Transcriptome profiling of developing testes and spermatogenesis in the Mongolian horse

Bei Li

College of Animal Science <https://orcid.org/0000-0002-2841-7700>

Xiaolong He

Inner Mongolia Academy of Agricultural and Animal Husbandry Sciences

Yiping Zhao

Inner Mongolia Agricultural University

Dongyi Bai

Inner Mongolia Agricultural University

Ming Du

Inner Mongolia Agricultural University

Lianjie Song

Inner Mongolia Agricultural University

Zhuang Liu

Inner Mongolia Agricultural University

Zhenchen Yin

Inner Mongolia Agricultural University

Dugarjaviin Manglai (✉ dmanglai@163.com)

<https://orcid.org/0000-0002-3152-8137>

Research article

Keywords: Mongolian horse, developing testis, spermatogenesis, RNA-Seq

Posted Date: January 8th, 2020

DOI: <https://doi.org/10.21203/rs.2.16493/v2>

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

[Read Full License](#)

Version of Record: A version of this preprint was published at BMC Genetics on April 28th, 2020. See the published version at <https://doi.org/10.1186/s12863-020-00843-5>.

Abstract

Background: Horse testis development and spermatogenesis are complex physiological processes. **Methods:** To study these processes, three immature and three mature testes were collected from the Mongolian horse, and six libraries were established using high-throughput RNA sequencing technology (RNA-Seq) to screen for genes related to testis development and spermatogenesis. **Results:** A total of 16,237 upregulated genes and 8,641 downregulated genes were detected in the testis of the Mongolian horse. These genes play important roles in different developmental stages of spermatogenesis and testicular development. Five genes with alternative splicing events that may influence spermatogenesis and development of the testis were detected. GO (Gene ontology) and KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway analyses were performed for functional annotation of the differentially expressed genes. Pathways related to “spermatogenesis,” male gamete generation,” “spermatid development” and “oocyte meiosis” were significantly involved in different stages of testis development and spermatogenesis. **Conclusion:** Genes, pathways and alternative splicing events were identified with inferred functions in the process of spermatogenesis in the Mongolian horse. The identification of these differentially expressed genetic signatures improves our understanding of horse testis development and spermatogenesis.

Background

The mammalian testis is an important male reproductive organ that produces sperm and androgens. Spermatogenesis is a complex developmental process with a unique mechanism of cell division—meiosis—as a major defining event. Sperm development consists of three stages: (1) proliferation and differentiation of spermatogonia, (2) meiosis of spermatocytes, and (3) sperm maturation. Spermatogonial stem cells differentiate into type B spermatogonia and, with replication of their DNA content, give rise to primary spermatocytes. Secondary spermatocytes and spermatids are derived from the primary spermatocytes by DNA replication and tetraploid meiosis. Haploid sperm cells undergo at least four distinct morphological changes into sperm cells, namely chromatin condensation, acrosome formation, flagellar formation, and cytoplasmic reduction. In addition to spermatogenic cells, spermatogenesis also involves multiple somatic cells in testicular tissues, such as Sertoli and Leydig cells. In mammalian testes, spermatogenesis occurs within seminiferous tubules, where germ cells associate with Sertoli cells, which provide the appropriate microenvironment for spermatogenesis. Related genes in these cells play important roles in specific stages of germ cell development, and the transcription and translation of these genes vary in different stages of spermatogenesis [1, 2].

RNA sequencing technology (RNA-Seq) technology, which provides a global view of gene expression profiles, has been developed for mapping and quantifying transcriptomes [3, 4]. This methodology has several advantages over other transcriptome technologies, such as higher resolution and sensitivity, a large dynamic range of gene expression levels, and the ability to identify novel transcribed regions and splicing isoforms of known genes [5]. Ramsköld [6] assessed different mammalian tissues by RNA-Seq and found that the vast majority of genes are specifically expressed in the testis. Furthermore, Djureinovic

[7] categorized 20,050 putative human genes based on RNA-Seq expression patterns. These genes were specifically expressed in normal human testicular tissues compared to 26 other normal human tissue types in seven individuals. The analysis showed that the testis has the highest number of tissue-specific genes by far. As an alternate, less sensitive methodology, Anand [8] identified differentially expressed genes (DEGs) in the human testis relative to other tissues by microarray analysis, and a total of 2,868 upregulated and 2,011 downregulated transcripts were identified. The testis thus seems to be highly metabolically active relative to other normal tissues as indicated by functional annotation. Currently, most of the studies on the relationship between the testis development and spermatogenesis have focused on humans or mice. However, little is known regarding the spermatogenesis process in the horse.

The Mongolian horse is one of the most famous native breeds in China, and in the world. It is a dual-purpose breed that can adapt to harsh climate and extensive breeding conditions. Normal development of the testis and spermatogenesis is extremely important to ensure high level semen production and protect population continuity, and genes that are important for different stages of testis development and spermatogenesis are likely to have roles in maintaining fertility. Unfortunately, however, the transcriptional expression of genes during testicular development and spermatogenesis in the horse is not well understood, thereby limiting the understanding of the mechanisms of deformation and maturation in spermatogenesis. This study aimed to study the transcriptome profiles of immature and mature male Mongolian horse testes utilizing high-throughput RNA-Seq technology and bioinformatics analysis. For this purpose, a testis transcript database was obtained, and DEGs and regulatory pathways were analyzed. This study provides new insights into gene expression and the regulatory mechanisms of horse testis development and spermatogenesis.

Materials And Methods

Ethics statement

This study was approved by and conducted according to the guidelines of the Institutional Animal Ethics Committee and Animal Care Guidelines of the Inner Mongolia Agricultural University. The animals did not unnecessarily suffer at any stage of this study.

Animals and sample collection

We received permission to geld six healthy male Mongolian horses in Xilingol League, Inner Mongolia, China. The ages of the horses were determined based on physical examination of their teeth, as well as information from the owner. Three were sexually immature colts (before sexual maturation; samples BSM1–3), between 11 and 13 months old. The other three were sexually mature stallions, between three and four years old (after sexual maturation; samples ASM1–3). We surgically collected the testes of all six horses. The excised testes were stored in cryogenic vials with RNA/DNA sample protector (Takara). Small aliquots (~1 mg) from the testes of each animal from the parenchyma, including the semiferous tubules and Leydig cells, were immediately frozen in liquid nitrogen for quantitative real-time polymerase chain reaction (qRT-PCR) analysis. All horses were reared by their owner after our study.

