

# LncRNAs induce oxidative stress and spermatogenesis by regulating endoplasmic reticulum genes and pathways

Tie-Cheng Sun<sup>1,2,\*</sup>, Yi Zhang<sup>5,\*</sup>, Kun Yu<sup>6,\*</sup>, Yao Li<sup>6</sup>, Hong Yu<sup>1</sup>, Shan-Jie Zhou<sup>1</sup>, Ya-Peng Wang<sup>4,&</sup>, Shou-Long Deng<sup>3</sup>, Li Tian<sup>1,2</sup>

<sup>1</sup>Reproductive Medical Center, Department of Obstetrics and Gynecology, Peking University International Hospital, Beijing 102206, China

<sup>2</sup>Reproductive Medicine Centre, Peking University Second Affiliated Hospital, Beijing 100044, China

<sup>3</sup>Institute of Laboratory Animal Sciences, Chinese Academy of Medical Sciences and Comparative Medicine Center, Peking Union Medical College, Beijing 100021, China

<sup>4</sup>Center of Reproductive Medicine, Peking University Third Hospital, Beijing 100191, China

<sup>5</sup>Department of Medicine, Panzihua University, Sichuan 16700, China

<sup>6</sup>Beijing Key Laboratory for Animal Genetic Improvement, National Engineering Laboratory for Animal Breeding, Key Laboratory of Animal Genetics and Breeding of the Ministry of Agriculture, College of Animal Science and Technology, China Agricultural University, Beijing 100193, China

\*Equal contribution

**Correspondence to:** Ya-Peng Wang, Li Tian, Shou-Long Deng; **email:** [yawang@163.com](mailto:yawang@163.com); <https://orcid.org/0000-0002-9677-6131>; [tianli@pkuih.edu.cn](mailto:tianli@pkuih.edu.cn); [dengshoulong@cnilas.org](mailto:dengshoulong@cnilas.org)

**Keywords:** long noncoding RNAs, oligozoospermia, spermatogenesis, oxidative stress, apoptosis

**Received:** June 24, 2020

**Accepted:** November 8, 2020

**Published:** May 17, 2021

**Copyright:** © 2021 Sun et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/3.0/) (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

## ABSTRACT

Oligozoospermia or low sperm count is a leading cause of male infertility worldwide. Despite decades of work on non-coding RNAs (ncRNAs) as regulators of spermatogenesis, fertilization, and male fertility, the literature on the function of long non-coding RNAs (lncRNAs) in human oligozoospermia is scarce. We integrated lncRNA and mRNA sequencing data from 12 human normozoospermic and oligozoospermic samples and comprehensively analyzed the function of differentially expressed lncRNAs (DE lncRNAs) and mRNAs (DE mRNAs) in male infertility. The target genes of DE lncRNAs were identified using a Gaussian graphical model. Gene ontology terms and Kyoto Encyclopedia of Genes and Genomes pathways were primarily enriched in protein transport and localization to the endoplasmic reticulum (ER). The lncRNA–mRNA co-expression network revealed cis- and trans-regulated target genes of lncRNAs. The transcriptome data implicated DE lncRNAs and DE mRNAs and their target genes in the accumulation of unfolded proteins in sperm ER, PERK-EIF2 pathway-induced ER stress, oxidative stress, and sperm cell apoptosis in individuals with oligozoospermia. These findings suggest that the identified lncRNAs and pathways could serve as effective therapeutic targets for male infertility.

## INTRODUCTION

Infertility is a global problem affecting human reproductive health. According to the World Health Organization, global fertility and population growth rates have continuously declined over the past 50 years

(Figure 1). It is estimated that about 10 to 15% of couples in the reproductive ages are infertile, with approximately 40 to 50% of infertility cases attributed to “male factor” [1–4]. The most common causes of male infertility are low sperm concentration (oligozoospermia) and reduced sperm motility

(asthenospermia), which often occur together [5, 6]. Oligozoospermia is characterized by qualitative and quantitative defects in spermatogenesis, manifested as poor sperm motility and morphology and sperm apoptosis [7]. Compared with healthy individuals, semen parameters, such as progressive rate, non-progressive rate, and DNA fragmentation index (DFI), are abnormal in those with oligozoospermia [7, 8], reducing the possibility of the sperm to reach the egg in the oviduct [6, 9–11]. Several studies have implicated chromosomal abnormalities, gene regulation disorders, environmental factors, infections, immune-related factors, and endocrine dysfunction in the etiology of oligozoospermia [6, 9].

