

## Article

# Integrated Analysis of Long Non-Coding RNA and mRNA Expression Profiles in Testes of Calves and Sexually Mature Wandong Bulls (*Bos taurus*)

Hongyu Liu <sup>1,2,†</sup>, Ibrar Muhammad Khan <sup>1,2,†</sup> , Huiqun Yin <sup>3</sup>, Xinqi Zhou <sup>1</sup>, Muhammad Rizwan <sup>1</sup>, Jingyi Zhuang <sup>1</sup> and Yunhai Zhang <sup>1,2,\*</sup> 

<sup>1</sup> College of Animal Science and Technology, Anhui Agricultural University, Hefei 230036, China; liuhongyu@ahau.edu.cn (H.L.); ibrar.pesh@gmail.com (I.M.K.); xinqizhou2021@163.com (X.Z.); rizwiktk007@hotmail.com (M.R.); Zhuang\_JY320@163.com (J.Z.)

<sup>2</sup> Anhui Provincial Laboratory of Local Livestock and Poultry Genetic Resource Conservation and Breeding, Anhui Agricultural University, Hefei 230036, China

<sup>3</sup> Reproductive Medicine Center, The 901st Hospital, Hefei 230031, China; milklecherry@163.com

\* Correspondence: yunhaizhang@ahau.edu.cn; Tel.: +86-551-6578-6357

† These authors contributed equally to this paper.

**Simple Summary:** The Chinese native cattle breeds are economically important, and these breeds are used extensively for beef production. However, their reproductive efficiency and the local elite genetic breeding populations, which is an important consideration in livestock production systems, have not been fully established yet. The literature recommends the application of molecular strategies that improve the production of good-quality sperm and increase the availability of elite genetics for improving local breed production. Therefore, this study comprehensively analyzed the mRNAs and long non-coding RNAs expression profiling, based on total RNA-Seq, of the testis samples of bulls of two distinct age groups. The findings of this study have strong implications, as they provide a suitable framework for understanding the mechanism of mRNAs and long non-coding RNAs in the development of testes and spermatogenesis.



**Citation:** Liu, H.; Khan, I.M.; Yin, H.; Zhou, X.; Rizwan, M.; Zhuang, J.; Zhang, Y. Integrated Analysis of Long Non-Coding RNA and mRNA Expression Profiles in Testes of Calves and Sexually Mature Wandong Bulls (*Bos taurus*). *Animals* **2021**, *11*, 2006. <https://doi.org/10.3390/ani11072006>

Academic Editors: Izhar Hyder Qazi and Guangbin Zhou

Received: 17 June 2021

Accepted: 30 June 2021

Published: 5 July 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 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/>).

**Abstract:** The mRNAs and long non-coding RNAs axes are playing a vital role in the regulating of post-transcriptional gene expression. Thereby, elucidating the expression pattern of mRNAs and long non-coding RNAs underlying testis development is crucial. In this study, mRNA and long non-coding RNAs expression profiles were investigated in 3-month-old calves and 3-year-old mature bulls' testes by total RNA sequencing. Additionally, during the gene level analysis, 21,250 mRNAs and 20,533 long non-coding RNAs were identified. As a result, 7908 long non-coding RNAs ( $p$ -adjust < 0.05) and 5122 mRNAs ( $p$ -adjust < 0.05) were significantly differentially expressed between the distinct age groups. In addition, gene ontology and biological pathway analyses revealed that the predicted target genes are enriched in the lysine degradation, cell cycle, propanoate metabolism, adherens junction and cell adhesion molecules pathways. Correspondingly, the RT-qPCR validation results showed a strong consistency with the sequencing data. The source genes for the mRNAs (*CCDC83*, *DMRTC2*, *HSPA2*, *IQCG*, *PACRG*, *SPO11*, *EHHADH*, *SPP1*, *NSD2* and *ACTN4*) and the long non-coding RNAs (*COX7A2*, *COX6B2*, *TRIM37*, *PRM2*, *INHBA*, *ERBB4*, *SDHA*, *ATP6VOA2*, *FGF9* and *TCF21*) were found to be actively associated with bull sexual maturity and spermatogenesis. This study provided a comprehensive catalog of long non-coding RNAs in the bovine testes and also offered useful resources for understanding the differences in sexual development caused by the changes in the mRNA and long non-coding RNA interaction expressions between the immature and mature stages.

**Keywords:** testicular growth; spermatogenesis; mRNAs and lncRNAs axes; total RNA-sequencing

## 1. Introduction

Improvement in the reproductive efficiency of local cattle bull via genetic selection is crucial, as individual bulls can breed multiple numbers of cows [1]. Advances in livestock genomics,

such as transcriptomics, proteomics and single-nucleotide polymorphism genotypes studies, are predicted the direct increasing rate of genetic gain in a generation [2,3]. Thus, it demands an extensive use of molecular tools to improve the reproductive traits of Chinese native cattle.

Mammalian spermatogenesis takes place by the subsequent division of the cell: self-restoration of gonocytes, mitotic proliferation of spermatogonia, meiotic division of primary spermatocytes and post-meiotic differentiation of secondary spermatocytes. Finally, the haploid spermatids transform into another process called spermiogenesis [4]. The process of spermatogenesis is more curious and involves several transcriptional factors that regulate by the testis-specific genes at the transcriptional and post transcriptional levels [5,6]. The interactive network of lncRNAs and mRNAs is thought to be involved in mammalian reproduction system regulation and sperm quality parameters [7].

The lncRNA profiles are comprehensively analyzed and identified at the 3-day-old neonatal and 13-month-old mature stages of Angus cattle [8]. The detailed catalog of known and novel lncRNAs was also identified at four different developmental stages, including 60, 90, 120 and 150 days in boar testes [9]. Non-coding RNAs (ncRNAs) with biological functions were identified nearly 60 years ago, and other functional ncRNAs were considered to occur as early as the 1980s [10]. Later, ncRNAs such as Xist were discovered to control the chromosome structure. Since then, the number of novel and potentially usable ncRNAs has grown dramatically [11,12]. However, new literature suggests that ncRNAs such as microRNAs [12], PIWI-interacting RNAs, small interfering RNAs and long non-coding RNAs (lncRNAs) play a wide range of functions [13–16]. These massive quantities of transcript species are referred to as ncRNAs.

In several animal species, the identification and functional validation of lncRNAs have been performed and characterized in spermatogenesis and testicular development [17,18]. Many lncRNAs have been found in mouse (e.g., 8265 lncRNAs and 18,563 mRNAs), *Drosophila* (128 testis-specific lncRNAs), pig (15,528 lncRNAs), chicken (2597 lncRNAs and 17,690 mRNAs) and in sheep (6460 lncRNAs and 42,300 mRNAs) at various stages of testis development and spermatogenesis [19–23].

Recently, lncRNAs have been recognized in various developmental stages of testes and spermatogenesis in rat, mouse and human models. The lncRNA link genes Membrane protein related to HSP30 (*MRH1*), testis-specific X-linked gene (*TSX*), DNA methylation regions (*DMR*) and *HONGRES2* are predicted to have critical roles in testicular growth and spermatogenesis [24,25]. The *DMRT1* protein has been shown with a role in spermatogonial growth and avoiding premature meiosis in spermatogonia, implying that the lncRNA *Dmr* could be involved in the transition between mitosis and meiosis during spermatozoa development [26]. The testis-specific X-linked lncRNAs were found to be expressed exclusively in pachytene spermatocytes but not in the spermatogonia or round spermatids, implying a regulatory role in germline meiotic division [23].

We investigated the supportive network of lncRNA-mRNAs and their expression profiles in immature and sexually mature bull testes. Our research identified the lncRNAs-mRNA integrative network during the sexual maturation of bulls (*Bos taurus*) and provided a useful landscape resource for reviewing the useful roles of ncRNA in testis development and spermatogenesis. We performed transcriptome sequencing and examined the lncRNA-mRNA expression design in the testis. The obtained results will provide potential molecular fingerprints for assessing the developmental status of bull testes and, also, facilitate the breeding soundness examination in local elite bulls for the AI program.

## 2. Materials and Methods

### 2.1. Experimental Animals and Sample Collection

Wandong cattle are an autochthonous Chinese breed used extensively for beef production. They tolerate extreme weather conditions and raw feed and are resistant to tick-borne diseases. Two different age group ( $n = 3$ ) of physically healthy and sexually mature Wandong bulls ( $3 \pm 0.014$  years) and immature calves ( $3 \pm 0.24$  months) were selected for the current research [27]. The sexually mature group was subjected for seminal quality traits, and the

semen volume, concentration, grass motility, progressive motility and spermatozoa defects were measured by specialized technicians (Table S1). Andrology testing was also carried out, where the selected bulls were kept in another pen to mount other bulls and become aroused. The bulls and calves were slaughtered for testicle samples collection at Fengyang County, Anhui, China. A total of six samples of testes from immature calves and mature Wandong bulls (*Bos taurus*) were collected (Figure S1). All pairs of testes were processed by incising the scrotum medially and uncovering the right and left testicles within the tunica vaginalis. The samples were immediately put in a cryogenic vial and frozen in liquid nitrogen ( $-195.8\text{ }^{\circ}\text{C}$ ) until use. The collection of experimental animals was carried out under strict ethical procedures and approved (SYXK 2016-007) by the Animal Care Unit and College of Animal Science and Technology, Anhui Agricultural University. All possible efforts were made to minimize any pain experienced by the experimental animals.

### 2.2. Histological Study of Testicular Tissues

The bulls and calves testes tissue samples were washed with 0.9% and preserved with Bouin's solution containing 75-mL saturated bitter acid solution, 25-mL formaldehyde and 5-mL glacial acetic acid for 48 h at room temperature [28]. Next, the tissues were embedded in paraffin and dissected at parts of 6- $\mu\text{m}$ -thick sections and stained with hematoxylin-eosin (HE). The histomorphology of the testicular tissues was analyzed using 100 $\times$ , 200 $\times$  and 400 $\times$  magnifying light microscopies by Nikon Eclipse C1; Nikon, Tokyo, Japan.

### 2.3. RNA Isolation and Integrity Assessment

The testicular samples were sent to a commercial sequencing facility (Beijing Novogene Corporation, Beijing, China) for Seq-analysis. Total RNA was isolated from the testis samples using Trizol reagent (Takara, Beijing, China) in sterile conditions and treated with DNase I enzyme to remove endogenous DNA contamination according to the manufacturer's protocols. RNA quality and contamination were assessed by 1% agarose gel electrophoresis. The purity was checked using a NanoPhotometer<sup>®</sup> spectrophotometer (IMPLEN, Westlake Village, CA, USA). The total RNA concentration was assessed using the Qubit<sup>®</sup> RNA Assay Kit in Qubit<sup>®</sup> 2.0 Fluorometer (Life Technologies, Carlsbad, CA, USA). RNA integrity was measured using the RNA Nano 6000 Assay Kit of the Bioanalyzer 2100 system (Agilent Technologies, Santa Clara, CA, USA).

