to better understand the genetic mechanisms of the complex forms of sheep. Genome-wide association analysis has been widely used in animal breeding. With the innovation of technology, artificial intelligence, machine learning, and the application of sequencing has enabled more accurate data sequencing results and identification of the more meaningful main genes. At the same time, an increase in sample size provides more accurate predictions. In addition, multi-omics combined analysis is also widely used, and the association analysis of the transcriptome, translation, proteome and metabolome of important trait indicators in sheep can explain problems from multiple directions.

5. Conclusions

Whole genome sequencing and whole genome association analysis revealed 466 genes with significant SNPs ($p < 1.0 \times 10^{-3}$). GO and KEGG enrichment analysis, protein–protein interaction network analysis, and preliminary investigations showed that SNPs on CTNBN1, CDH8, APOD, HCLS1, Tet2, MTSS1 and YAP1 genes may be associated with the antagonism and susceptibility of Hu sheep lambs to *E. coli* F17. This study provides genetic loci and candidate genes for resistance of Hu sheep lambs to *E. coli* F17 infection and provides a genetic basis for breeding disease-resistant sheep.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ani14010161/s1>. Figure S1: The density of SNPs distributed across chromosomes; Figure S2: Genome-wide coverage of every sample; Table S1: Variant location information; Table S2: GO level 2 enrichment; Table S3: KEGG enrichment.

Author Contributions: Y.D. participated in the main experiment and wrote the paper, P.S. and Y.G. were involved in executing the study, X.L., X.C., S.W., Z.Y. and W.S. guided the experiment and revised the paper. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China-CGIAR (32061143036); the National Natural Science Foundation of China (31872333 and 32172689); Major New Varieties of Agricultural Projects in Jiangsu Province (PZCZ201739); Jiangsu 333 Distinguished Talents Project Foundation [(2022) 2-323]; the Open Project Program Of the Joint International Research Laboratory of Agriculture and Agri-Product Safety, the Ministry of Education of China, Yangzhou University (JILAR-KF202205); the Natural Science Foundation of the Jiangsu Higher Education Institutions of China (22KJA230001); the Jiangsu Agricultural Science and Technology Innovation Fund (CX (23)1036); the Major Project of Natural Science Foundation of Xinjiang Uyghur Autonomous Region (2022D01D47); the Distinguished Talents Project Foundation of Yangzhou University; and the High-end Foreign Expert Introduction Project (G2022014148L).

Institutional Review Board Statement: The experiment was approved by the animal welfare committee of Yangzhou University (license No.: syxk (Su) 2012-0029) and conducted according to the guidelines of the Animal Use Committee of the Ministry of Agriculture of China (Beijing, China). All efforts were made to minimize the suffering of animals.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are openly available in the SRA database, accession number PRJNA996327.

Acknowledgments: Thank you to teachers for their support and classmates for their help.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Zhao, X.; Zhao, H.; Zhou, Z.; Miao, Y.; Li, R.; Yang, B.; Cao, C.; Xiao, S.; Wang, X.; Liu, H.; et al. Characterisation of Extended-Spectrum β -Lactamase-Producing *Escherichia coli* Isolates That Cause Diarrhea in Sheep in Northwest China. *Microbiol. Spectr.* **2022**, *10*, e0159522. [CrossRef] [PubMed]
2. Kaper, J.B.; Nataro, J.P.; Mobley, H.L. Pathogenic *Escherichia coli*. *Nat. Rev. Microbiol.* **2004**, *2*, 123–140. [CrossRef] [PubMed]
3. Le Bougu  nec, C.; Bertin, Y. AFA and F17 adhesins produced by pathogenic *Escherichia coli* strains in domestic animals. *Vet. Res.* **1999**, *30*, 317–342. [PubMed]
4. Lintermans, P.F.; Pohl, P.; Bertels, A.; Charlier, G.; Vandekerckhove, J.; Van Damme, J.; Schoup, J.; Schlicker, C.; Korhonen, T.; De Greve, H.; et al. Characterization and purification of the F17 adhesin on the surface of bovine enteropathogenic and septicemic *Escherichia coli*. *Am. J. Vet. Res.* **1988**, *49*, 1794–1799.