Epigenetic changes are inherited modifications that regulate gene expression, without any alteration in the underlying nucleotide sequence [12–21]. Several external factors, such as diet, air pollution, and chemical exposure, influence epigenetic modifications [12, 17, 22]. Moreover, genetic studies report that various environmental factors affect gametogenesis, early sperm development, and sperm apoptosis by epigenetically reprogramming the genome and thus reducing the sperm count [12–20].

Non-coding RNAs (ncRNAs) regulate transcription and translation by inhibiting the binding of transcription factors to their target DNA. They can also epigenetically regulate gene expression by activating epigenetic modifiers to enhance DNA methylation and

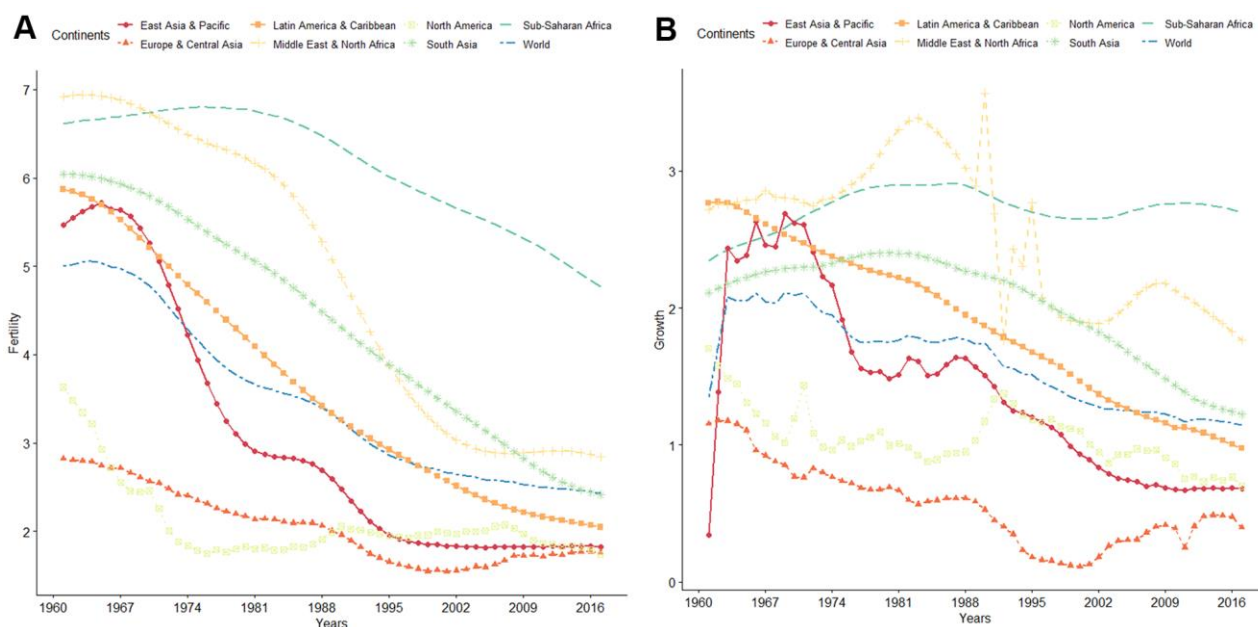
histone modifications [23–26]. Based on their length, ncRNAs are categorized as long ncRNAs (lncRNAs), short ncRNAs, and pseudogenes. Short ncRNAs such as miRNAs and PIWI-interacting RNAs (piRNAs) regulate the development of germ cells, primordial germ cell specialization, and early and late spermatogenesis, and are involved in the pathogenesis of various reproductive disorders [27–29]. lncRNAs—the largest category of ncRNAs—control the transcription factors linked to primordial germ cell specialization, such as BLIMP1, PRDM1, and DAZL, and maintain the survival of spermatogonial cells [27, 30–36].

High-throughput screening of bull semen transcriptome identified the function of ncRNAs in sperm motility [37]. We compared the sperm transcriptome profiles of patients with obstructive azoospermia (OA) ( $n = 3$ ) and individuals with normal spermatogenesis (including the control group) to systematically define the expression profiles of lncRNAs and mRNAs. We next used these expression data to identify oligozoospermia-related lncRNAs and functional genes and their involvement in male infertility.