### 2.4. cDNA Library Construction for lncRNA-mRNAs Sequencing Analysis

A total of 3 $\mu\text{g}$  of RNA from each testis sample of 3-month-old calves and 3-year-old bulls were used in construction of the cDNA library. In the present study, six (calves:  $n = 3$  and bulls:  $n = 3$ ) seq-libraries were generated by using the NEB-Next<sup>®</sup> Ultra-TM RNA Library Prep Kit for Illumina<sup>®</sup> (NEB, San Diego, CA, USA). The Novogene Bioinformatics Institute constructed a chain-specific library by removing the ribosomal RNA [28]. First of all, the ribosomal RNA was removed from the total RNA, and then, the RNA was divided into many short slices of 250–300 bp. The first cDNA as a template strand was manufactured from fragmented RNA and the random use of oligonucleotides as primers. In the next step, the RNA strand was degraded with RNase H, and then, the second cDNA strand was manufactured along with dNTPs (dATP, dGTP, dUTP and dCTP) as a material underneath the DNA polymerase I system. After the purification, double-stranded cDNA was fixed while followed by a tail step and connected to the sequencing connector. A cDNA around 200 bp was separated with AM-Pure XP beads. The USER enzyme degraded the second strand of the cDNA, which contained U, and PCR amplification was performed to obtain the library. Qubit 2.0 (<https://www.thermofisher.com>, accessed on 11 April 2019) was used to quantify the library, which was then diluted to 1.5 ng/ $\mu\text{L}$ , whereas the total insert size of the library was detected by Agilent 2100. When the library qualified the inspection process, the Illumina PE150 (Hi-Seq 4000 System) (<https://www.iefitd.com/biotech>, accessed on 11 April 2019) sequencing machine was run to identify the data output demand, and 150-base pair (bp) paired-end reads were generated. The cDNA library construction and quality inspection are shown in Figure S2.

### 2.5. Transcriptome Assembly

After isolation of the reads with adaptors and a poly-N greater than 10% and poor-quality reads from raw data using in-house perl scripts produced by the Bioinformatics Institute (Beijing, China), clean transcriptomic data were obtained. The ribosomal RNAs free reads of each calf and bull sample were charted to reference genome *Bos taurus*-UMD 3.1.1 by TopHat2 (v. 2.0.3.12), as described in reference [29,30]. Bowtie v2.0.6 was used to build an index of the bovine reference genome [31], and the paired-end clean reads were aligned using TopHat [32] v2.0.9 (<https://ccb.jhu.edu/software/tophat/index.shtml>, accessed on 11 November 2019). The mapped reads of each sample were assembled by both Scripture (beta2) [33] and Cufflinks v2.1.1 [34].

### 2.6. LncRNAs and mRNAs Identification and Coding Potential

LncRNA has a variety of classification methods, and Novogene references HGNC to classify the newly screened lncRNAs into the following four types, according to their positional relationship with known mRNAs. In order to cut down the false discovery rate (FDR), we adopted these subsequent steps to classify the diverse classes of lncRNAs, including lincRNA, intronic lncRNA, antisense lncRNA and sense-overlapping [35,36]. All the assembled transcripts of the six sequencing libraries were associated when using cuffcompare software [37], and those transcripts that were assembled by Scripture (beta2) [33] were discarded. The transcripts possessing only single exon and less than 200 bp were also discarded, and the reads coverage of every transcript was designed using Cufflinks v2.1.1 (<http://cole-trapnell-lab.github.io/cufflinks/releases/v2.1.1/>, accessed on 12 February 2019) [33]. All those transcripts that had less than three-reads coverage and an expression value (FPKM) less than 2 (log<sub>2</sub>-fold change) were removed, and the remaining known transcripts were blasted with bovine known lncRNAs in ALDB [38] by cuffcompare software. Coding potential of lncRNAs was determined by coding non coding Index (CNCI) software (<https://github.com/www-bioinfo-org/>, accessed on 5 June 2020) [39], Pfam-scansoftware (<https://www.ebi.ac.uk/pfamscan/>, accessed on 5 June 2020) [40,41] and coded potential calculator (CPC) software (<http://cpc2.gao-lab.org>, accessed on 5 June 2020) [42]. As a result of this software, all transcripts with coding potential were categorized as putative lncRNAs, whereas those lacking coding potentials were classified as non-coding RNAs. After filtering out, the database of known transcripts was categorized into a novel class of lncRNAs and mRNAs, and the different steps are shown in Figure S3.

### 2.7. Expression Pattern of Differentially Expressed lncRNAs and mRNAs

The lncRNA and mRNA gene expression levels were assessed using the FPKMs (fragments per kilobase of transcript sequence per millions of base pairs sequenced) value. A significant analysis of the expression differences (ED) at the gene or transcript level was done, and the functional genes or transcripts relevant to the testis developmental process were found. The Novogene Bioinformatics Institute (Beijing, China) used Cuffdiff (v2.1.1) (<https://bioinformaticscenter.com/tools/rna-seq/descriptions/Cufflinks.html>, accessed on 11 November 2019) software and projected the FPKM levels of the lncRNAs and mRNAs [34]. The DE genes were considered statistically significant by a log<sub>2</sub>-fold change greater or equal to 2 ( $p$ -value < 0.05) or  $p$ -adjust < 0.05 using Ballgown [43].

### 2.8. GO and KEGG Pathways Enrichments

The GO enrichment analysis provided significant source genes that involved different biological functions, and lncRNA DE genes were employed by the GO-seq R package [44]. All the source genes in the pathways and the background genes were charted by GO terms into the database (<http://www.geneontology.org/>, accessed on 11 November 2019). GO terms with corrected  $p$  < 0.05 were considered significantly enriched by DE genes as per the definition of the hypergeometric test. Following that, the term enrichment significance analysis was corrected using the false discovery rate (FDR), and the corrected  $p$ -value (Q-value) was acquired [45]. The KEGG pathways-based study added an explanation of the source

genes and their biological functions (<http://www.genome.jp>, accessed on 11 November 2019). KOBAS 2.0 (<http://kobas.cbi.pku.edu.cn>, accessed on 11 November 2019) software was suggested by reference [46] to test the enriched DE source genes in the KEGG pathways. The enrichment analysis significantly identified the signaling and metabolic pathways where the lncRNAs source genes and total background genes were charted.

### 2.9. Cis and Trans-Regulated Genes Prediction

A classic instance of *trans*-acting is transcription factors and lncRNAs [47], and DE lncRNAs were chosen for forecasting target genes. Increasing proof has shown that lncRNAs can impact the expression of neighboring coding genes through a *cis*-regulating role. We sought to discover the coding genes located both 10 and 100 Kb upstream and downstream of the lncRNAs and figured out their *cis*-functions [48]. The *trans*-role in lncRNAs is to communicate by the level of expression with target genes. The expressed correlation between lncRNAs and coding genes with custom scripts and target genes identified by a Pearson correlation coefficient >0.95 was calculated [49].

### 2.10. Validation of DE lncRNAs and mRNAs Gene via RT-qPCR

Whole RNA was extracted from the immature and mature groups of testes tissues using a Trizol reagent (Life Technologies, USA), and its concentration was measured using a Nanodrop spectrophotometer. The good-quality RNAs were then converted into cDNAs using a QuantiTect Reverse Transcription Kit (Qiagen, Hilden, Germany) in accordance with the manufacturer's protocols. The software programs Primer 3 web version 4.0.0 and Basic Local Alignment Search Tool (BLAST; <https://blast.ncbi.nlm.nih>, accessed on 11 June 2020) were used to design gene-specific primers. The primers used in this study are presented in Table S2. The PCR primers were generated using FastStart SYBR Green Master Mix (Roche, Germany) and the StepOne Plus Real-Time PCR System (Applied Biosystems, Waltham, MA, USA). Each PCR reaction comprised of 7.5- $\mu$ L 2 $\times$ SYBR Green PCR Master Mix, 1.5- $\mu$ L cDNA, 0.5  $\mu$ L each of the reverse and forward primers, 4.7- $\mu$ L nuclease-free water and 10.3  $\mu$ L of ROX dye. The following q-PCR thermal cycles were carried out: (I) pre-denaturation at 95 °C for 10 min, (II) followed by 45 denaturation amplification cycles at 95 °C for 15 s, (III) annealing at 60 °C for 10 s and (IV) an extension cycle at 72 °C for 20 s. The quantification cycle (Cq) of each target gene was normalized to that of the reference *GAPDH* gene. The q-PCR experiment was performed in triplicate to minimize the risk of error in the experiments. The Cq values were obtained and transferred to a Microsoft Excel sheet for a further relative quantification analysis using the  $2^{-\Delta\Delta C_t}$  method [50].

### 2.11. Statistical Analysis

Data on the lncRNA and mRNA transcripts of interest from immature and mature testicular tissues were analyzed using the Student's *t*-test (SPSS 17.0) and presented as the mean  $\pm$  SEM. Before conducting the *t*-test, the data distribution and variances between the two groups were evaluated and found to be normal and homogenous, respectively. The mean values were believed to be significantly different with \*  $p < 0.05$  and \*\*  $p < 0.01$ .

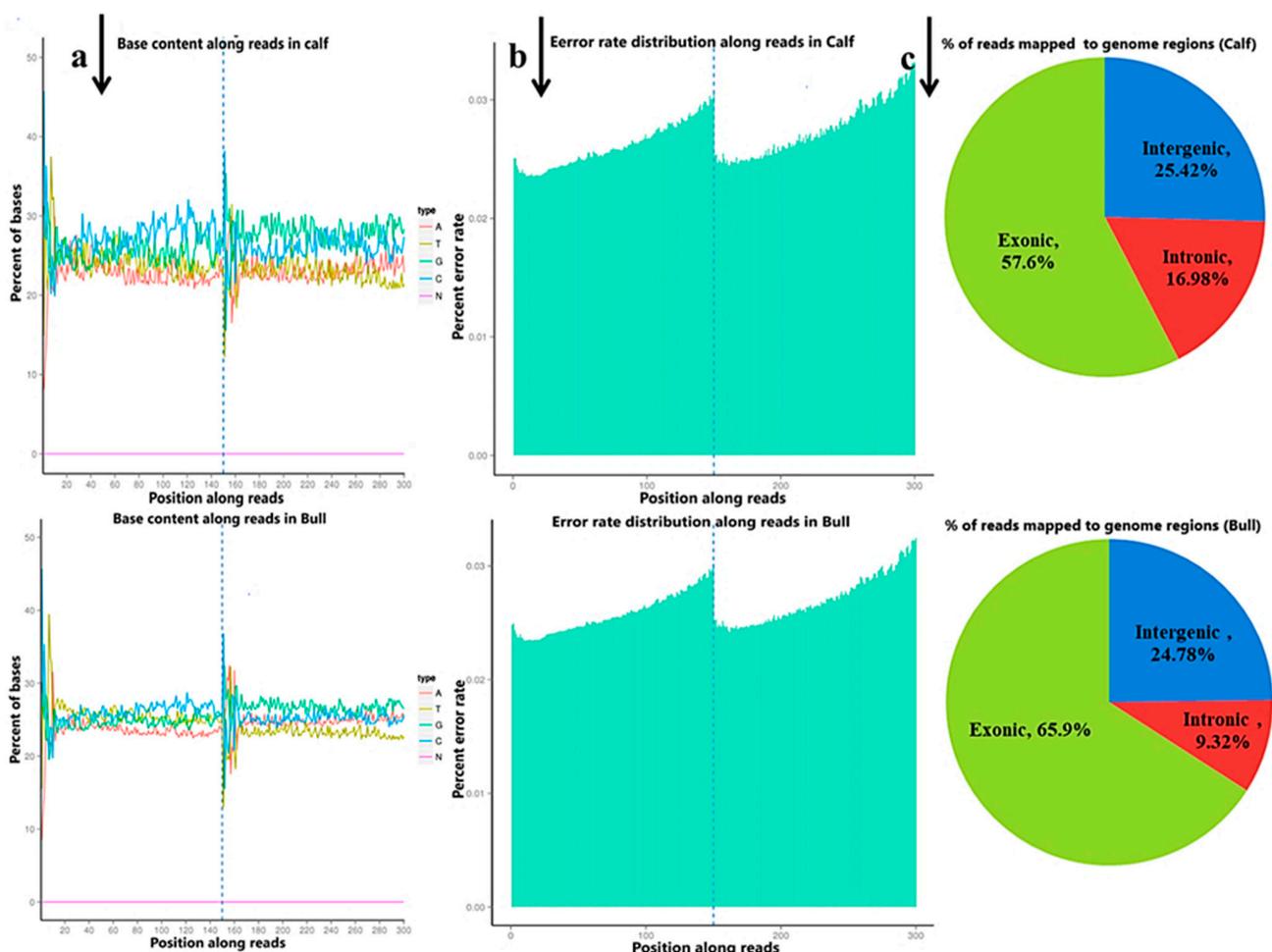
## 3. Results

### 3.1. The Morphological Study of Testicular Tissues

The histomorphology of the bull (*Bos taurus*) testes showed a noteworthy difference during the immature and mature stages. Under a 100 $\times$  microscopic study, the diameter of the seminiferous tubules was much smaller in the immature as compared to the mature testes. At the same time, similar circumstances were found in the interstitial connective tissue of the immature and mature testes. Under a 400 $\times$  microscopic study, we got further details and found more developmental stages of biological spermatozoa in the mature testes than the immature. The Sertoli cells were significantly increased in size and found very near to the basement line of the seminiferous tubules in the mature testes, as shown in Figure S4.