RNA quantification and qualification

Total RNA was extracted using RNeasy Mini Kits (Invitrogen) according to the manufacturer's protocols and resuspended in nuclease-free water. Then, the samples were treated with DNase I (Invitrogen) to eliminate any contaminating genomic DNA. RNA degradation and contamination were monitored on 1% agarose gels. RNA purity was assessed using the NanoPhotometer® spectrophotometer (IMPLEN). RNA concentrations were measured using the Qubit® RNA Assay Kit with the Qubit® 2.0 Fluorometer (Life Technologies). RNA integrity was assessed using the RNA Nano 6000 Assay Kit with the Bioanalyzer 2100 system (Agilent Technologies).

Library preparation for strand-specific transcriptome sequencing

Six libraries were constructed. The three immature testes libraries were designated BSM1, BSM2, and BSM3; and the three mature testes libraries were designated ASM1, ASM2, and ASM3. Approximately 3 µg RNA per sample were used as input material for the RNA sample preparations. Sequencing libraries were generated using NEBNext® Ultra™ Directional RNA Library Prep Kit for Illumina® (NEB) according to the manufacturer's recommendations, and index codes were added to attribute sequences to each sample. Briefly, mRNA was purified from total RNA using poly-T oligo-conjugated magnetic beads. Fragmentation was conducted using divalent cations under elevated temperature in NEBNext First Strand Synthesis Reaction Buffer (5X). First-strand cDNA was synthesized using random hexamer primer and M-MuLV reverse transcriptase (RNaseH-). Second-strand cDNA synthesis was subsequently performed using DNA polymerase I and RNase H. In the reaction buffer, dNTPs with dTTP were replaced by dUTP. Remaining overhangs were converted into blunt ends via exonuclease/polymerase activity. After adenylation of the 3' ends of the DNA fragments, NEBNext Adaptors with hairpin loop structure were ligated for hybridization. To select cDNA fragments of preferentially 150~200 bp in length, the libraries were purified with the AMPure XP system (Beckman Coulter). Then 3 µL of USER Enzyme (NEB) were used with size-selected, adaptor-ligated cDNA at 37°C for 15 min, followed by 5 min at 95°C before PCR. PCR was performed with Phusion High-Fidelity DNA polymerase, Universal PCR primers, and Index (X) Primer. Finally, the products were purified (AMPure XP system), and the library quality was assessed using the Agilent Bioanalyzer 2100 system.

Clustering and sequencing

Clustering of the index-coded samples was performed on a cBot Cluster Generation System using the TruSeq PE Cluster Kit v3-cBot-HS (Illumina) according to the manufacturer's instructions. After cluster generation, the library preparations were sequenced on an Illumina HiSeq platform and 150 bp paired-end reads were generated.

Quality control

In order to guarantee data quality in our analysis, a Perl script was used to filter the original data (Raw Data). The steps were as follows: 1) the Smart-seq2 public primer sequence was trimmed from the

Reads; 2) the contaminated reads were separated from the adapters; 3) the low quality reads were removed (defined by > 20% Read Bases with phred Quality value of ≤ 19 at either end); 4) the reads with N base > 5% were removed.

Read mapping to the reference genome

The reference genome and gene model annotation files were directly downloaded from the genome website. The index of the reference genome was built using Bowtie v2.2.3, and paired-end clean reads were aligned to the reference genome using TopHat v2.0.12 [9, 10]. TopHat was selected as the mapping tool because it can generate a database of splice junctions based on the gene model annotation file and thus provides a better mapping result than other non-splice mapping tools [11].

Quantification of gene expression

HTSeq v0.6.1 was used to count the read numbers mapped to each gene [12]. Then, the expected number of Fragments Per Kilobase of transcript sequence per Million base pairs sequenced (FPKM) of each gene was calculated based on the length of the gene and read count mapped to the gene. The FPKM simultaneously considers the effect of the sequencing depth and gene length for the read count, and is currently the most commonly used method for estimating gene expression levels [13].

Differential expression analysis

Differential expression analysis of two conditions/groups (two biological replicates per condition) was performed using the DESeq R package (1.18.0). DESeq provides statistical routines for determining differential expression within digital gene expression data using a model based on the negative binomial distribution. The resulting P-values were adjusted using the Benjamin and Hochberg's approach for controlling the false discovery rate. Genes with an adjusted P-value < 0.05 found by DESeq were assigned as differentially expressed [14, 15].

Gene ontology and KEGG enrichment analyses of DEGs

Gene ontology (GO) enrichment analysis of DEGs was implemented by the Goseq R package, in which the gene length bias was corrected. GO terms with corrected P-value < 0.05 were considered significantly enriched DEGs [16, 17]. Kyoto Encyclopedia of Genes and Genomes (KEGG) is a database resource for understanding high-level functions and utilities of biological systems, such as the cell, the organism and the ecosystem, based on molecular-level information, especially large-scale molecular datasets generated by genome sequencing and other high-throughput experimental technologies (<http://www.genome.jp/kegg/>). We used KOBAS software to test the statistical enrichment of DEGs in the KEGG pathways [18].

Novel transcript prediction and alternative splicing analysis

The Cufflinks v2.1.1 Reference Annotation Based Transcript (RABT) assembly method was used to construct and identify both known and novel transcripts from TopHat alignment results. Alternative splicing (AS) events were classified into 12 basic types by the software Asprofile v1.0. The number of AS events in each sample was estimated separately.

PPI analysis of DEGs

Protein-protein interaction (PPI) analysis of DEGs was based on the STRING database, which includes known and predicted interactions. For species represented in the database, we constructed networks by extracting the target gene list from the database. Otherwise, Blastx (v2.2.28) was used to align the target gene sequences to the selected reference protein sequences, and then the network was built according to the known interaction of the selected reference species.

Quantitative real time PCR (qRT-PCR) validation

A piece of testis was excised and frozen immediately in liquid nitrogen for qRT-PCR analysis. Total RNA was extracted using TRIzol (TaKaRa) according to the manufacturer's protocols. The RNA was resuspended in nuclease-free water, and the concentration was measured using NanoDrop (Thermo Scientific Nanodrop 2000). About 0.5 µg of total RNA was used as template to synthesize first-strand cDNA using a PrimerScript RT reagent kit (TaKaRa) according to the manufacturer's protocols. The resultant cDNA was diluted to 0.1 µg/µL for further analysis by qRT-PCR (Bio-Rad) using SYBR Green Realtime PCR Master Mix (TaKaRa). β-Actin was chosen as an internal reference gene to eliminate sample-to-sample variations based on evidence of its role as a constitutive house-keeping gene for qRT-PCR and its use in normalizing changes in specific gene expressions[19]. The primers of tested genes were designed by Primer 5.0 according to alternative splicing events of those genes (Table S8). The relative gene expression levels were calculated using the $2^{-\Delta\Delta C_t}$ method. The DEGs between ASM and BSM testis were analyzed by ANOVA using SAS software 9.0.