## RESULTS

### RNA sequencing and identification of DE mRNAs and DE lncRNAs

We performed a case–control study involving six normozoospermic (NS) samples from healthy fertile



**Figure 1.** (A) WHO data on global trends of fertility from 1960 to 2016. (B) WHO data on global trends of population growth rate from 1960 to 2016.

controls and six oligozoospermic (OS) samples from infertile patients to identify the major perturbations in oligozoosperms. Cases and controls were closely matched for age, body mass index (BMI), and sperm DFI (Table 1). The levels of reproductive hormones, such as follicle-stimulating hormone (FSH), luteinizing hormone (LH), and testosterone (T), were analyzed. As summarized in Table 1, we did not observe any differences in the levels of these hormones between NS and OS samples.

We next assessed the association between semen characteristics and oligozoospermia. Five of the six examined semen characteristics (concentration, progressive rate, non-progressive rate, motility, and normal morphology) differed between the two groups ( $p = 0.0001$ – $0.048$ ; Table 1). Compared with individuals in the OS group, those in the NS group had a 9-fold higher concentration of collected sperms (NS:  $81.78 \pm 37.21$ ; OS:  $8.83 \pm 3.55$ ). Similar patterns were observed for other characteristics (progressive rate, non-progressive rate, motility, and normal morphology) using the *t*-test (Table 1). The remaining characteristic, ejaculate volume, was not associated with the development of oligozoospermia ( $p = 0.939$ ).

We performed a high-throughput whole transcriptome shotgun sequencing to compare the sperm RNA expression patterns between NS and OS groups. We extracted high-quality RNA from each sperm sample and constructed cDNA libraries. Next, we calculated the levels of transcripts of lncRNAs and mRNAs on an Illumina HiSeq X-Ten Platform, which supported a read length of  $2 \times 150$  bp with quality scores of  $\geq 75\%$  of bases above Q30. After filtering the raw data, 24.4 to 79.8 million clear reads and 2,501 lncRNAs and mRNAs were identified. Distribution of base composition and distribution of quality are shown in Figure 2. Figure 2 shows the characteristics of identified lncRNAs and mRNAs. The lncRNA transcripts were shorter in length (Figure 2D) and contained fewer exons than mRNAs (Figure 2A). The majority of identified lncRNAs contained one to two transcripts, whereas most mRNAs had one to four transcripts (Figure 2C)—findings that are in agreement with those of previous studies [30–32]. The length of open reading frames (ORFs) of mRNAs was shorter than that of lncRNA ORFs (Figure 2B).

### **LncRNA and mRNA characterization**

RNA sequencing analysis revealed that of the 55,278 identified lncRNAs and mRNAs, 4,963 were differentially expressed (DE) in OS samples as compared with NS samples (2,364 lncRNAs and 2,599 mRNAs; Supplementary Tables 1, 2). Of these, only

464 lncRNAs and 875 mRNAs were downregulated; 1,900 lncRNAs and 1,724 mRNAs were upregulated (Supplementary Tables 1, 2). Moreover, 772 lncRNAs and 410 mRNAs were identified in oligozoosperms with a  $\log_2$  fold change (FC)  $> 6$  or  $< -6$  compared with normozoosperms (Supplementary Table 2 and Figure 3). The most upregulated lncRNA was lnc-GNS-3 (FC = 10.0921, FDR =  $2.91E-08$ ) and the most downregulated lncRNA was ITGA9-AS1 (FC =  $-6.2743$ , FDR =  $1.60E-07$ ). The most upregulated mRNA was ZNF350 (FC = 7.3363, FDR =  $3.16E-08$ ) and the most downregulated mRNA was ASPH (FC =  $-6.4339$ , FDR =  $7.92E-08$ ). The DE lncRNAs and DE mRNAs were widely scattered among all chromosomes although the distribution was unequal (Supplementary Tables 3, 4). Chromosome 1 had the highest number of altered RNAs, 213 lncRNAs and 247 mRNAs that accounted for 9.01% (213/2,364) and 9.50% (247/2,599) of all DE lncRNAs and DE mRNAs, respectively (Figure 4A, 4B). The X chromosome had the highest number of DE lncRNAs and DE mRNAs (52 and 82) than the Y chromosome (5 and 6) (Figure 4A, 4B). Box plots revealed the lncRNA and mRNA expression patterns in each sperm sample (Figure 2E). Furthermore, analysis using PhastCons software revealed that lncRNA transcripts were less conserved than mRNA transcripts (Figure 2F).