### 3.2. Outline of lncRNAs and mRNAs Sequencing in Calf and Bull Testes

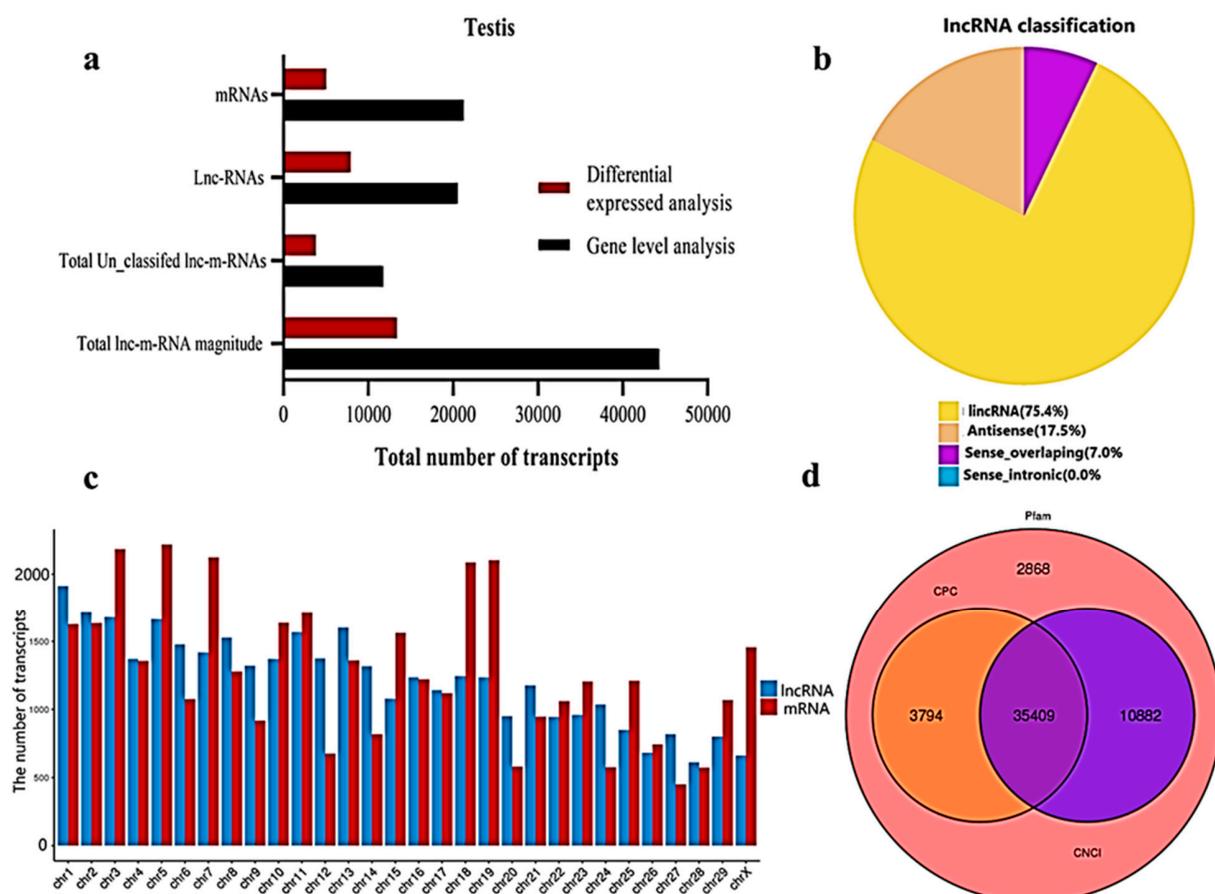
The raw reads were filtered and classified into different segments; consequently, 98.95% (70,333,218 clean reads) of the sequencing data was screened, containing N 0.05% (38,769) reads, low-quality 0.18% (126,760) reads and adapter-related 0.82% (584,096) reads for the bulls, and similar classifications were done for the calves, including clean reads (70,218,119, 98.48%), containing N (37,396, 0.05%), low-quality (141,228, 0.20%) and adapter-related (905,921, 1.27%) (Figure S5). The raw reads for the total samples were found between 118,785,372 and 142,165,686, while the clean reads per sample were 116,004,150 and 140,666,436. Thus, the raw reads and clean reads yielded together 90 GB data, and the GC contents were 47.09–53.57% (Figure 1a). The sequencing process itself has the possibility of machine errors. The distribution check of the sequencing error rate can reflect the quality of the sequencing data, and the sequencing error rate distribution test is used to detect the error % in the sequencing of samples (Figure 1b). All the % Q20 values of the subjected reads in six samples were surpassed by 97%, and all the details are shown in Table S3. More than 95.5% of the clean reads were aligned to *B. taurus* UMD3.1 and are shown in (Figure 1c).



**Figure 1.** The quality summary of the clean reads used in the subsequent analysis were obtained after the original data filtering, sequencing error rate checking and GC content distribution checking. (a) The abscissa shows the base position of the reads, and the ordinate is the proportion of a single base, whereas different colors represent different base types. (b) The abscissa represents the bases position of the reads, and the ordinate is the average error rate of all the reads at this location. The 0–150 bp to the left of the dotted line is the error rate distribution of read1, while the 150–300 bp to the right is the error rate distribution of read2. (c) The percentage of the total number of clean reads mapped to the reference genome. The proportion of reads in the exon region, intron region and intergene region of the genome were calculated.

### 3.3. Identification and Characterizations of lncRNAs and mRNAs

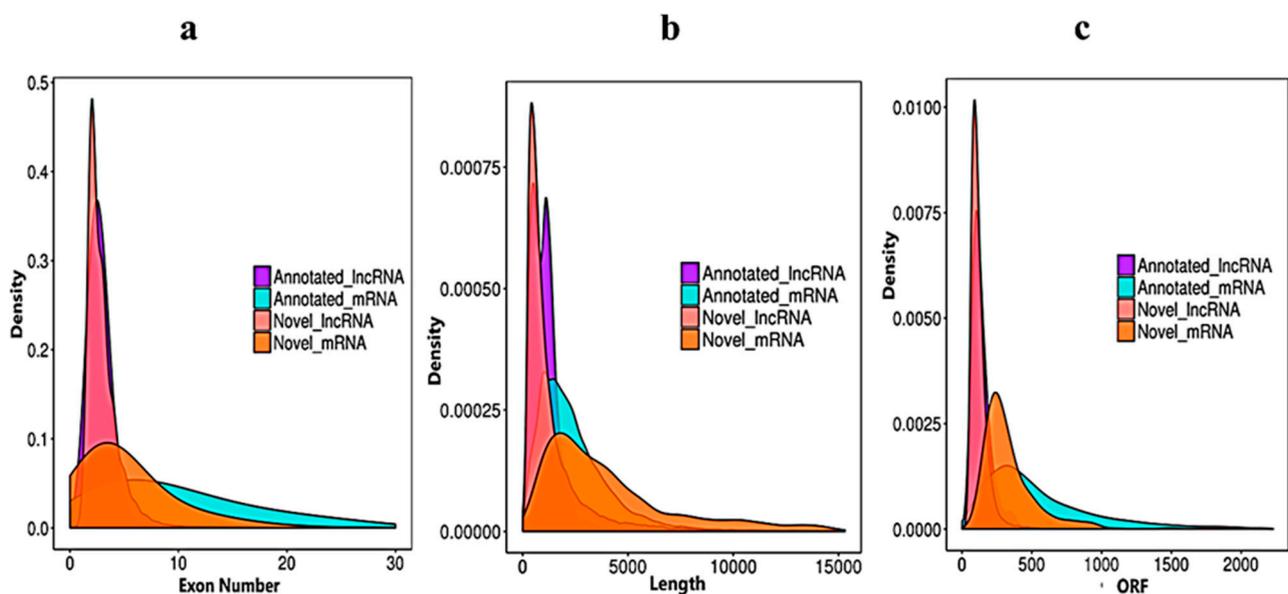
We classified and identified lncRNAs based on the position and direction of their genomic regions relative to mRNAs. A total lnc-mRNA magnitude of 44,279 was found in the calf and bull testes, where, at the gene level analysis, 21,250 mRNAs and 20,533 lncRNAs are identified (Figure 2a). We established an accurate putative lncRNAs and mRNAs network from the assembled transcripts with a stringent pipeline. Based on the structural characteristics and genomic position, the lncRNAs were classified into four subclasses, though denoted as lincRNA (75.4%), antisense (17.6%), sense-overlapping (7.0%) and sense-intronic (0.0%) (Figure 2b). The lncRNAs were overlapped in same or in the opposite direction with the exonic or intronic parts of the coding gene; thus, they became exonic-sense, exonic-antisense, intronic-sense or intronic-antisense, respectively. The genomic locus that was more than 1 kb from any mRNA-encoding genomic region was considered as a long intergenic ncRNA (lincRNA). According to this classification scheme, we found that most testis-specific lncRNAs (75.4%) were lincRNA. We examined the chromosomal localization of the testicular lncRNAs and mRNAs and identified the loci that expressed these transcripts. Our research confirmed that the transcripts were uniformly spread throughout the chromosomes of the bulls and calves (Figure 2c). The coding potential of the lncRNA and mRNA were analyzed by using CPC2, CNCI and Pfam software, and a total of 35,409 coding transcripts were identified (Figure 2d).



**Figure 2.** General properties of the lncRNAs and mRNAs in the premature and mature bovine testes were identified. (a) The transcripts were categorized as the total magnitude, unclassified and specified lncRNAs and mRNAs while analyzed at the gene level and differentially expressed patterns. (b) The distribution of lncRNAs at the genome region at a high magnitude were identified from the exon, intergenic and intron parts of the genome. (c) The distribution of testis mRNAs and lncRNAs in bovine chromosomal sets. (d) Venn diagram showing the coding potential results of the lncRNAs and intersection of the coding tools, such as CNCI, PFAM and CPC2.