Results

mRNA expression profiles in Mongolian horse testis development

To screen for genes related to testis development and spermiogenesis, a total of six libraries—three from immature colts (BSM1, BSM2, and BSM3) and three from mature stallions (ASM1, ASM2, and ASM3)—were constructed and sequenced on the Illumina Hiseq platform. An overview of the information for these libraries is listed in Table 1. The libraries were found to contain 58,277,598 (BSM1), 59,921,568 (BSM2), 53,371,478 (BSM3), 59,523,578 (ASM1), 61,727,216 (ASM2), and 53,956,064 (ASM3) raw reads. To ensure the quality of subsequent analysis, the adapter reads with ratio > 10% or with incomplete base information and low-quality reads were removed; the results were 57,482,858 (BSM1), 58,972,988 (BSM2), 52,475,320 (BSM3), 58,411,056 (ASM1), 60,775,868 (ASM2), and 53,135,736 (ASM3) clean reads. The error rate and GC content of each library were calculated to control for the quality of libraries (Table 1). These results verify that the six libraries were of high quality.

The clean reads were aligned to the reference genome of *E. caballus*, which is the most accurate and commonly used reference genome for horses[20, 21], using TopHat v2.0.12 (Table 2). In total, 84.85% (BSM1), 85.05% (BSM2), 84.25% (BSM3), 84.38% (ASM1), 84.14% (ASM2), and 84.23% ASM3) clean reads were uniquely mapped to the reference genome. In addition, more than 40% of clean reads were non-splice reads in each sample. Additionally, 27.98% (BSM1), 27.35% (BSM2), 26.92% (BSM3), 36.17% (ASM1), 34.89% (ASM2), and 34.74% (ASM3) of the clean reads were mapped to the borders of exons (also called “junction reads”). All RNA-Seq data have been submitted to The Gene Expression Omnibus, accession number GEO: GSE101697).

Analysis of alternative splicing

Alternative splicing (AS) is a common occurrence in most eukaryotic cells and can result in the translation of different forms of proteins at various time points and environments, which increases the adaptability and physical fitness of the species. In our study, the AS events were classified into five kinds by rMATS (<http://rnaseq-mats.sourceforge.net/index.html>): (1) SE: Skipped exon, (2) MXE: Mutually exclusive exon, (3) A5SS: Alternative 5' splice site, (4) A3SS: Alternative 3' splice site, and (5) RI: Retained intron. The kinds and quantities of AS events were counted, and the expression of each type of AS event was analyzed. The results indicate that numerous and varied AS events exist in the spermatogenesis process Table 3.

Analysis of DEGs

The DEGs between libraries were screened based on DEGSeq analysis, taking $|\log_2(\text{fold-change})| > 1$ and $P < 0.05$ as the cut-off. By considering ad-joining libraries, 16,237 up- and 8,641 downregulated genes were identified in total that were up-regulated in ASM versus BSM (Table S1, Fig. 1a). Venn diagrams were constructed to summarize the data, which show that the ASM and BSM libraries shared 16,327 genes that were expressed in both BSM and ASM (Fig. 1b). Figure 1c shows the results of hierarchical clustering, with the DEGs of the six libraries divided into two clusters. These results show that the ASM and BSM libraries had overall distinct patterns of differential expression but that the replicates within each developmental stage were similar.

GO pathway enrichment analysis of DEGs

GO pathway enrichment analysis was used to explore the functions of the DEGs in testis development. A total of 709 GO terms related to “biological processes”, “molecular functions”, and “cellular components” were significantly enriched across the ASM vs BSM. DEG groups (Table S2, upregulated GO terms; Table S3, downregulated GO terms, Fig. 2a). Approximately 510 GO terms belong to “biological processes”, and these were mainly focused on “spermatogenesis” (36 genes), “male gamete generation” (36 genes), “spermatid development” (18 genes), “metabolic processes” (2,454 genes), “organic substance metabolic process” (2,044 genes), and “cellular metabolic process” (1,948 genes). Around 84 GO terms belong to the functional category of “cellular component”, and these were mainly concentrated in “intracellular” (1,951 genes), “organelle” (1,844 genes), and “intracellular part” (1,668 genes) classification groups.

Approximately 115 GO terms belong to “molecular function”, which centered on “binding” (3,037), “protein binding” (1,719), and “catalytic activity” (1,659 genes) (Fig. 2b). The differential up- and down-regulation may due to the gene ontology associations with spermatogenesis or, conversely, spermatogonia development. Further research will aid in clarifying this discrepancy.

KEGG pathway enrichment analysis of DEGs

KEGG pathway enrichment analysis was performed on these differently expressed genes with a cut-off criterion of corrected P-values < 0.05. A total of 10 enriched pathways were detected (Table S4, upregulated KEGG pathways; Table S5). Six KEGG pathways were upregulated, including “ubiquitin-mediated proteolysis”, “protein processing in endoplasmic reticulum”, “RNA transport”, “metabolic pathways”, “oocyte meiosis”, and “cell cycle”. Four pathways were downregulated, including “ribosome”, “focal adhesion”, “proteoglycans in cancer”, and “ECM-receptor interactions” (Fig. 3).

Differential gene transcription factor analysis

Putative transcription factors were analyzed using the animal transcription factor database TFDB 2.0. For genes included in the database, the transcription factors were directly screened by Ensembl geneID; for genes not contained in database, the transcription factors were identified by hmmsearch according to transcription factor family Pfam files. Non-Ensembl geneID genes were screened through the known species transcription factor protein sequence database by BLASTX. In our study, 917 known gene transcription factors and 402 unknown transcription factors were detected, belonging to 57 gene families (Tables S6 and S7).