5. El Mazouari, K.; Oswald, E.; Hernalsteens, J.P.; Lintermans, P.; De Greve, H. F17-like fimbriae from an invasive *Escherichia coli* strain producing cytotoxic necrotizing factor type 2 toxin. *Infect. Immun.* **1994**, *62*, 2633–2638. [CrossRef]
6. Bertin, Y.; Girardeau, J.P.; Darfeuille-Michaud, A.; Contrepois, M. Characterization of 20K fimbria, a new adhesin of septicemic and diarrhea-associated *Escherichia coli* strains, that belongs to a family of adhesins with N-acetyl-D-glucosamine recognition. *Infect. Immun.* **1996**, *64*, 332–342. [CrossRef]
7. Bertin, Y.; Martin, C.; Oswald, E.; Girardeau, J.P. Rapid and specific detection of F17-related pilin and adhesin genes in diarrheic and septicemic *Escherichia coli* strains by multiplex PCR. *J. Clin. Microbiol.* **1996**, *34*, 2921–2928. [CrossRef]
8. Bihannic, M.; Ghanbarpour, R.; Auvray, F.; Cavali  , L.; Ch  tre, P.; Boury, M.; Brug  re, H.; Madec, J.Y.; Oswald, E. Identification and detection of three new F17 fimbrial variants in *Escherichia coli* strains isolated from cattle. *Vet. Res.* **2014**, *45*, 76. [CrossRef]
9. Cid, D.; Sanz, R.; Mar  n, I.; de Greve, H.; Ruiz-Santa-Quiteria, J.A.; Amils, R.; de la Fuente, R. Characterization of nonenterotoxigenic *Escherichia coli* strains producing F17 fimbriae isolated from diarrheic lambs and goat kids. *J. Clin. Microbiol.* **1999**, *37*, 1370–1375. [CrossRef]
10. Osman, K.; Mustafa, A.; Elhariri, M.; Abdelhamed, G.J.T.; diseases, e. The distribution of *Escherichia coli* serovars, virulence genes, gene association and combinations and virulence genes encoding serotypes in pathogenic *E. coli* recovered from diarrhoeic calves, sheep and goat. *Transbound. Emerg. Dis.* **2013**, *60*, 69–78. [CrossRef]
11. Zhang, Y.; Tan, P.; Zhao, Y.; Ma, X. Enterotoxigenic *Escherichia coli*: Intestinal pathogenesis mechanisms and colonization resistance by gut microbiota. *Gut Microbes* **2022**, *14*, 2055943. [CrossRef] [PubMed]
12. Kagnoff, M.F.; Eckmann, L. Epithelial cells as sensors for microbial infection. *J. Clin. Investig.* **1997**, *100*, 6–10. [CrossRef] [PubMed]

13. Li, M.; Wang, Y.; Sun, Y.; Cui, H.; Zhu, S.J.; Qiu, H.J. Mucosal vaccines: Strategies and challenges. *Immunol. Lett.* **2020**, *217*, 116–125. [[CrossRef](#)] [[PubMed](#)]
14. Liu, G.; Gu, K.; Wang, F.; Jia, G.; Zhao, H.; Chen, X.; Wu, C.; Zhang, R.; Tian, G.; Cai, J.; et al. Tryptophan Ameliorates Barrier Integrity and Alleviates the Inflammatory Response to Enterotoxigenic *Escherichia coli* K88 Through the CaSR/Rac1/PLC- γ 1 Signaling Pathway in Porcine Intestinal Epithelial Cells. *Front. Immunol.* **2021**, *12*, 748497. [[CrossRef](#)] [[PubMed](#)]
15. Chen, Q.; Wang, P.; Wang, J.; Xu, J.; Liu, C.; Qiao, H.; Gan, L.; Duan, E.; Zhang, Y.; Wang, M.; et al. Zinc Laurate Protects against Intestinal Barrier Dysfunction and Inflammation Induced by ETEC in a Mice Model. *Nutrients* **2022**, *15*, 54. [[CrossRef](#)]
16. Peng, H.; Shen, Y.; Zhang, Q.; Liu, J.; Wang, Z.; Huang, L.; Zhou, F.; Yu, J.; Liu, M.; Yuan, Y.; et al. Qihuang decoction promotes the recovery of intestinal immune barrier dysfunction after gastrectomy in rats. *Am. J. Transl. Res.* **2018**, *10*, 827–836.