### 3.4. Characteristics Comparison between lncRNAs and mRNAs

We verified the general characteristics of the novel lncRNAs and mRNAs by comparing the transcript length, exon number and ORF length, which also differentiated between the lncRNAs and mRNAs. The number of exons of mRNAs about (30) was higher than of the lncRNAs (10). The majority of lncRNAs have three or less than three exons, whereas mRNAs have five or more exons (Figure 3a). In the length comparison, we observed that the lncRNAs (an average of 2000 bp) were significantly shorter than the mRNAs (an average of 3000 bp), as shown in (Figure 3b). The open reading frame (ORF) is the normal nucleotide sequence of a structural gene. From the initiation codon to the termination codon, the open reading frame encodes a complete polypeptide chain, without any interruption in the translation. When we compared the lncRNA with mRNA ORFs, the sequences of annotated mRNAs and annotated lncRNAs were extracted by annotations of the known gene structure. The obtained ORF sequence was converted into a protein sequence, and the length distribution was obtained (Figure 3c).

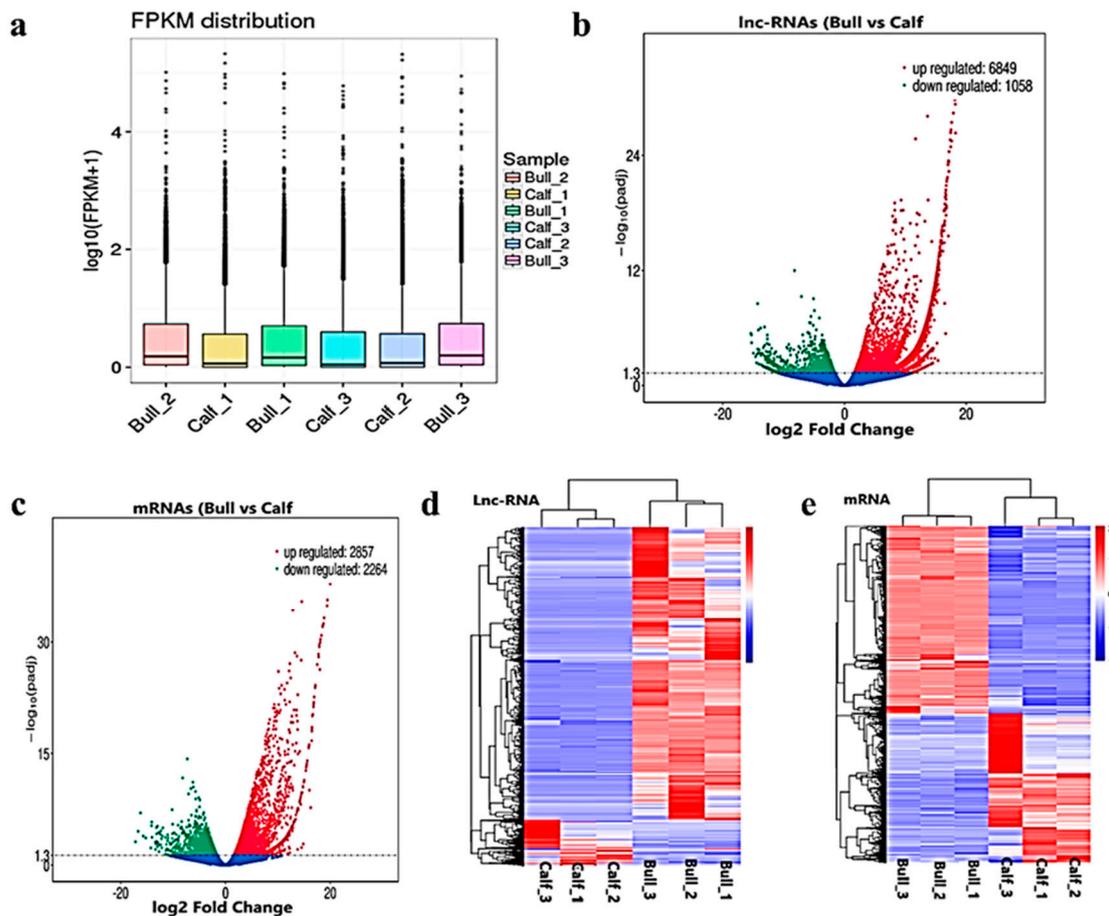


**Figure 3.** Sequencing characteristics of the lncRNAs and mRNAs have shown. (a) Exon number distribution of the annotated and novel lncRNAs (purple and pink) vs. the mRNAs denoted blue and dark yellow. (b) Transcript length comparisons of the annotated and novel lncRNAs (purple and pink) vs. the mRNAs (blue and dark yellow). (c) Open reading frame nucleotide sequence comparisons of the annotated and novel lncRNAs (purple and pink) vs. the mRNAs (blue and dark yellow).

### 3.5. Differentially Expressed Gene-Level lncRNAs and mRNAs in Testicular Tissues

After the quantification process, the expression pattern of the lncRNAs and mRNAs were identified through the cuff-diff and Ballgown tools. The average transcripts expression level was higher in the bulls compared to the calves group (Figure 4a). Thus, we recorded the differentially expressed (DE) lncRNAs and mRNAs between immature and mature testes by using Ballgown. The significant levels of lncRNAs and mRNAs among the testes samples were calculated, and the parameter of significance was taken into account, whereas the log<sub>2</sub>-fold change was considered higher than two or equal to and *p*-adjusted < 0.05. As a result, 7908 lncRNAs and 5122 mRNAs were found to be differentially expressed (DE) in immature and mature bovine testes. Further, we identified 6849 lncRNAs substantially upregulated and 1058 are downregulated in the testes (*p*-value < 0.05). In the adult testes group, *p*-adjust < 0.05 revealed that 2857 mRNAs were upregulated and 2264 were downregulated. These up and down highlights are displayed in volcanic plots (Figure 4b,c), respectively. We also analyzed the DE lncRNAs and mRNAs by using the hierarchical clustering method, and the clustering was another way to display the DE genes, which

bring together genes with similar expression patterns that may have common functions or participate in common metabolic and signaling pathways. The genes cluster on the left side was made due to similar expressions (fold change >2,  $p < 0.05$ ), and the calves and bulls were represented by columns, while the expression from blue to red represented them gradually upregulated (Figure 4d,e).



**Figure 4.** Differential expressed level of the transcripts and protein-coding genes in the testes. (a) The distribution of the expression levels of the genes or transcripts in different samples is shown by a box graph. The boxplot of each region corresponds to five statistics, e.g., the maximum, upper quartile, median, lower quartile and minimum, respectively. (b) The volcanic map visualizes the overall distribution of the lncRNAs genes with significant differences in their expression. (c) Differential mRNA expressions between the calves and bulls. The upregulated genes are represented by red dots, downregulated genes by green dots and the blue dots are genes without significant changes. (d) Clustering map of DE lncRNAs among the calf and bull testes sample. (e) Clustering map of the DE mRNAs among the samples. Red indicates upregulated coding genes, and blue indicates downregulated gene products.

### 3.6. Functional Annotation and Enrich KEGG Pathways Analysis of DE mRNAs

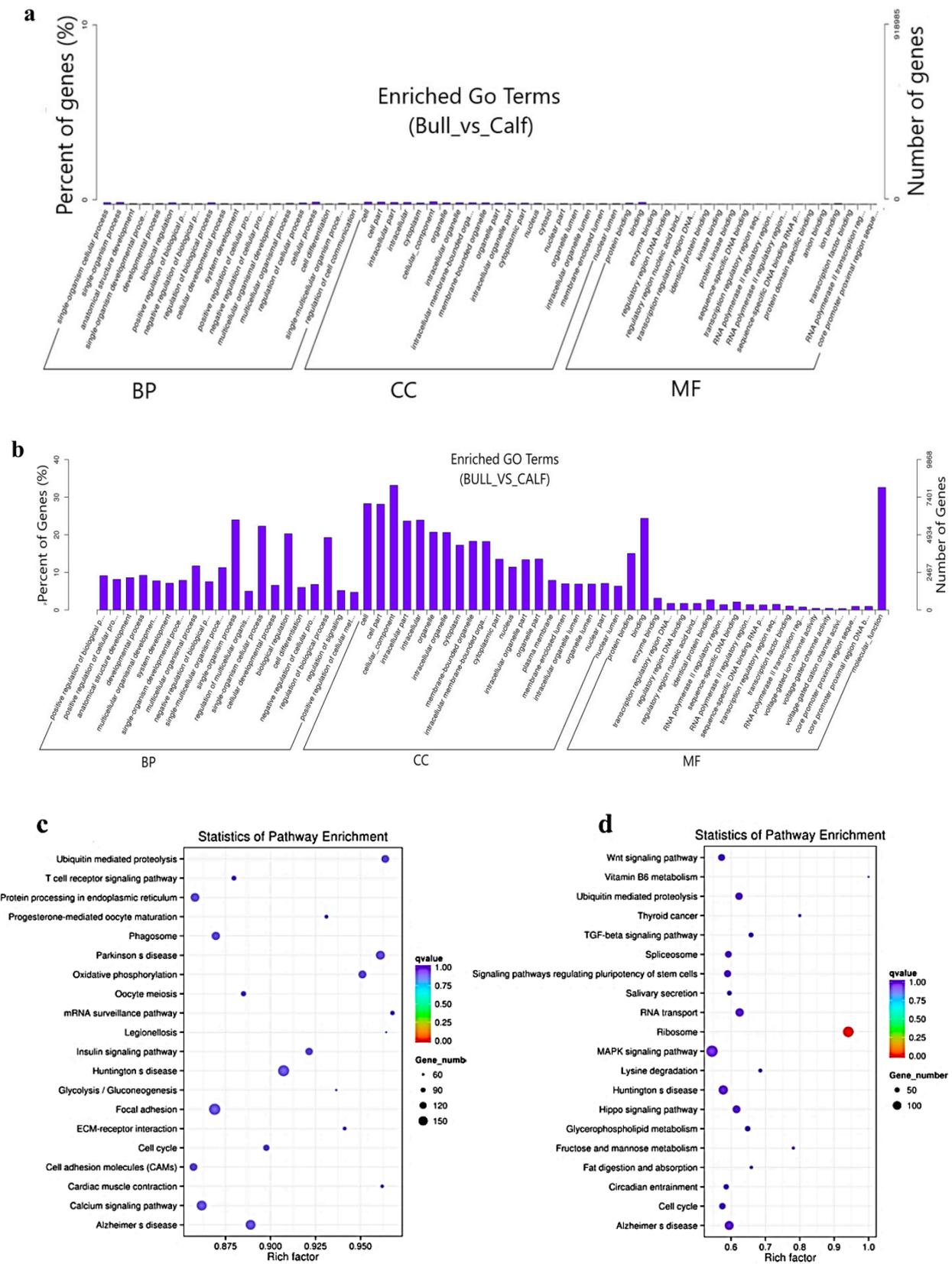
The transcriptome level examination showed 3857 DE mRNAs (2352 upregulated and 1504 downregulated) were marked, and these mRNAs were preferred for the enrichment analysis. The DE mRNAs were substantially enriched in GO terms, including the reproduction, male gamete generation, spermatogenesis, multicellular organism reproduction, single-organism reproductive process, spermatids differentiation, spermatids development and sexual reproduction, which were the top eight GO annotations terms (Figure 5a). Detailed descriptions of the GO terms are available in Table S4. As for as the statistics of the pathway enrichment, the DE mRNAs were enriched in the top twenty KEGG-rich factors pathways, such as adherens junctions, cell adhesion molecules (CAMs) and lysine degradation (Figure 5b). Table S5 represents the top 10 most significant enriched pathways