Quantitative real time PCR (qRT-PCR) validation of differential expression

To verify the differential expression of genes in ASM vs. BSM testis, we selected eight DEGs, including *MEI1* (two different sites), *EHMT2*, *BTRC*, *CLOCK*, *IQCG*, *ANAPC5* and *SKIP*, to validate the expression patterns by qRT-PCR. These genes belong to “oocyte meiosis” KEGG pathways and GO terms “spermatogenesis”, “male gamete generation”, and “spermatid development” (Table S8). There was a relatively high correlation between the mRNA expression levels of the eight genes, as detected by qRT-PCR and RNA-Seq (correlation coefficient = 0.853), and all 9 genes were upregulated in ASM. This suggests that our RNA-Seq results are reliable.

Discussion

Currently, with the development of improved detection techniques, a growing number of sperm-specific mRNAs have been identified. RNA-Seq technology has become an efficient and inexpensive method for the discovery of novel transcripts and genes. In our study, three ASM testes and three BSM testes were used as research samples, and 24,878 genes were screened by RNA-Seq. The results led to the identification of 16,236 genes that were upregulated and 8,641 genes that were downregulated during development. Using the next-generation platform, we found that most of the highly expressed genes with

p_{adj} values of < 0.05 are differentiation-related with identifiable roles in the development of the testis and spermatogenesis. The *SOX8*, *SOX9*, *SOX11*, and *SOX13* genes encode members of the SOX (SRY-related HMG-box) family of transcription factors that are involved in the regulation of embryonic development and in the determination of cell fate. The encoded proteins act as transcriptional activators that form complexes with other proteins [22]. However, in our study, the *SOX5*, *SOX6*, *SOX10* and *SOX30* genes were upregulated in ASM. Every family member may have its own role in determining testis-related functions, and thus, further investigations are warranted. The testicular hormone *Insl3* (insulin-like 3), which is highly expressed during BMS, plays a crucial role in mouse gubernacular development and testis descent [23]. Our analysis also identified related genes that were upregulated in ASM, such as *SPATA3*, *SPATA45*, *SPATC1*, *SPATA18*, *SPEM1*, *SPERT*, *SHBG*, *TEX35*, *TEX43*, *GSG1*, and *SPATA19*. Among these genes, *SPATA3*, *SPATA45*, *SPATC1*, *SPATA18*, and *SPATA19* have known roles in spermatogenesis; *TEX35* and *TEX43* are associated with testis expression; *GSG1* is related to germ cell function; *SPEM* influences spermatid maturation; and *SHBG* controls sex hormone-binding globulin. In addition, most of the novel genes detected are likely to be related to spermatogenesis.

Alternative splicing (AS) is a crucial mechanism for regulating gene expression and generating proteomic diversity. Recent estimates indicate that the expression of nearly 95% of human multi-exon genes involves AS [24, 25]. In metazoans, AS plays an important part in generating different protein products that function in diverse cellular processes, including cell growth, differentiation, and death [26]. In our study, five kinds of AS events were detected, most of them involving skipped exons (SEs). The SE events and the results of GO and KEGG analyses were comprehensively analyzed to determine the influence of AS events on the function of related genes. In our study, IQ motif-containing G (*IQCG*) is regulated by SE events as categorized under the GO terms “spermatogenesis”, “male gamete generation”, and “spermatic development”. The causality of the mutation was confirmed with a targeted null allele. Loss of *IQCG* disrupts spermiogenesis wherein tail formation is either incomplete or breaks apart from the sperm heads. Because IQ motif-containing genes typically regulate calmodulin, which in turn can impact the actin cytoskeleton, these findings suggest a potential role for localized calcium signaling in sperm flagellum morphogenesis, Ca^{2+} signaling is important for regulating sperm motility, most notably in the context of flagellar function and regulation [6, 27-32]. ES restores the reading frame within a gene. In our study, *IQCG* incurred SE events from 33,961,411 bp to 33,961,540 in ASM, whereas *IQCG* was upregulated in ASM. SE event also occurred from 38,149,431 to 38,149,546 of the meiotic double-stranded break formation protein 1 (*Mei1*) gene in ASM. *Mei1* was upregulated in ASM and is related to the GO terms “spermatogenesis”, “male gamete generation”, and “sperm development”. *MEI1* is expressed specifically in the testis and the first meiosis-specific mutation identified by forward genetic approaches in mammals. Mutations in the *MEI1* gene disrupt meiosis during spermatogenesis, especially in Americans of European descent [33-35].

Other enriched genes in this study that belong to the “spermatogenesis”, “male gamete generation”, and “spermatid development” GO terms include *SLC26A8*, *MEI1*, *EHMT2*, *DPY19L2*, *HIST1H2BA*, *CAPZA3*, *TSSK2*, *TSSK1B*, *KDM3A*, *ALMS1*, and *IQCG*. Testis anion transporter 1 is a protein that, in humans, is

encoded by the *SLC26A8* gene [36, 37]. This gene is expressed primarily in spermatocytes. Two transcript variants encoding different isoforms have been reported [38-40]. It acts as a DIDS-sensitive anion exchanger that mediates chloride, sulfate, and oxalate transport. It also has critical anion exchange functions in male germ lines during meiosis, and therefore, may play a role in spermatogenesis [41]. The DPY19L2 protein, which belongs to the dpy-19 family, is highly expressed in testis and is required for sperm head elongation and acrosome formation during spermatogenesis. Mutations in *DPY19L2* are associated with an infertility disorder known as spermatogenic failure type 9 (SPGF9) [42-48]. TSSK2 (testis specific serine kinase 2), and TSSK1B (testis-specific serine/threonine-protein kinase 1) are required during spermatid development. *DPY19L2* is also involved in the late stages of spermatogenesis and the reconstruction of the cytoplasm. During spermatogenesis, *TSSK2* is highly expressed in the testis and is required for the transformation of a ring-shaped structure around the base of the flagellum that originates from the chromatoid body [49-54]. The above genes have well-established functions in spermatogenesis; however, the regulatory mechanisms in the horse have not been verified and thus require further investigation.