17. Hao, W.; Hao, C.; Wu, C.; Xu, Y.; Jin, C. Aluminum induced intestinal dysfunction via mechanical, immune, chemical and biological barriers. *Chemosphere* **2022**, *288*, 132556. [[CrossRef](#)]
18. Chen, W.; Lv, X.; Zhang, W.; Hu, T.; Cao, X.; Ren, Z.; Getachew, T.; Mwacharo, J.M.; Haile, A.; Sun, W. Insights Into Long Non-Coding RNA and mRNA Expression in the Jejunum of Lambs Challenged With *Escherichia coli* F17. *Front. Vet. Sci.* **2022**, *9*, 819917. [[CrossRef](#)]
19. Chen, S.; Zhou, Y.; Chen, Y.; Gu, J. fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **2018**, *34*, i884–i890. [[CrossRef](#)]
20. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **2009**, *25*, 1754–1760. [[CrossRef](#)]
21. 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](#)] [[PubMed](#)]
22. McKenna, A.; Hanna, M.; Banks, E.; Sivachenko, A.; Cibulskis, K.; Kernysky, A.; Garimella, K.; Altshuler, D.; Gabriel, S.; Daly, M.; et al. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **2010**, *20*, 1297–1303. [[CrossRef](#)] [[PubMed](#)]
23. Wang, K.; Li, M.; Hakonarson, H. ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **2010**, *38*, e164. [[CrossRef](#)] [[PubMed](#)]
24. Sahana, G.; Guldbrandtsen, B.; Bendixen, C.; Lund, M.S. Genome-wide association mapping for female fertility traits in Danish and Swedish Holstein cattle. *Anim. Genet.* **2010**, *41*, 579–588. [[CrossRef](#)] [[PubMed](#)]
25. Turner, S.D.J.B. qqman: An R package for visualizing GWAS results using Q-Q and manhattan plots. *bioRxiv* **2014**. bioRxiv:005165.
26. Gene Ontology Consortium. The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res.* **2019**, *47*, D330–D338. [[CrossRef](#)]
27. Kanehisa, M.; Araki, M.; Goto, S.; Hattori, M.; Hirakawa, M.; Itoh, M.; Katayama, T.; Kawashima, S.; Okuda, S.; Tokimatsu, T.; et al. KEGG for linking genomes to life and the environment. *Nucleic Acids Res.* **2008**, *36*, D480–D484. [[CrossRef](#)]
28. Szklarczyk, D.; Franceschini, A.; Wyder, S.; Forslund, K.; Heller, D.; Huerta-Cepas, J.; Simonovic, M.; Roth, A.; Santos, A.; Tsafou, K.P.; et al. STRING v10: Protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* **2015**, *43*, D447–D452. [[CrossRef](#)]
29. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **2003**, *13*, 2498–2504. [[CrossRef](#)]
30. Silva-García, O.; Valdez-Alarcón, J.J.; Baizabal-Aguirre, V.M. Wnt/ β -Catenin Signaling as a Molecular Target by Pathogenic Bacteria. *Front. Immunol.* **2019**, *10*, 2135. [[CrossRef](#)]
31. Chan, H.; Li, Q.; Wang, X.; Liu, W.Y.; Hu, W.; Zeng, J.; Xie, C.; Kwong, T.N.Y.; Ho, I.H.T.; Liu, X.; et al. Vitamin D(3) and carbamazepine protect against *Clostridioides difficile* infection in mice by restoring macrophage lysosome acidification. *Autophagy* **2022**, *18*, 2050–2067. [[CrossRef](#)] [[PubMed](#)]
32. Hu, X.; Zhou, J.; Chen, X.; Zhou, Y.; Song, X.; Cai, B.; Zhang, J.; Lu, X.; Ying, B. Pathway Analyses Identify Novel Variants in the WNT Signaling Pathway Associated with Tuberculosis in Chinese Population. *Sci. Rep.* **2016**, *6*, 28530. [[CrossRef](#)] [[PubMed](#)]