was within 100 Kb, and DE *cis*-acting lncRNAs were shown (Table 1). The co-expression network built for bull and calf testes comprised 1168 mRNAs and 410 lncRNAs, which have regulatory potential. Taking 100 Kb as the cutoff, 410 out of 550 lncRNAs were found as the nearest neighbors of 1168 mRNAs. The findings of the GO enrichment study revealed that there were 70 GO descriptions significantly observed ( $p$ -adj-value < 0.05). The top five GO terms were cellular process, cell, cell part, cellular component and regulation of the biological process (Figure 6a). The top 10 significantly enriched GO descriptions of BP, CC and MF, which regulated the target genes of the lncRNAs, are shown in Table S6. We also studied the co-location target genes in a *trans* way for the lncRNAs. A total of 550 lncRNAs and 11,470 genes were detected using Pearson correlation 0.95 as the cutoff. The GO enrichment analysis for co-located target genes identified 190 significant GO terms ( $p < 0.05$ ), while the top five GO terms were spermatogenesis, sexual reproduction, male gamete generation, gamete generation and reproduction (Figure 6b). The KEGG analysis for co-expression identified 10 pathways ( $p < 0.05$ ), and the top 20 pathways are shown in (Figure 6c), such as the T-cell receptor signaling pathway, Ubiquitin-mediated proteolysis, mRNA surveillance pathway, ECM–receptor interaction, focal adhesion and cell cycle pathway. Table S7 shows the top 10 most significant enriched pathways regulated by lncRNAs in Wandong testes development. Concerning the pathways, the co-located *trans*-target genes for the lncRNAs were supplemented in several KEGG pathways—for example, the ribosome pathway, Lysine degradation, fructose and mannose metabolism, Vitamin B6 metabolism and TGF-beta signaling pathway (Figure 6d).

**Table 1.** Top 20 DE *cis*-acting long non-coding RNAs in calf and bull testes.

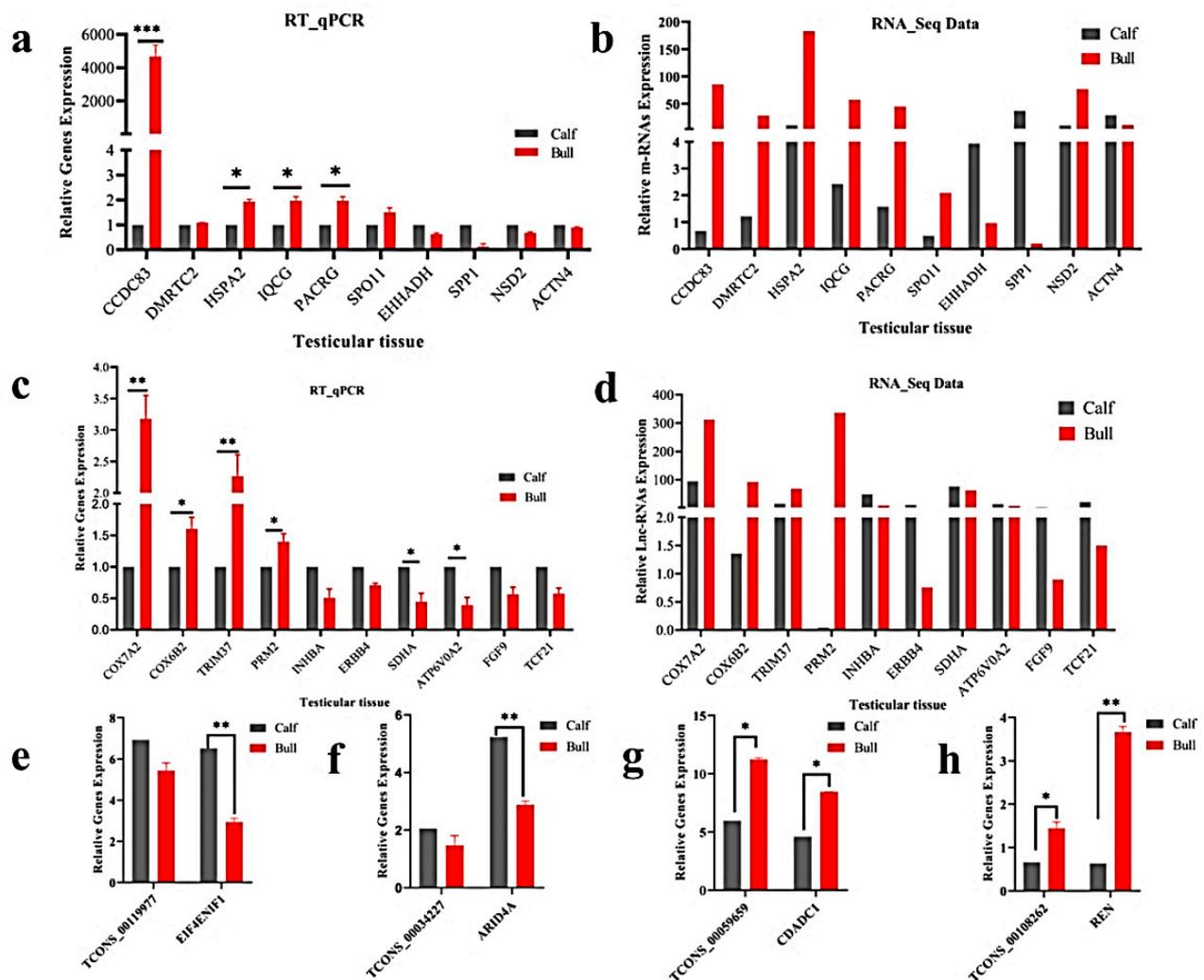
DE-lnc-RNA/ID	Chromosome/Location	Lnc-RNA-Status	Regulation	DE <i>cis</i> -Regulated Genes/ID	Gene-Symbol	Chromosome/Location	Enrich Pathways
TCONS_00223595	25: 40984697–41226938	Novel-lncRNA	Up	ENSBTAG00000002474	<i>MAD1L1</i>	25: 40984696–41222673	Cell Cycle
TCONS_00218376	24: 47452790–47459159	Novel-lncRNA	Down	ENSBTAG00000039916	<i>SMAD2</i>	24: 47500431–47586992	Cell Cycle
TCONS_00285708	4: 49484422–49563305	Novel-lncRNA	Down	ENSBTAG00000006732	<i>NRCAM</i>	4: 49248702–49563305	Cell Adhesion Molecules (CAMs)
TCONS_00265960	3: 3819912–3824649	Novel-lncRNA	Down	ENSBTAG00000012025	<i>LMX1A</i>	3: 3653710–3828022	Cell Adhesion Molecules (CAMs)
TCONS_00263842	3: $1.02 \times 10^8$ – $1.02 \times 10^8$	Novel-lncRNA	Up	ENSBTAG00000000253	<i>PTPRF</i>	3: 102414018–102498351	Adherens Junction
TCONS_00092191	15: 30147243–30156216	Novel-lncRNA	Down	ENSBTAG00000015939	<i>NECTIN1</i>	15: 30211169–30282836	Adherens Junction
TCONS_00226910	25: 35113818–35117275	Novel-lncRNA	Up	ENSBTAG00000026273	<i>MYL10</i>	25: 35177976–35190739	Tight Junction
TCONS_00003412	1: 53213665–53221154	Novel-lncRNA	Up	ENSBTAG00000018399	<i>MYH15</i>	1: 53114698–53273181	Tight Junction
TCONS_00314718	6: 98071217–98093465	Novel-lncRNA	Down	ENSBTAG00000005745	<i>HPSE</i>	6: 98071126–98118656	Metabolic Pathway
TCONS_00305722	5: 76121498–76122025	Novel-lncRNA	Down	ENSBTAG00000030632	<i>ALG10</i>	5: 76104065–76113301	Metabolic Pathway
TCONS_00106232	16: 56589861–56596165	Novel-lncRNA	Down	ENSBTAG00000011706	<i>TNR</i>	16: 56545152–56764975	ECM-receptor Interaction
TCONS_00340198	7: 75125739–75137127	Novel-lncRNA	Up	ENSBTAG00000014773	<i>HMMR</i>	7: 75137017–75184925	ECM-receptor Interaction
TCONS_00282996	4: 11670810–11680838	Novel-lncRNA	Down	ENSBTAG00000013472	<i>COL1A2</i>	4: 11776162–11823181	Focal Adhesion
ENSBTAT00000080812	5: 32333604–32349222	Annotated-lncRNA	Up	ENSBTAG00000013155	<i>COL2A1</i>	5: 32283180–32313172	Focal Adhesion
TCONS_00098825	15: 37889979–37909497	Novel-lncRNA	Up	ENSBTAG00000005218	<i>PDE3B</i>	15: 37935445–38104280	cAMP Signaling Pathway
TCONS_00267075	3: 16683665–16685238	Novel-lncRNA	Up	ENSBTAG00000006234	<i>NPR1</i>	3: 16683665–16700685	cAMP Signaling Pathway
TCONS_00118360	17: 53611398–53625902	Novel-lncRNA	Down	ENSBTAG00000004457	<i>ORAI1</i>	17: 53610136–53626039	Calcium Signaling Pathway
TCONS_00038312	11: 22365812–22406197	Novel-lncRNA	Down	ENSBTAG00000013861	<i>SLC8A1</i>	11: 22408368–22787993	Calcium Signaling Pathway
TCONS_00310078	6: 11726041–11732204	Novel-lncRNA	Down	ENSBTAG00000014463	<i>CAMK2D</i>	6: 11800357–12189732	GnRH Signaling Pathway
TCONS_00321828	6: $1.01 \times 10^8$ – $1.01 \times 10^8$	Novel-lncRNA	Down	ENSBTAG00000020048	<i>MAPK10</i>	6: 100907869–101279063	GnRH Signaling Pathway



**Figure 6.** The co-location and co-expression methods were used and identified the lncRNA *cis*- and *trans*-regulated genes. (a) The scheme represents the co-expressed (*trans*-regulated) genes of the lncRNAs. (b) Co-located (*cis*-regulated) genes of the lncRNAs. (c) The top 20 significant enriched KEGG pathways for co-expressed lncRNA target genes are listed. (d) The top 20 significant enriched KEGG pathways for co-located lncRNA target genes are listed.

### 3.8. Validation of Differentially Expressed LncRNAs and mRNAs

We randomly selected 10 DE mRNA-predicted genes (e.g., *CCDC83*, *DMRTC2*, *HSPA2*, *IQCG*, *PACRG*, *SPO11*, *EHHADH*, *SPP1*, *NSD2* and *ACTN4*) and 10 lncRNA-predicted genes (*COX7A2*, *COX6B2*, *TRIM37*, *PRM2*, *INHBA*, *ERBB4*, *SDHA*, *ATP6VOA2*, *FGF9* and *TCF21*) for the immature and mature testes samples, using RT-qPCR for the purpose of validating the genes that identified in RNA-Seq results. The RT-qPCR relative expression results were parallel to the RNA-seq results as shown (Figure 7a–d), whereas the primer details are available in Table S2. The achieved results demonstrated that the RNA-seq data was valid and dependable, with a specificity in the bull testes developmental stages. To further elucidate the *cis*-acting potential of the lncRNAs and their target genes in the testicular growth of bulls, four candidate lncRNA: TCONS\_00119977, TCONS\_00034227, TCONS\_00059659 and TCONS\_00108262, along with *cis*-acting genes: *EIF4ENIF1*, *ARID4A*, *CDADC1* and *REN* were selected, while the RT-qPCR and the obtained results were in line with the RNA-Seq findings (Figure 7e–h).