In our study, six upregulated KEGG pathways were detected. Though the “oocyte meiosis” pathway is named according to its role in female fertility, the process of meiosis in spermatogenesis has obvious similarities. Meiosis plays a critical role in the formation of sperm. In males, meiosis occurs during spermatogenesis in the seminiferous tubules of the testicles and is specific spermatocytes, which later mature into spermatozoa. A total of 47 genes in the oocyte meiosis pathway were detected, with functions that influence meiosis I and II. *PKMYT1*, *CDC25C*, *CCNB2*, and *FBX043* are upregulated in meiosis I, with a central position for *CDC25C*. *CDC25C* (M-phase inducer phosphatase 3) is an enzyme that is highly conserved in evolution and plays a key role in the regulation of cell division. On the other hand, *BUB1*, *MAD2L1*, *SGOL1*, *PLK1*, *CCNB1*, *CCNE1*, *CCNE2*, *PPP2R5E*, *PPP2R5C*, *PPP2CB*, *PPP2R1B*, *PLCZ1*, *ITPR1*, *CALM2*, and *CAMK2D* are upregulated in meiosis II. The G2/mitotic-specific cyclin-B1 (CCNB) is encoded by the *CCNB1* gene in humans [55]. The CCNE protein belongs to the highly conserved cyclin family, whose activity is required for cell cycle G1/S transition. This protein accumulates during the G1-S phase and is degraded as cells progress through the S phase [56]. *PLK1* is an early trigger for G2/M transition and supports the functional maturation of the centrosome in late G2/early prophase. *PLK1* phosphorylates and activates components of the anaphase-promoting complex (APC). The function of APC can maintain cohesion of sister chromatids and anaphase inhibitors. [57]. There is evidence that *PLK1* might function during human meiosis S [58]. These genes are affected the G1-S and G2/M phase in meiosis II. However, alternate transcription initiation sites exist in these genes. Further investigations may reveal changes in the function of these genes caused by AS.

The *BTRC*, *CUL1*, *AURKA*, *FBX043*, *ANAPC11*, *CDC27*, *CDC20*, *ANAPC5*, *ESPL1*, *SMX1B*, and *REC8* genes influence both meioses I and II. Aurora kinase a, also known as serine/threonine-protein kinase 6, is an enzyme that, in humans, is encoded by the *AURKA* gene and is a member of a family of mitotic serine/threonine kinases [59, 60]. This gene is implicated to have an essential role during mitosis and meiosis [61], and its activity peaks during the G2 phase to M phase transition in the cell cycle [62]. Meiotic

recombination protein REC8 homolog is a meiosis-specific component of the [cohesin](#) complex that binds sister chromatids in preparation for the two divisions of meiosis.[63, 64].

Our study focused on increasing understanding of the process of testis development and spermatogenesis in the Mongolian horse, which ultimately may lead to the development of approaches to improve fertility. Due to the lack of information available in regards to the genetics of horse spermatogenesis, we have inferred the function of certain DEGs according to available information in humans and mice. Therefore, the application of the results to situations of infertility in the horse, such as azoospermia and asthenozoospermia, necessitates further species-specific validation of gene functions.

Conclusions

In our study, six mRNA libraries were constructed and sequenced from horse testes from ASM and BSM. A total of 16,237 upregulated and 8,641 downregulated genes were identified; most genes had functions within biological process categories and are known to participate in spermatogenesis and testis development. We analyzed a large number of genes related to spermatogenesis and somatic development of the horse testis, including those that play important roles in different developmental stages. We also found that the AS events might influence the function of these genes and indirectly affect the spermatogenesis and development of the horse testis. These findings will help us to further understand how gene expression is regulated during testis development and spermatogenesis. Future research to evaluate the influence of AS events on testis development and spermatogenesis-related genes, or studies that combine other experimental results, such as miRNA and piRNA expression patterns, will provide more fundamental information in understanding the regulatory mechanisms of equine testis development and spermatogenesis at the molecular level.

Declarations

Ethics approval and consent to participate

All procedures involving animals were approved and authorized by the Inner Mongolia Agricultural University. The animals used in our study were consented by their owners.

Availability of data and material

The accession number for the raw and processed profiling by high-throughput RNA sequencing technology reported in this paper is GEO: GSE101697.

Consent for publication

Not applicable

Competing interests

The authors declare no competing interests.

Authors' contributions

BL and XH conceived the research. DM and DB collected the testicular samples of the Mongolian horse. BL, YZ, and XH performed RNA extraction, MD and LS conducted sequence preprocessing and functional annotation, ZL and ZY performed qRT-PCR validation. BL and DM wrote the paper. All authors have read and approved the final manuscript.

Funding

This study was supported by the Youth Fund of College of Animal Science, Inner Mongolia Agricultural University (QN201907,BZCG201901), the Natural Science Foundation of Inner Mongolia (2015BS0314, 2017ZD06 and 2019ZD03), National Natural Science Foundation of China (3191101008, 31960657 and 31860636), People's Republic of China Specific Scientific and Technological Cooperation (2017YFE0108700) and Special Fund for Transformation of Scientific and Technological Achievements for Breeding of New Animals and Plants in Inner Mongolia Agricultural University (YZGC2017001). Funding bodies had no influence on the design of the study, collection of the data, analysis, interpretation of the data or the writing of the manuscript.

Acknowledgment

Not Applicable

References

1. Hecht NB: **Molecular mechanisms of male germ cell differentiation.** *BioEssays* 1998, **20**(7):555-561.
2. EM E, DA. OB: **Gene expression during mammalian meiosis.** *Curr Top Dev Biol* 1998, **37**:141-200.
3. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B: **Mapping and quantifying mammalian transcriptomes by RNA-Seq.** *Nat Methods* 2008, **5**(7):621-628.
4. Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y: **RNA-seq: An assessment of technical reproducibility and comparison with gene expression arrays.** *Genome Research* 2008, **18**(9):1509-1517.
5. Wang Z, Gerstein M, Snyder M: **RNA-Seq: a revolutionary tool for transcriptomics.** *Nature Reviews Genetics* 2009, **10**(1):57-63.
6. Ramsköld D, Wang ET, Burge CB, Sandberg R: **An Abundance of Ubiquitously Expressed Genes Revealed by Tissue Transcriptome Sequence Data.** *PLoS Computational Biology* 2009, **5**(12):e1000598.
7. Djureinovic D, Fagerberg L, Hallström B, Danielsson A, Lindskog C, Uhlén M, Pontén F: **The human testis-specific proteome defined by transcriptomics and antibody-based profiling.** *Mol Hum Reprod* 2014 **29**(6):476-488.