33. Queirós, J.; Alves, P.C.; Vicente, J.; Gortázar, C.; de la Fuente, J. Genome-wide associations identify novel candidate loci associated with genetic susceptibility to tuberculosis in wild boar. *Sci. Rep.* **2018**, *8*, 1980. [[CrossRef](#)] [[PubMed](#)]
34. Kawamura, A.; Katayama, Y.; Kakegawa, W.; Ino, D.; Nishiyama, M.; Yuzaki, M.; Nakayama, K.I. The autism-associated protein CHD8 is required for cerebellar development and motor function. *Cell Rep.* **2021**, *35*, 108932. [[CrossRef](#)] [[PubMed](#)]
35. Pagnamenta, A.T.; Khan, H.; Walker, S.; Gerrelli, D.; Wing, K.; Bonaglia, M.C.; Giorda, R.; Berney, T.; Mani, E.; Molteni, M.; et al. Rare familial 16q21 microdeletions under a linkage peak implicate cadherin 8 (CDH8) in susceptibility to autism and learning disability. *J. Med. Genet.* **2011**, *48*, 48–54. [[CrossRef](#)]
36. Frei, J.A.; Niescier, R.F.; Bridi, M.S.; Durens, M.; Nestor, J.E.; Kilander, M.B.C.; Yuan, X.; Dykxhoorn, D.M.; Nestor, M.W.; Huang, S.; et al. Regulation of Neural Circuit Development by Cadherin-11 Provides Implications for Autism. *eNeuro* **2021**, *8*, 4. [[CrossRef](#)] [[PubMed](#)]
37. Jacobsen, M.; Kracht, S.S.; Estes, G.; Cirera, S.; Edfors, I.; Archibald, A.L.; Bendixen, C.; Andersson, L.; Fredholm, M.; Jørgensen, C.B. Refined candidate region specified by haplotype sharing for *Escherichia coli* F4ab/F4ac susceptibility alleles in pigs. *Anim. Genet.* **2010**, *41*, 21–25. [[CrossRef](#)]
38. Rossi, U.A.; Hasenauer, F.C.; Caffaro, M.E.; Neumann, R.; Salatin, A.; Poli, M.A.; Rossetti, C.A. A haplotype at intron 8 of PTPRT gene is associated with resistance to *Brucella* infection in Argentinian creole goats. *Vet. Microbiol.* **2017**, *207*, 133–137. [[CrossRef](#)]

39. Aytekin, C.; Germeshausen, M.; Tuygun, N.; Tanir, G.; Dogu, F.; Ikinciogullari, A. Kostmann disease with developmental delay in three patients. *Eur. J. Pediatr.* **2010**, *169*, 759–762. [\[CrossRef\]](#)
40. Qin, W.; Brands, X.; Van't Veer, C.; de Vos, A.F.; Scicluna, B.P.; van der Poll, T. Bronchial Epithelial Tet2 Maintains Epithelial Integrity during Acute Pseudomonas aeruginosa Pneumonia. *Infect. Immun.* **2020**, *89*, 10–1128. [\[CrossRef\]](#)
41. Zhu, C.; Cai, Y.; Mo, S.; Zhu, J.; Wang, W.; Peng, B.; Guo, J.; Zhang, Z.; Chen, X. NF- κ B-mediated TET2-dependent TNF promoter demethylation drives Mtb-upregulation TNF expression in macrophages. *Tuberculosis* **2021**, *129*, 102108. [\[CrossRef\]](#) [\[PubMed\]](#)
42. Maddugoda, M.P.; Stefani, C.; Gonzalez-Rodriguez, D.; Saarikangas, J.; Torrino, S.; Janel, S.; Munro, P.; Doye, A.; Prodon, F.; Aurrand-Lions, M.; et al. cAMP signaling by anthrax edema toxin induces transendothelial cell tunnels, which are resealed by MIM via Arp2/3-driven actin polymerization. *Cell Host Microbe* **2011**, *10*, 464–474. [\[CrossRef\]](#) [\[PubMed\]](#)
43. Brooks, J.F., 2nd; Behrendt, C.L.; Ruhn, K.A.; Lee, S.; Raj, P.; Takahashi, J.S.; Hooper, L.V. The microbiota coordinates diurnal rhythms in innate immunity with the circadian clock. *Cell* **2021**, *184*, 4154–4167.e12. [\[CrossRef\]](#) [\[PubMed\]](#)