**Figure 7.** Validation of the RNA-Seq results via RT-qPCR in the immature and mature stages of bulls. The panel figures (a–d) compared the DE mRNA and lncRNA sequencing results and the RT-qPCR findings. The panel figures (e–h) showed the expression decoration of four lncRNAs and their *cis*-acting genes in two different developmental stages. The data were measured by the  $2^{-\Delta\Delta Ct}$  method, and *GAPDH* was used as the housekeeping gene. The data are presented as the mean  $\pm$  SEM, where \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

#### 4. Discussion

In the present research, we studied the DE lncRNAs and mRNAs during the testicular growth of local Wandong cattle and also investigated the potential regulatory roles of lncRNAs and *cis*-acting genes in cattle bull sexual development. The histological transition in testes occurs primarily at three developmental phases: the premature, mature and adulthood periods. Testicle growth is usually followed by spermatogenesis, so we found only spermatogonia, spermatocytes and Sertoli cells in the seminiferous tubules in the 3-month-old group, while spermatids and spermatozoa were also observed in the 3-year-old adult group. The very similar differences were also noticed in 3-day-old neonatal and 13-month-old mature Angus bulls [8]. lncRNAs have received a lot of coverage as a novel regulatory player in cellular growth [51]. lncRNAs are abundant in the testes, and their average number fluctuates in a similar pattern to mRNAs during the developmental changes [39,52,53]. lncRNAs were found, in several experiments, to play a significant role in mammalian testes formation and spermatogenesis. Numerous lncRNAs from testicular tissues or germ cells of mice [54], porcine [21] and humans [35,55] have been identified. Weng et al. [55] isolated 240 proven and 537 putative lncRNAs from the testes of 30- and 180-day-old male Shaziling Chinese pigs.

The accurate putative lncRNAs and mRNAs combined network was assembled and characterized. A total of 21,250 mRNA and 20,533 lncRNA transcripts were identified at two developmental stages of local bulls. Gao et al. [8] identified a total of 23,735 lncRNAs and 22,118 mRNAs in two developmental stages of Angus cattle testes, in which 540 lncRNAs and 3525 mRNAs were differentially expressed (DE) between the neonatal and adult stages. Furthermore, we discovered that the mRNAs and lncRNAs had identical chromosomal distributions at the autosomes and the X-chromosome, suggesting that the lncRNAs interacted closely with the mRNAs. The higher numbers of lncRNAs (6460) and mRNAs (42,300) were also characterized and identified in sheep testes during sexual maturation [23]; similarly, 36,727 lncRNAs were found in the postnatal (30-day and 180-day-old) pig testes [21]. The maximal numbers of both the lncRNAs and mRNAs were observed in the testes compared to the other tissues, suggesting that the testes are characterized by a high level of transcriptomic diversity and complexity [56].

The number of lncRNAs discovered in our study was substantially higher than that observed in bovine skeletal muscle (4155) [57]; from ovaries of Duroc pig (2076) lncRNAs [58], (3481) lncRNAs in milk exosomes [59], in mammary gland (6450) lncRNAs [60] and bovine ileum (1568) lncRNAs [61], it is revealing that lncRNAs are testes-specific. We differentiated between lncRNAs and mRNAs via comparing the transcript length, exon number and ORF length. The lncRNA exons were found relatively shorter as compared to mRNAs. We found that this genomic comparison between the lncRNAs and mRNAs was in line with recent studies on other mammals [20,22].

A total of 5122 mRNAs and 7908 known lncRNAs were significantly DE between the immature and mature groups. Further, we identified 6849 lncRNAs significantly up-regulated and 1058 downregulated, whereas 2857 mRNAs were upregulated and 2264 were downregulated between the calf and bull groups. A magnitude of 23,735 lncRNAs and 22,118 mRNAs were detected in the testes of Angus cattle at two developmental phases, where 540 lncRNAs and 3525 mRNAs were discovered DE in 3-day-old calves and 13-month-old bulls [8]. Weng et al. [9] identified 15,528 lncRNAs, while these transcripts consisted of 5032 known and 10,496 novel lncRNAs in four different postnatal developments of testes in porcine. Bao et al. [53] investigated the expression profiling of thousands of lncRNAs and mRNAs and hundreds of lincRNAs in male spermatozoa progress and testes development. According to the previous studies, lncRNA expressions in porcine and mouse testes were strongly associated with specific developmental stages [19,21]. Our systematic studies also proposed that the lncRNAs and mRNAs were specifically expressed in the two developmental stages of bovine testes.

To thoroughly analyze the biological functions of lncRNAs and mRNAs in local bull testes, GO and KEGG studies were done for the linked genes of DE lncRNAs and

mRNAs. Through these bioinformatic analyses, candidate genes were found, and these genes were linked to male reproduction. The results showed that 247 mRNA transcripts were enriched for the following GO descriptions: biological process, molecular function and cellular component. The source genes linked to mRNAs, such as *CCDC83*, *DMRTC2*, *HSPA2*, *IQCG*, *PACRG*, *SPO11*, *EHHADH*, *SPP1*, *NSD2* and *ACTN4*, and genes linked to lnc-RNA transcripts, like *COX7A2*, *COX6B2*, *TRIM37*, *PRM2*, *INHBA*, *ERBB4*, *SDHA*, *ATP6VOA2*, *FGF9* and *TCF21*, were involved in the GO sub-descriptions like spermatid development, spermatogenesis, spermatid differentiation, reproductive process, male gonad development, cellular process involved in reproduction, sperm motility, inactivation of MAPK activity and steroid hormone-mediated signaling pathway, which was similar to the GO terms analysis in sheep testes [23]. A catalog of candidate genes was discovered in the testes of Angus cattle, and some of these genes were linked to male reproduction. The *FSCN3*, *TNP2*, *IQCF1*, *PRM2* and *MORC1* were related to spermatogenesis. *KLHL10*, *TCF21* and *SFRP1* may be involved in the development of male gonads and the androgen receptor [8]. Furthermore, the biological pathway reports verified that many DE source genes actively participated in the reproduction-linked signaling pathways, like adherens junctions, cell adhesion molecules (CAMs) and lysine degradation, and these pathways were also associated with male reproduction [62,63].

The cell cycle, lysine degradation, tight junction and adherence junction pathways and the host genes that link to ncRNA transcripts work together in mammalian testes growth, spermatogenesis and primary sexual activity [64]. It is well-known that gene expression profiling regulates testes development and spermatogenesis in mammals; we identified many DE mRNA and lncRNA genes that may take part in bovine reproduction. The candidate genes like the double sex and mab-3-related transcription factors (*DMRT*) and the family C2 (*DMRTC2*) are required for mammalian testicular functions and spermatogenesis, and *DMRTC2* is highly conserved in sheep spermatogenesis and testis growth [65] and also plays a significant role in fox hybrids during spermatogenesis [66]. The known candidate *HSPA2* gene is enriched in testes and is a member of the 70-kDa heat shock protein family that facilitates the folding, delivery and assembly of protein complexes and has been linked to in vitro fertilization (IVF) performances in both humans and rodents [67,68]. These source gene *IQCG* was mainly involved in the tight junction and adherens junction pathways and regulated the testes growth and reproduction in bovines [69] and are also found in sperm morphogenesis to give rise to mature spermatozoon in mammals [70]. The transcript *SPO11* showed the highest expression in human fetal gonadal development [71]. Another study highlighted that the *INHBA* gene belongs to the major *TGFB* pathway and regulates early germ cell proliferation in rainbow trout testes [72]. Gene expressions are sex-specific in the sex development of mammals. Different genes express in different phases; correspondingly, the *FGF9* gene has been defined in male gonadal morphogenesis [73].

The past literature proposed that lncRNA expression can manipulate and highly interact with neighboring protein-coding gene expression via transcriptional coactivation [74,75]. Gao et al. [8] predicted the potential role of *cis*-target genes of differentially expressed lncRNAs in the regulation of testes development and spermatogenesis. In this study, we selected four differentially expressed lncRNA transcripts, along with their *cis*-target genes, for RT-qPCR validation. We found consistency in the expression results, and thus, the genetic regulatory network plays a vital role in cattle testes development and germ cell proliferation [76].

## 5. Conclusions

The lncRNAs and mRNAs were abundantly expressed in Wandong cattle testes development, and many of them were significantly differentially expressed between the calves and bulls. The lncRNAs may have significant biological functions in bovine testicular development. Thus, we thoroughly analyzed the GO descriptions and found that the source genes were enriched in gonad development, spermatid development, spermatogenesis, spermatid differentiation and the reproductive process. The lncRNA-linked genes *COX7A2*,

*COX6B2*, *TRIM37*, *PRM2*, *INHBA*, *ERBB4*, *SDHA*, *ATP6VOA2*, *FGF9* and *TCF21* and mRNAs *CCDC83*, *DMRTC2*, *HSPA2*, *IQCG*, *PACRG*, *SPO11*, *EHHADH*, *SPP1*, *NSD2* and *ACTN4* were found to be actively associated with bull sexual maturity and spermatogenesis. The obtained results can provide a molecular fingerprint for assessing the developmental status of bull testes and also facilitate the selection of local elite bulls.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/ani11072006/s1>: Figure S1: The testes samples were collected from the 3 months old experimental calf and 3 years old sexually mature bull. Figure S2: Flow chart was adopted for cDNA library construction, and quality inspection. Figure S3: The different steps are involved in screening of novel lnc-RNAs and mRNAs. Figure S4: The histomorphological analysis of bull testicular tissues at three months and three years ages were observed under microscope at  $\times 100$ ,  $\times 200$ , and  $\times 400$  magnifications. Bars = 100  $\mu\text{m}$ . The segments (a), (c) and (e) represent the morphology of 3-month-old testes, whereas (b) (d), and (f) represents the morphology of 3 years bulls. The red arrows indicate different cells types. 1: spermatogonia, 2: Leydig cells, 3: Sertoli cells, 4: spermatocytes, 5: round spermatids, 6: elongated spermatids, 7: sperm. Figure S5: Pie charts represent the raw data classification in to different filtering reads in calf and bull, whereas different color ratios represent the proportions of different components. Table S1: Scrotal circumference, grass motility threshold and progressive motility were evaluated for the sexually mature bulls. Table S2: Primers used for RT-qPCR validation of m-RNA and lnc-RNAs sequencing data. Table S3: Raw data quality summery of calves and bulls testes sequencing data. Table S4: The top 10 significantly enriched GO description of BP, CC and MF for mRNAs targets genes. Table S5: The top 10 most significant enriched pathways regulated by mRNAs in the bovine testes development. Table S6: The top 10 significantly enriched GO descriptions of BP, CC and MF and regulated the targets genes of lnc-RNAs. Table S7: The top 10 most significant enriched pathways regulated by lncRNAs in the bovine testes development.

**Author Contributions:** Conceptualization, I.M.K. and H.L.; methodology, H.Y. and J.Z.; software, I.M.K.; validation, X.Z., J.Z. and I.M.K.; formal analysis, M.R.; investigation, H.L.; resources, Y.Z.; data curation, H.L.; the original draft was prepared by I.M.K.; final draft and editing, Y.Z.; visualization, M.R. and H.Y.; supervision, H.L.; project administration, Y.Z. and funding acquisition, Y.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Research and Development Program of China (2018YFD0501700); the National Natural Science Foundation of China (31101696) and the Anhui Province Modern Agriculture Industry (cattle, goat and sheep) Technology System (AHCYJSTX-07).