8. Anand M PB: **The computational analysis of human testis transcriptome reveals closer ties to pluripotency.** *J Hum Reprod Sci* 2012, **5**(3):266-273.
9. Langmead B, Trapnell C, Pop M, Salzberg SL: **Ultrafast and memory-efficient alignment of short DNA sequences to the human genome.** *Genome Biology* 2009, **10**(3):R25.
10. Langmead B, Salzberg SL: **Fast gapped-read alignment with Bowtie 2.** *Nature Methods* 2012, **9**(4):357-359.
11. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky 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 Research* 2010, **20**(9):1297-1303.
12. Anders S, Pyl PT, Huber W: **HTSeq—a Python framework to work with high-throughput sequencing data** *Bioinformatics* 2015, **31**(2):166-169.
13. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L: **Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation.** *Nature Biotechnology* 2010, **28**(5):511-515.
14. Anders S, Huber W: **Differential expression analysis for sequence count data.** *Genome Biology* 2010, **11**(10):R106.
15. Wang L, Feng Z, Wang X, Wang X, Zhang X: **DEGseq: an R package for identifying differentially expressed genes from RNA-seq data.** *Bioinformatics* 2010 **26**(1):136-138.
16. Young MD, Wakefield MJ, Smyth GK, Oshlack A: **Gene ontology analysis for RNA-seq: accounting for selection bias.** *Genome Biology* 2010, **11**(2):R14.
17. 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 Research* 2007, **36**(Database):D480-D484.
18. Mao X, Cai T, Olyarchuk JG, Wei L: **Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary.** *Bioinformatics* 2005, **21**(19):3787-3793.
19. Mori R, Wang Q, Danenberg K, Pinski J, Danenberg P: **Both beta-actin and GAPDH are useful reference genes for normalization of quantitative RT-PCR in human FFPE tissue samples of prostate cancer.** *The Prostate* 2008, **68**:1555-1560.
20. Jun J, Cho Y, Hu H, Kim H-M, Jho S, Gadhvi P, Park K, Lim J, Paek W, Han K *et al*: **Whole genome sequence and analysis of the Marwari horse breed and its genetic origin.** *BMC Genomics* 2014, **15**:S4.
21. Doan R, Cohen N, Sawyer J, Ghaffari N, Johnson C, Dindot S: **Whole-Genome Sequencing and Genetic Variant Analysis of a Quarter Horse Mare.** *BMC genomics* 2012, **13**:78.
22. Pfeifer D, Poulat F, Holinski-Feder E, Kooy F, Scherer G: **The SOX8 gene is located within 700 kb of the tip of chromosome 16p and is deleted in a patient with ATR-16 syndrome.** *Genomics* 2000, **63**(1):108-116.

23. Sozubir S, Barber T, Wang Y, Ahn C, Zhang S, Verma S, Lonergan D, Lorenzo AJ, Nef S, Baker LA: **Loss of *Ins13*: a potential predisposing factor for testicular torsion.** *J Urol* 2010, **183**(6):2373-2379.
24. DL B: **Mechanisms of alternative pre-messenger RNA splicing.** *Annu Rev Biochem* 2003, **72**:291-336.
25. Wahl MC WC, Lührmann R.: **The spliceosome: design principles of a dynamic RNP machine.** *Cell* 2009 **136**(4):701-718.
26. Chen M MJ: **Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches.** *Nat Rev Mol Cell Biol* 2009, **10**(11):741-754.
27. Harris TP, Schimenti KJ, Munroe RJ, Schimenti JC: **IQ Motif-Containing G (*Iqcg*) Is Required for Mouse Spermiogenesis.** *G3: Genes/Genomes/Genetics* 2014, **4**(2):367-372.
28. Becky Marquez GI, and Susan S. Suarez: **Contributions of Extracellular and Intracellular Ca²⁺ to Regulation of Sperm Motility: Release of Intracellular Stores Can Hyperactivate CatSper1 and CatSper2 Null Sperm.** *Dev Biol* 2007, **303**(1):214-221.
29. Lindemann CB, Goltz JS, Kanous KS: **Regulation of activation state and flagellar wave form in epididymal rat sperm: evidence for the involvement of both Ca²⁺ and cAMP.** *Cell Motil Cytoskeleton* 1987, **8**(4):24-32.
30. Igotz GG, Suarez SS: **Calcium/calmodulin and calmodulin kinase II stimulate hyperactivation in demembranated bovine sperm.** *Biol Reprod* 2005, **73**(3):519-526.
31. Ho HC GK, Suarez SS.: **Hyperactivated motility of bull sperm is triggered at the axoneme by Ca²⁺ and not cAMP.** *Dev Biol* 2002 **250**(1):208-217.
32. Darszon A NT, Beltran C, Treviño CL.: **Calcium channels in the development, maturation, and function of spermatozoa.** *Physiol Rev* 2011, **91**(4):1305-1355.
33. Sato H, Miyamoto T, Yogev L, Namiki M, Koh E, Hayashi H, Sasaki Y, Ishikawa M, Lamb DJ, Matsumoto N *et al*: **Polymorphic alleles of the human *MEI1* gene are associated with human azoospermia by meiotic arrest.** *J Hum Genet* 2006, **51**(6):533-540.
34. Reinholdt LG, Schimenti JC: ***Mei1* is epistatic to *Dmc1* during mouse meiosis.** *Chromosoma* 2005, **114**(2):127-134.
35. Libby BJ, Reinholdt LG, Schimenti JC: **Positional cloning and characterization of *Mei1*, a vertebrate-specific gene required for normal meiotic chromosome synapsis in mice.** *Proceedings of the National Academy of Sciences* 2003, **100**(26):15706-15711.
36. Lohi H, Kujala M, Mäkelä S, Lehtonen E, Kestilä M, Saarialho-Kere U, Markovich D, Kere J: **Functional Characterization of Three Novel Tissue-specific Anion Exchangers SLC26A7, -A8, and -A9.** *Journal of Biological Chemistry* 2002, **277**(16):14246-14254.
37. Dirami T, Rode B, Jollivet M, Da Silva N, Escalier D, Gaïtch N, Norez C, Tuffery P, Wolf J-P, Becq F *et al*: **Missense Mutations in SLC26A8, Encoding a Sperm-Specific Activator of CFTR, Are Associated with Human Asthenozoospermia.** *The American Journal of Human Genetics* 2013, **92**(5):760-766.
38. Mäkelä S ER, Lähdetie J, Mikkola M, Hovatta O, Kere J.: **Mutational analysis of the human SLC26A8 gene: exclusion as a candidate for male infertility due to primary spermatogenic failure.** *Mol Hum*

Reprod 2005, **11**(2):129-132.