44. Zhang, C.; Wang, K.; Hu, Z.; Yang, L.; Wei, B.; Li, S.; Qin, X.; Yang, P.; Yu, H. SIRT5 is important for bacterial infection by regulating insulin secretion and glucose homeostasis. *Protein Cell* **2020**, *11*, 846–851. [\[CrossRef\]](#) [\[PubMed\]](#)
45. Wu, H.; Xie, S.; Miao, J.; Li, Y.; Wang, Z.; Wang, M.; Yu, Q. Lactobacillus reuteri maintains intestinal epithelial regeneration and repairs damaged intestinal mucosa. *Gut Microbes* **2020**, *11*, 997–1014. [\[CrossRef\]](#)
46. Rogan, M.R.; Patterson, L.L.; Wang, J.Y.; McBride, J.W. Bacterial Manipulation of Wnt Signaling: A Host-Pathogen Tug-of-Wnt. *Front. Immunol.* **2019**, *10*, 2390. [\[CrossRef\]](#) [\[PubMed\]](#)
47. Chen, Y.; Sjölander, M.; Wang, X.; Altenbacher, G.; Hagner, M.; Berglund, P.; Gao, Y.; Lu, T.; Jonsson, A.B.; Sjölander, H. Thyroid hormone enhances nitric oxide-mediated bacterial clearance and promotes survival after meningococcal infection. *PLoS ONE* **2012**, *7*, e41445. [\[CrossRef\]](#)
48. Chen, Y.; Jing, H.; Chen, M.; Liang, W.; Yang, J.; Deng, G.; Guo, M. Transcriptional Profiling of Exosomes Derived from Staphylococcus aureus-Infected Bovine Mammary Epithelial Cell Line MAC-T by RNA-Seq Analysis. *Oxidative Med. Cell. Longev.* **2021**, *2021*, 8460355. [\[CrossRef\]](#)
49. Ma, Y.C.; Yang, Z.S.; Ma, L.Q.; Shu, R.; Zou, C.G.; Zhang, K.Q. YAP in epithelium senses gut barrier loss to deploy defenses against pathogens. *PLoS Pathog.* **2020**, *16*, e1008766. [\[CrossRef\]](#)
50. LaCanna, R.; Liccardo, D.; Zhang, P.; Tragesser, L.; Wang, Y.; Cao, T.; Chapman, H.A.; Morrissey, E.E.; Shen, H.; Koch, W.J.; et al. Yap/Taz regulate alveolar regeneration and resolution of lung inflammation. *J. Clin. Investig.* **2019**, *129*, 2107–2122. [\[CrossRef\]](#)
51. Salhi, I.; Rabti, A.; Dhehibi, A.; Raouafi, N. Sandwich-Based Immunosensor for Dual-Mode Detection of Pathogenic F17-Positive Escherichia coli Strains. *Int. J. Mol. Sci.* **2022**, *23*, 6028. [\[CrossRef\]](#) [\[PubMed\]](#)
52. Chen, W.; Lv, X.; Zhang, W.; Hu, T.; Cao, X.; Ren, Z.; Getachew, T.; Mwacharo, J.M.; Haile, A.; Sun, W. Non-Coding Transcriptome Provides Novel Insights into the Escherichia coli F17 Susceptibility of Sheep Lamb. *Biology* **2022**, *11*, 348. [\[CrossRef\]](#) [\[PubMed\]](#)
53. Jiang, J.; Cao, Y.; Shan, H.; Wu, J.; Song, X.; Jiang, Y. The GWAS Analysis of Body Size and Population Verification of Related SNPs in Hu Sheep. *Front. Genet.* **2021**, *12*, 642552. [\[CrossRef\]](#) [\[PubMed\]](#)
54. Bolormaa, S.; Swan, A.A.; Stothard, P.; Khansefid, M.; Moghaddar, N.; Duijvesteijn, N.; van der Werf, J.H.J.; Daetwyler, H.D.; MacLeod, I.M. A conditional multi-trait sequence GWAS discovers pleiotropic candidate genes and variants for sheep wool, skin wrinkle and breech cover traits. *Genet. Sel. Evol.* **2021**, *53*, 58. [\[CrossRef\]](#)
55. Gebreselassie, G.; Berihulay, H.; Jiang, L.; Ma, Y. Review on Genomic Regions and Candidate Genes Associated with Economically Important Production and Reproduction Traits in Sheep (Ovis aries). *Animals* **2019**, *10*, 33. [\[CrossRef\]](#)

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.