**Institutional Review Board Statement:** The collection of the experimental animals was carried out under strict ethical procedures and approved (SYXK 2016-007) by the Animal Care Unit of the College of Animal Science and Technology of the University of Anhui Agriculture.

**Data Availability Statement:** All data generated or analyzed during this study are available from the corresponding authors upon reasonable request.

**Acknowledgments:** We highly acknowledge Fengyang County Daming Agriculture and Animal Husbandry Technology Development Co., Ltd., which provided the animals and physical support.

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

## References

1. García-Ruiz, A.; Cole, J.B.; VanRaden, P.M.; Wiggans, G.R.; Ruiz-López, F.J.; Van Tassell, C.P. Changes in genetic selection differentials and generation intervals in US Holstein dairy cattle as a result of genomic selection. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E3995–E4004. [CrossRef] [PubMed]
2. Meuwissen, T.H.; Hayes, B.J.; Goddard, M.E. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* **2001**, *157*, 1819–1829. [CrossRef] [PubMed]
3. König, S.; Simianer, H.; Willam, A. Economic evaluation of genomic breeding programs. *J. Dairy Sci.* **2009**, *92*, 382–391. [CrossRef] [PubMed]
4. Hecht, N.B. Molecular mechanisms of male germ cell differentiation. *Bioessays* **1998**, *20*, 555–561. [CrossRef]
5. Luk, A.; Chan, W.-Y.; Rennert, O.M.; Lee, T.-L. Long noncoding RNAs in spermatogenesis: Insights from recent high-throughput transcriptome studies. *Reproduction* **2014**, *147*, R131–R141. [CrossRef]

6. Chocu, S.; Calvel, P.; Rolland, A.D.; Pineau, C. Spermatogenesis in mammals: Proteomic insights. *Syst. Biol. Reprod. Med.* **2012**, *58*, 179–190. [[CrossRef](#)] [[PubMed](#)]
7. Wang, X.; Yang, C.; Guo, F. Integrated analysis of mRNAs and long noncoding RNAs in the semen from Holstein bulls with high and low sperm motility. *Sci. Rep.* **2019**, *9*, 2092. [[CrossRef](#)]
8. Gao, Y.; Li, S.; Lai, Z.; Zhou, Z.; Wu, F.; Huang, Y.; Lan, X.; Lei, C.; Chen, H.; Dang, R. Analysis of Long Non-Coding RNA and mRNA Expression Profiling in Immature and Mature Bovine (*Bos taurus*) Testes. *Front. Genet.* **2019**, *10*, 646. [[CrossRef](#)]
9. Weng, B.; Ran, M.; Chen, B.; He, C.; Dong, L.; Peng, F. Genome-wide analysis of long non-coding RNAs and their role in postnatal porcine testis development. *Genomics* **2017**, *109*, 446–456. [[CrossRef](#)]
10. Yang, H.; Wang, F.; Li, F.; Ren, C.; Pang, J.; Wan, Y.; Wang, Z.; Feng, X.; Zhang, Y. Comprehensive analysis of long noncoding RNA and mRNA expression patterns in sheep testicular maturation. *Biol. Reprod.* **2018**, *99*, 650–661. [[CrossRef](#)]
11. Brockdorff, N.; Ashworth, A.; Kay, G.F.; McCabe, V.M.; Norris, D.P.; Cooper, P.J.; Swift, S.; Rastan, S. The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. *Cell* **1992**, *71*, 515–526. [[CrossRef](#)]
12. Rinn, J.; Guttman, M. RNA and dynamic nuclear organization. *Science* **2014**, *345*, 1240–1241. [[CrossRef](#)]
13. Wu, J.; Zhu, H.; Song, W.; Li, M.; Liu, C.; Li, N.; Tang, F.; Mu, H.; Liao, M.; Li, X. Identification of conservative microRNAs in Saanen dairy goat testis through deep sequencing. *Reprod. Domest. Anim.* **2014**, *49*, 32–40. [[CrossRef](#)]
14. Guan, Y.; Liang, G.; Hawken, P.A.; Malecki, I.A.; Cozens, G.; Vercoe, P.E.; Martin, G.B. Roles of small RNAs in the effects of nutrition on apoptosis and spermatogenesis in the adult testis. *Sci. Rep.* **2015**, *5*, 1–14. [[CrossRef](#)]
15. Zimmermann, C.; Romero, Y.; Warnefors, M.; Bilican, A.; Borel, C.; Smith, L.B.; Kotaja, N.; Kaessmann, H.; Nef, S. Germ cell-specific targeting of DICER or DGCR8 reveals a novel role for endo-siRNAs in the progression of mammalian spermatogenesis and male fertility. *PLoS ONE* **2014**, *9*, e107023. [[CrossRef](#)] [[PubMed](#)]
16. Chen, L.-L.; Zhao, J.C. Functional analysis of long noncoding RNAs in development and disease. *Syst. Biol. RNA Bind. Proteins* **2014**, 129–158. [[CrossRef](#)]
17. Xu, J.; Bai, J.; Zhang, X.; Lv, Y.; Gong, Y.; Liu, L.; Zhao, H.; Yu, F.; Ping, Y.; Zhang, G. A comprehensive overview of lncRNA annotation resources. *Brief. Bioinform.* **2017**, *18*, 236–249. [[CrossRef](#)] [[PubMed](#)]
18. Dianatpour, A.; Ghafouri-Fard, S. Long non coding RNA expression intersecting cancer and spermatogenesis: A systematic review. *Asian Pac. J. Cancer Prev. APJCP* **2017**, *18*, 2601. [[PubMed](#)]
19. Sun, J.; Lin, Y.; Wu, J. Long non-coding RNA expression profiling of mouse testis during postnatal development. *PLoS ONE* **2013**, *8*, e75750. [[CrossRef](#)]
20. Wen, K.; Yang, L.; Xiong, T.; Di, C.; Ma, D.; Wu, M.; Xue, Z.; Zhang, X.; Long, L.; Zhang, W. Critical roles of long noncoding RNAs in *Drosophila* spermatogenesis. *Genome Res.* **2016**, *26*, 1233–1244. [[CrossRef](#)] [[PubMed](#)]
21. Ran, M.; Chen, B.; Li, Z.; Wu, M.; Liu, X.; He, C.; Zhang, S.; Li, Z. Systematic identification of long noncoding RNAs in immature and mature porcine testes. *Biol. Reprod.* **2016**, *94*, 71–79. [[CrossRef](#)]
22. Liu, Y.; Sun, Y.; Li, Y.; Bai, H.; Xue, F.; Xu, S.; Xu, H.; Shi, L.; Yang, N.; Chen, J. Analyses of long non-coding RNA and mRNA profiling using RNA sequencing in chicken testis with extreme sperm motility. *Sci. Rep.* **2017**, *7*, 9055. [[CrossRef](#)] [[PubMed](#)]
23. Zhang, Y.; Yang, H.; Han, L.; Li, F.; Zhang, T.; Pang, J.; Feng, X.; Ren, C.; Mao, S.; Wang, F. Long noncoding RNA expression profile changes associated with dietary energy in the sheep testis during sexual maturation. *Sci. Rep.* **2017**, *7*, 5180. [[CrossRef](#)] [[PubMed](#)]
24. Arun, G.; Akhade, V.S.; Donakonda, S.; Rao, M.R.S. mrhl RNA, a long noncoding RNA, negatively regulates Wnt signaling through its protein partner Ddx5/p68 in mouse spermatogonial cells. *Mol. Cell. Biol.* **2012**, *32*, 3140–3152. [[CrossRef](#)]
25. Anguera, M.C.; Ma, W.; Clift, D.; Namekawa, S.; Kelleher, R.J., III; Lee, J.T. Tsx produces a long noncoding RNA and has general functions in the germline, stem cells, and brain. *PLoS Genet.* **2011**, *7*, e1002248. [[CrossRef](#)] [[PubMed](#)]
26. Agbor, V.A.; Tao, S.; Lei, N.; Heckert, L.L. A Wt1-Dmrt1 transgene restores DMRT1 to Sertoli cells of Dmrt1<sup>-/-</sup> testes: A novel model of DMRT1-deficient germ cells. *Biol. Reprod.* **2013**, *88*. [[CrossRef](#)]
27. Lunstra, D.-D.; Echterkamp, S. Puberty in beef bulls: Acrosome morphology and semen quality in bulls of different breeds. *J. Anim. Sci.* **1982**, *55*, 638–648. [[CrossRef](#)] [[PubMed](#)]
28. Fischer, A.H.; Jacobson, K.A.; Rose, J.; Zeller, R. Hematoxylin and eosin staining of tissue and cell sections. *Cold Spring Harb. Protoc.* **2008**, 2008, pdb-prot4986. [[CrossRef](#)]
29. Parkhomchuk, D.; Borodina, T.; Amstislavskiy, V.; Banaru, M.; Hallen, L.; Krobitch, S.; Lehrach, H.; Soldatov, A. Transcriptome analysis by strand-specific sequencing of complementary DNA. *Nucleic Acids Res.* **2009**, *37*, e123. [[CrossRef](#)]
30. Kim, D.; Pertea, G.; Trapnell, C.; Pimentel, H.; Kelley, R.; Salzberg, S.L. TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **2013**, *14*, R36. [[CrossRef](#)]
31. Langmead, B.; Trapnell, C.; Pop, M.; Salzberg, S.L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **2009**, *10*, R25. [[CrossRef](#)]
32. Trapnell, C.; Pachter, L.; Salzberg, S.L. TopHat: Discovering splice junctions with RNA-Seq. *Bioinformatics* **2009**, *25*, 1105–1111. [[CrossRef](#)]
33. Guttman, M.; Garber, M.; Levin, J.Z.; Donaghey, J.; Robinson, J.; Adiconis, X.; Fan, L.; Koziol, M.J.; Gnirke, A.; Nusbaum, C. Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. *Nat. Biotechnol.* **2010**, *28*, 503–510. [[CrossRef](#)]