39. Mocellin R BK: **Late banding operation in children with ventricular septal defect and pulmonary arterial hypertension.** *Eur J Cardiol* 1975 **3**(3):205-211.
40. Rode B, Dirami T, Bakouh N, Rizk-Rabin M, Norez C, Lhuillier P, Lores P, Jollivet M, Melin P, Zvetkova I *et al*: **The testis anion transporter TAT1 (SLC26A8) physically and functionally interacts with the cystic fibrosis transmembrane conductance regulator channel: a potential role during sperm capacitation.** *Hum Mol Genet* 2012, **21**(6):1287-1298.
41. Touré A, Morin L, Pineau C, Becq F, Dorseuil O, Gacon G: **Tat1, a Novel Sulfate Transporter Specifically Expressed in Human Male Germ Cells and Potentially Linked to RhoGTPase Signaling.** *Journal of Biological Chemistry* 2001, **276**(23):20309-20315.
42. Ghédir H, Ibala-Romdhane S, Okutman O, Viot G, Saad A, Viville S: **Identification of a new DPY19L2 mutation and a better definition of DPY19L2 deletion breakpoints leading to globozoospermia.** *Mol Hum Reprod* 2016, **22**(1):35-45.
43. Coutton C, Abada F, Karaouzene T, Sanlaville D, Satre V, Lunardi J, Jouk P-S, Arnoult C, Thierry-Mieg N, Ray PF: **Fine Characterisation of a Recombination Hotspot at the DPY19L2 Locus and Resolution of the Paradoxical Excess of Duplications over Deletions in the General Population.** *PLoS Genetics* 2013, **9**(3):e1003363.
44. Zhu F, Gong F, Lin G, Lu G: **DPY19L2 gene mutations are a major cause of globozoospermia: identification of three novel point mutations.** *Mol Hum Reprod* 2013, **19**(6):395-404.
45. Ellnati E, Kuentz P, Redin C, Jaber S, Vanden Meerschaut F, Makarian J, Kosciński I, Nasr-Esfahani MH, Demirel A, Gurgan T *et al*: **Globozoospermia is mainly due to DPY19L2 deletion via non-allelic homologous recombination involving two recombination hotspots.** *Hum Mol Genet* 2012, **21**(16):3695-3702.
46. Coutton C, Zouari R, Abada F, Ben Khelifa M, Merdassi G, Triki C, Escalier D, Hesters L, Mitchell V, Levy R *et al*: **MLPA and sequence analysis of DPY19L2 reveals point mutations causing globozoospermia.** *Hum Reprod* 2012 **27**(8):2549-2558.
47. Kosciński I, Ellnati E, Fossard C, Redin C, Muller J, Velez de la Calle J, Schmitt F, Ben Khelifa M, Ray P, Kilani Z *et al*: **DPY19L2 Deletion as a Major Cause of Globozoospermia.** *The American Journal of Human Genetics* 2011, **88**(3):344-350.
48. Coutton C EJ, Martinez G, Arnoult C, Ray PF: **Teratozoospermia: spotlight on the main genetic actors in the human.** *Hum Reprod Update* 2015 **21**(4):455-485.
49. Zhang H, Su D, Yang Y, Zhang W, Liu Y, Bai G, Ma M, Ma Y, Zhang S: **Some single-nucleotide polymorphisms of the TSSK2 gene may be associated with human spermatogenesis impairment.** *Journal of Andrology* 2010, **31**(4):388-392.
50. Shang P HJ, Baarends WM, Grootegoed JA: **Evolution of testis-specific kinases TSSK1B and TSSK2 in primates.** *Andrology* 2013, **1**(1):160-168.
51. Xu B, Hao Z, Jha KN, Zhang Z, Urekar C, Digilio L, Pulido S, Strauss JF, Flickinger CJ, Herr JC: **Targeted deletion of Tssk1 and 2 causes male infertility due to haploinsufficiency.** *Developmental*

Biology 2008, **319**(2):211-222.

52. Hao Z JK, Kim YH, Vemuganti S, Westbrook VA, Chertihin O, Markgraf K, Flickinger CJ, Coppola M, Herr JC, Visconti PE.: **Expression analysis of the human testis-specific serine/threonine kinase (TSSK) homologues. A TSSK member is present in the equatorial segment of human sperm.** *Mol Hum Reprod* 2004 **10**(6):433-444.
53. Kueng P NZ, Djonov V, Hemphill A, Rohrbach V, Boehlen D, Zuercher G, Andres AC, Ziemiecki A.: **A novel family of serine/threonine kinases participating in spermiogenesis.** *J Cell Biol* 1997, **139**(7):1851-1859.
54. Li Y, Sosnik J, Brassard L, Reese M, Spiridonov NA, Bates TC, Johnson GR, Anguita J, Visconti PE, Salicioni AM: **Expression and localization of five members of the testis-specific serine kinase (Tssk) family in mouse and human sperm and testis.** *Molecular Human Reproduction* 2010, **17**(1):42-56.
55. Sartor H EF, Grzeschik KH, Müller R, Adolph S.: **Assignment of two human cell cycle genes, CDC25C and CCNB1, to 5q31 and 5q12, respectively.** *Genomics* 1992, **13**(3):911-912.
56. Nakayama K, Rahman M, Rahman M, Nakamura K, Ishikawa M, Katagiri H, Sato E, Ishibashi T, Iida K, Ishikawa N *et al*: **CCNE1 amplification is associated with aggressive potential in endometrioid endometrial carcinomas.** *International Journal of Oncology* 2016.
57. van de Weerd BCM, Medema RH: **Polo-like kinases: a team in control of the division.** *Cell Cycle* 2006, **5**(8):853-864.
58. Soung N-K, Park J-E, Yu L-R, Lee KH, Lee J-M, Bang JK, Veenstra TD, Rhee K, Lee KS: **Plk1-Dependent and -Independent Roles of an ODF2 Splice Variant, hCenexin1, at the Centrosome of Somatic Cells.** *Developmental Cell* 2009, **16**(4):539-550.
59. Sen S, Zhou H, White RA: **A putative serine/threonine kinase encoding gene BTAK on chromosome 20q13 is amplified and overexpressed in human breast cancer cell lines.** *Oncogene* 1997 **14**(18):2195-2200.
60. Zhou H, Kuang J, Zhong L, Kuo W-I, Gray J, Sahin A, Brinkley B, Sen S: **Tumour amplified kinase STK15/BTAK induces centrosome amplification, aneuploidy and transformation.** *Nat Genet* 1998, **20**(2):189-193.
61. Crane R, Gadea B, Littlepage L, Wu H, Ruderman JV: **Aurora A, meiosis and mitosis.** *Biol Cell* 2004 **96**(3):215-229.
62. Hannak E, Kirkham M, Hyman AA, Oegema K: **Aurora-A kinase is required for centrosome maturation in *Caenorhabditis elegans*.** *The Journal of Cell Biology* 2001, **155**(7):1109-1116.
63. Xu H, Beasley MD, Warren WD, van der Horst GTJ, McKay MJ: **Absence of Mouse REC8 Cohesin Promotes Synapsis of Sister Chromatids in Meiosis.** *Developmental Cell* 2005, **8**(6):949-961.
64. Lee J IT, Yokota T, Yamashita M.: **Temporally and spatially selective loss of Rec8 protein from meiotic chromosomes during mammalian meiosis.** *J Cell Sci* 2003, **116**(Pt 13):2781-2790.