34. Trapnell, C.; Williams, B.A.; Pertea, G.; Mortazavi, A.; Kwan, G.; Van Baren, M.J.; Salzberg, S.L.; Wold, B.J.; Pachter, L. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **2010**, *28*, 511–515. [[CrossRef](#)] [[PubMed](#)]
35. Cabili, M.N.; Trapnell, C.; Goff, L.; Koziol, M.; Tazon-Vega, B.; Regev, A.; Rinn, J.L. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.* **2011**, *25*, 1915–1927. [[CrossRef](#)]
36. Harrow, J.; Frankish, A.; Gonzalez, J.M.; Tapanari, E.; Diekhans, M.; Kokocinski, F.; Aken, B.L.; Barrell, D.; Zadissa, A.; Searle, S. GENCODE: The reference human genome annotation for The ENCODE Project. *Genome Res.* **2012**, *22*, 1760–1774. [[CrossRef](#)]
37. Hong, S.-E.; Song, H.K. Identification of tissue-enriched novel transcripts and novel exons in mice. *BMC Genomics* **2014**, *15*, 592. [[CrossRef](#)] [[PubMed](#)]
38. Li, A.; Zhang, J.; Zhou, Z.; Wang, L.; Liu, Y.; Liu, Y. ALDB: A domestic-animal long noncoding RNA database. *PLoS ONE* **2015**, *10*, e0124003. [[CrossRef](#)] [[PubMed](#)]
39. Sun, L.; Luo, H.; Bu, D.; Zhao, G.; Yu, K.; Zhang, C.; Liu, Y.; Chen, R.; Zhao, Y. Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts. *Nucleic Acids Res.* **2013**, *41*, e166. [[CrossRef](#)] [[PubMed](#)]
40. Bateman, A.; Coin, L.; Durbin, R.; Finn, R.D.; Hollich, V.; Griffiths-Jones, S.; Khanna, A.; Marshall, M.; Moxon, S.; Sonnhammer, E.L. The Pfam protein families database. *Nucleic Acids Res.* **2004**, *32*, D138–D141. [[CrossRef](#)] [[PubMed](#)]
41. Finn, R.D.; Bateman, A.; Clements, J.; Coghill, P.; Eberhardt, R.Y.; Eddy, S.R.; Heger, A.; Hetherington, K.; Holm, L.; Mistry, J. Pfam: The protein families database. *Nucleic Acids Res.* **2014**, *42*, D222–D230. [[CrossRef](#)] [[PubMed](#)]
42. Kong, L.; Zhang, Y.; Ye, Z.-Q.; Liu, X.-Q.; Zhao, S.-Q.; Wei, L.; Gao, G. CPC: Assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Res.* **2007**, *35*, W345–W349. [[CrossRef](#)]
43. Frazee, A.C.; Pertea, G.; Jaffe, A.E.; Langmead, B.; Salzberg, S.L.; Leek, J.T. Flexible analysis of transcriptome assemblies with Ballgown. *Biorxiv* **2014**. [[CrossRef](#)]
44. Young, M.D.; Wakefield, M.J.; Smyth, G.K.; Oshlack, A. Gene ontology analysis for RNA-seq: Accounting for selection bias. *Genome Biol.* **2010**, *11*, R14. [[CrossRef](#)]
45. Ashburner, M.; Ball, C.A.; Blake, J.A.; Botstein, D.; Butler, H.; Cherry, J.M.; Davis, A.P.; Dolinski, K.; Dwight, S.S.; Eppig, J.T. Gene ontology: Tool for the unification of biology. *Nat. Genet.* **2000**, *25*, 25–29. [[CrossRef](#)] [[PubMed](#)]
46. Mao, X.; Cai, T.; Olyarchuk, J.G.; Wei, L. Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. *Bioinformatics* **2005**, *21*, 3787–3793. [[CrossRef](#)] [[PubMed](#)]
47. Reuveni, E.; Getselter, D.; Oron, O.; Elliott, E. Differential contribution of cis and trans gene transcription regulatory mechanisms in amygdala and prefrontal cortex and modulation by social stress. *Sci. Rep.* **2018**, *8*, 6339. [[CrossRef](#)]
48. Ghanbarian, A.T.; Hurst, L.D. Neighboring genes show correlated evolution in gene expression. *Mol. Biol. Evol.* **2015**, *32*, 1748–1766. [[CrossRef](#)]
49. Huang, D.W.; Sherman, B.T.; Lempicki, R.A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protocol.* **2009**, *4*, 44–57. [[CrossRef](#)]
50. Livak, K.J.; Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods.* **2001**, *25*, 402–408. [[CrossRef](#)]
51. Veneziano, D.; Nigita, G.; Ferro, A. Computational approaches for the analysis of ncRNA through deep sequencing techniques. *Front. Bioeng. Biotechnol.* **2015**, *3*, 77. [[CrossRef](#)]
52. Zhang, C.; Gao, L.; Xu, E.Y. LncRNA, a new component of expanding RNA-protein regulatory network important for animal sperm development. *Semin. Cell Dev. Biol.* **2016**, *59*, 110–117. [[CrossRef](#)] [[PubMed](#)]
53. Bao, J.; Wu, J.; Schuster, A.S.; Hennig, G.W.; Yan, W. Expression profiling reveals developmentally regulated lncRNA repertoire in the mouse male germline. *Biol. Reprod.* **2013**, *89*, 107-1. [[CrossRef](#)] [[PubMed](#)]
54. Yoneda, R.; Satoh, Y.; Yoshida, I.; Kawamura, S.; Kotani, T.; Kimura, A.P. A genomic region transcribed into a long noncoding RNA interacts with the Prss42/Tessp-2 promoter in spermatocytes during mouse spermatogenesis, and its flanking sequences can function as enhancers. *Mol. Reprod. Dev.* **2016**, *83*, 541–557. [[CrossRef](#)]
55. Derrien, T.; Johnson, R.; Bussotti, G.; Tanzer, A.; Djebali, S.; Tilgner, H.; Guernec, G.; Martin, D.; Merkel, A.; Knowles, D.G. The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression. *Genome Res.* **2012**, *22*, 1775–1789. [[CrossRef](#)] [[PubMed](#)]
56. Hong, S.H.; Kwon, J.T.; Kim, J.; Jeong, J.; Kim, J.; Lee, S.; Cho, C. Profiling of testis-specific long noncoding RNAs in mice. *BMC Genomics* **2018**, *19*, 539. [[CrossRef](#)] [[PubMed](#)]
57. Jia, X.; He, Y.; Chen, S.-Y.; Wang, J.; Hu, S.; Lai, S.-J. Genome-wide identification and characterisation of long non-coding RNAs in two Chinese cattle breeds. *Ital. J. Anim. Sci.* **2020**, *19*, 383–391. [[CrossRef](#)]
58. Liu, Y.; Li, M.; Bo, X.; Li, T.; Ma, L.; Zhai, T.; Huang, T. Systematic analysis of Long non-coding RNAs and mRNAs in the ovaries of Duroc pigs during different follicular stages using RNA sequencing. *Int. J. Mol. Sci.* **2018**, *19*, 1722. [[CrossRef](#)]
59. Zeng, B.; Chen, T.; Xie, M.-Y.; Luo, J.-Y.; He, J.-J.; Xi, Q.-Y.; Sun, J.-J.; Zhang, Y.-L. Exploration of long noncoding RNA in bovine milk exosomes and their stability during digestion in vitro. *J. Dairy Sci.* **2019**, *102*, 6726–6737. [[CrossRef](#)] [[PubMed](#)]
60. Cai, W.; Li, C.; Liu, S.; Zhou, C.; Yin, H.; Song, J.; Zhang, Q.; Zhang, S. Genome wide identification of novel long non-coding RNAs and their potential associations with milk proteins in Chinese Holstein cows. *Front. Genet.* **2018**, *9*, 281. [[CrossRef](#)]

61. Ibeagha-Awemu, E.M.; Do, D.N.; Dudemaine, P.-L.; Fomenky, B.E.; Bissonnette, N. Integration of lncRNA and mRNA transcriptome analyses reveals genes and pathways potentially involved in calf intestinal growth and development during the early weeks of life. *Genes* **2018**, *9*, 142. [[CrossRef](#)] [[PubMed](#)]
62. Ding, H.; Luo, Y.; Liu, M.; Huang, J.; Xu, D. Histological and transcriptome analyses of testes from Duroc and Meishan boars. *Sci. Rep.* **2016**, *6*, 20758. [[CrossRef](#)] [[PubMed](#)]
63. Yang, J.; Hong, Y.; Yuan, C.; Fu, Z.; Shi, Y.; Zhang, M.; Shen, L.; Han, Y.; Zhu, C.; Li, H. Microarray analysis of gene expression profiles of *Schistosoma japonicum* derived from less-susceptible host water buffalo and susceptible host goat. *PLoS ONE* **2013**, *8*, e70367. [[CrossRef](#)]
64. Young, J.C.; Wakitani, S.; Loveland, K.L. TGF- $\beta$  superfamily signaling in testis formation and early male germline development. *Semin. Cell Dev. Biol.* **2015**, *45*, 94–103. [[CrossRef](#)]
65. Li, T.; Zhang, H.; Wang, X.; Yin, D.; Chen, N.; Kang, L.; Zhao, X.; Ma, Y. Cloning, molecular characterization and expression patterns of DMRTC2 implicated in germ cell development of male Tibetan sheep. *Int. J. Mol. Sci.* **2020**, *21*, 2448. [[CrossRef](#)]
66. Yang, K.; Wang, J.; Wang, K.; Luo, Y.; Tang, Q.; Liu, X.; Fang, M. Integrated Analysis of miRNA-mRNA Network Reveals Different Regulatory Patterns in the Endometrium of Meishan and Duroc Sows during Mid-Late Gestation. *Animals* **2020**, *10*, 420. [[CrossRef](#)]
67. Nixon, B.; Bromfield, E.G.; Dun, M.D.; Redgrove, K.A.; McLaughlin, E.A.; Aitken, R.J. The role of the molecular chaperone heat shock protein A2 (HSPA2) in regulating human sperm-egg recognition. *Asian J. Androl.* **2015**, *17*, 568. [[CrossRef](#)]
68. Scieglinska, D.; Krawczyk, Z. Expression, function, and regulation of the testis-enriched heat shock HSPA2 gene in rodents and humans. *Cell Stress Chaperones* **2015**, *20*, 221–235. [[CrossRef](#)] [[PubMed](#)]
69. Gao, Y.; Wu, M.; Fan, Y.; Li, S.; Lai, Z.; Huang, Y.; Lan, X.; Lei, C.; Chen, H.; Dang, R. Identification and characterization of circular RNAs in Qinchuan cattle testis. *R. Soc. Open Sci.* **2018**, *5*, 180413. [[CrossRef](#)]
70. Li, R.-K.; Tan, J.-L.; Chen, L.-T.; Feng, J.-S.; Liang, W.-X.; Guo, X.-J.; Liu, P.; Chen, Z.; Sha, J.-H.; Wang, Y.-F. Iqcg is essential for sperm flagellum formation in mice. *PLoS ONE* **2014**, *9*, e98053. [[CrossRef](#)] [[PubMed](#)]
71. Houmard, B.; Small, C.; Yang, L.; Naluai-Cecchini, T.; Cheng, E.; Hassold, T.; Griswold, M. Global gene expression in the human fetal testis and ovary. *Biol. Reprod.* **2009**, *81*, 438–443. [[CrossRef](#)] [[PubMed](#)]
72. Sambroni, E.; Rolland, A.D.; Lareyre, J.-J.; Le Gac, F. Fsh and Lh have common and distinct effects on gene expression in rainbow trout testis. *J. Mol. Endocrinol.* **2013**, *50*, 1–18. [[CrossRef](#)] [[PubMed](#)]
73. Wu, J.; Wang, F. Genes related with male gonadal morphogenesis in mammals. *Zhonghua Nan Ke Xue Natl. J. Androl.* **2008**, *14*, 356–359.
74. Ponting, C.P.; Oliver, P.L.; Reik, W. Evolution and functions of long noncoding RNAs. *Cell* **2009**, *136*, 629–641. [[CrossRef](#)]
75. Engreitz, J.M.; Haines, J.E.; Perez, E.M.; Munson, G.; Chen, J.; Kane, M.; McDonel, P.E.; Guttman, M.; Lander, E.S. Local regulation of gene expression by lncRNA promoters, transcription and splicing. *Nature* **2016**, *539*, 452–455. [[CrossRef](#)]
76. Wu, Q.; Song, R.; Yan, W. *SPATA3 and SPATA6 Interact with KLHL10 and Participate in Spermatogenesis*; Oxford University Press: Oxford, UK, 2010.