Tables

Table 1 List of data output quality

Sample name	Raw reads	Clean reads	Error rate(%)	Q20(%)	Q30(%)	GC content(%)
BS1	58277598	57482858	0.01	98.37	95.91	48.89
BS2	59921568	58972988	0.01	98.45	96.1	48.46
BS3	53371478	52475320	0.01	98.48	96.17	48.98
AS1	59523578	58411056	0.01	98.43	96.03	49.97
AS2	61727216	60775868	0.01	98.33	95.81	49.6
AS3	53956064	53135736	0.01	98.38	95.91	49.6

Note: 1. Raw reads—Statistical raw sequence data,

2. Clean reads: Statistics for the filtered data

3. Q20 – Q30: Calculate the percentage of the total base of Phred value greater than 20 and 30, respectively

4. GC content: Calculate the total number of base G and C and the percentage of total base number

Table 2 Comparison of Reads and reference genome

Sample name	BS1	BS2	BS3	AS1	AS2	AS3
Total reads	57482858	58972988	52475320	58411056	60775868	53135736
Total mapped	49544664	50977489	44970066	50154994	52106893	45639892
Multiple mapped	769537	820207	757219	865718	967532	881231
Uniquely mapped	48775127	50157282	44212847	49289276	51139361	44758661
Read-1	24956766	25616683	22540529	25259215	26257895	22962720
Read-2	23818361	24540599	21672318	24030061	24881466	21795941
Reads map to '+'	24358384	25055369	22098091	24639652	25572721	22376022
Reads map to '-'	24416743	25101913	22114756	24649624	25566640	22382639
Non-splice reads	32691913	34028728	30088902	28163504	29933197	26300556
Splice reads	16083214	16128554	14123945	21125772	21206164	18458105

Note:

1. Total reads—Statistics for the filtered data

2. Total mapped—The number of sequences that can be localized to the genome;

3 Multiple mapped—The number of sequencing sequences with multiple alignment positions on a reference sequence;

4. Uniquely mapped—The number of sequences in a reference sequence with a unique alignment.

5. Reads map to '+'—Reads map to '-'—The comparison of sequence alignment to positive and negative chains in genome.

6. Splice reads—means reads mapped to the border of exon, also called junction reads.

Non-splice reads—means reads for the entire sequence are mapped to one exon.

Table 3 Classification and quantity statistics of AS events

Event Type	NumEvents JC.only	SigEvents.JC.only	NumEvents JC.only+reads On Target	SigEvents.JC.only+reads On Target
SE	16702	987(431:556)	16724	991(434:557)
MXE	1888	341(172:169)	1891	311(163:148)
A5SS	43	2(0:2)	43	2(0:2)
A3SS	75	6(4:2)	76	6(4:2)
RI	40	1(1:0)	40	1(1:0)

Note:1.Event Type: AS event type(SE, MXE, A5SS, A3SS, RI)

2. NumEvents JC.only: Total number of differential AS events. Only use junction counts to AS event detection

3. SigEvents.JC.only: Total number of differential AS events. The figures in brackets before the colon is mean numbers of up-regulated genes, figures in brackets after the colon is mean numbers of down-regulated genes,

4. NumEvents JC.only+reads On Target: Total number of differential AS events. simultaneous use junction counts and reads on target to AS event detection

5. SigEvents.JC.only+reads On Target: Total number of differential AS events, meaning as (3)

Figures

Figure 1

Fig1a Volcanic map of differential genes Note: There are significant differences in the expression of genes with red (up) and green (down) point. No significant difference of the expression of gene expression with blue point. The abscissa represents the fold change of genes in different samples; ordinate represents the significant statistical difference of gene expression changes. Fig1b Differential gene cluster Note: The overall FPKM hierarchical clustering diagram, the log₁₀ (FPKM+1) value for the normalization of the conversion (scale number) and clustering, the red represents a high expression of genes, blue expression of low expression genes. Color from red to blue, indicating that log₁₀ (FPKM+1) from large to small. Fig1c The venn diagram of gene expression The sum of the numbers in each large

circle represents the total number of genes expressed by the group, and the overlapping part of the circle represents the common expression gene between group and fpkm>1.

Figure 2

Fig2a The most enriched GO term Note: The ordinate is the enriched GO term, and the abscissa is the number of differential genes in term. Different colors were used to distinguish biological processes, cellular components and molecular functions. With "*" as a significant enrichment of GO term Fig2b The most enriched GO term Note: GO in figure2A, According to the biological process, cellular component and molecular function of three categories and the differences by painting histogram classification.

Figure 3

Scatter plot of differential gene KEGG Note: The vertical axis represents the name of the pathway, the horizontal axis represents the Rich factor. The differential size of points show number of differential expression genes, and the color of points mean different range of Qvalue.

Oocyte meiosis pathways Note: Differentially expressed genes in horse testis and their involvement in the Oocyte meiosis pathway. Genes in red boxes are up-regulated and genes in green boxes are down-regulated in ASM testis versus BSM testis. This figure were cited in <http://www.kegg.jp/>

Supplementary Files

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

- [Supplementaltable5.xls](#)
- [Supplementaltable7.xls](#)
- [Supplementaltable1.xls](#)
- [Supplementaltable6.xls](#)
- [Supplementaltable3.xls](#)
- [Supplementaltable8.xls](#)
- [Supplementaltable2.xls](#)
- [Supplementaltable4.xls](#)