### **Classification of DE lncRNAs and DE mRNAs**

We next classified DE lncRNAs and DE mRNAs and performed subgroup analysis to study their potential functions (Tables 2, 3). The DE lncRNAs were distributed into four subgroups: sense, antisense, intron, and intergenic (Supplementary Tables 5, 6). Antisense lncRNAs accounted for 48.1% of the total lncRNAs (1,077/1,900 upregulated and 60/464 downregulated), whereas intergenic lncRNAs constituted the second largest category (315/1,900 upregulated and 233/464 downregulated) (Figure 4C). Approximately 10.6% of the lncRNAs belonged to the intron subgroup (199/1,900 upregulated and 51/464 downregulated) (Figure 4C) and 18.1% lncRNAs were identified with sense regulation (309/1,900 upregulated and 120/464 downregulated) (Figure 4C).

### **Enrichment analysis of DE mRNAs and DE lncRNAs**

We next performed GO and KEGG pathway enrichment analyses to study the functions of DE genes corresponding to DE mRNAs and DE lncRNAs. Overrepresentation (ora) GO analysis revealed 777, 196, and 188 terms enriched in biological processes (BP), cell components (CC), and molecular functions (MF), respectively. Similar results were obtained in the

**Table 1. Clinical data of participants.**

Parameters	Normospermia fertile controls ( <i>n</i> = 6)	Oligozoospermia infertile group ( <i>n</i> = 6)	<i>p</i> -Value
Age (y)	32.83 ± 4.96	37.67 ± 3.56	0.081
BMI (kg/m <sup>2</sup> )	24.32 ± 2.44	21.31 ± 2.68	0.069
DFI (%)	18.82 ± 14.58	35.54 ± 19.78	0.127
FSH (IU mL <sup>-1</sup> )	5.33 ± 2.17	9.11 ± 4.19	0.078
LH (IU mL <sup>-1</sup> )	4.11 ± 0.84	5.8 ± 4.3	0.366
T (ng mL <sup>-1</sup> )	4.0 ± 2.21	4.1 ± 2.34	0.943
Volume (mL)	4.22 ± 2.38	4.32 ± 2.01	0.939
Concentration (Million mL <sup>-1</sup> )	81.78 ± 37.21	8.83 ± 3.55	0.000
PR (%)	56.82 ± 17.16	32.01 ± 19.07	0.039
NP (%)	16.04 ± 6.05	8.6 ± 5.39	0.048
Motility (%)	72.86 ± 16.59	40.62 ± 23.31	0.020
Normal morphology (%)	4.83 ± 0.52	3.93 ± 0.71	0.031

Abbreviations: BMI, body mass index; DFI, DNA fragmentation index; FSH, follicle-stimulating hormone; LH, luteinizing hormone; PR, progressive rate; NP, non-progressive rate; T, testosterone.

pathway-level analysis (plage). These GO terms were primarily related to intracellular protein transport and localization. GO terms with the highest differential expression were enriched in SRP-dependent co-translational protein targeting to the membrane, co-translational protein targeting to the membrane, and protein targeting to the ER (Figure 5A). Table 4 lists the DE genes associated with these processes, including CHMP4B, SEC61A1, SEC61A2, SEC61G, SEC62, SGTB, SPCS1, SRP9, SRP19, RYR2, ANK2, DDRGK1, INSIG1, KDELR2, RER1, and RTN4. Furthermore, differential GO terms were enriched in ER stress, oxidative stress, protein unfolding, and cell apoptosis (Figure 5B, Supplementary Table 7). Twenty-six over-represented GO terms were directly related to negative regulation of response to ER stress (Supplementary Table 7). In the case of protein unfolding, unfold protein binding had a significant difference (Supplementary Table 7). Five GO terms were related to oxidative stress, including neuron death in response to aerobic stress, negative regulation of aerobic stress-induced intrinsic acoustic signaling pathway, positive regulation of aerobic-stress induced intrinsic acoustic pathway, and regulation of oxidative stress-induced neuron death (Supplementary Table 7). Other apoptosis GO terms were those involved in the negative regulation of apoptosis (Supplementary Table 7).

Differentially expressed genes involved in these processes are listed in Tables 5, 6. These included upregulated EIF2a, ATF5, ATF4, XBP1, SERP1, and CEBPZ associated with ER stress. Only a few genes, such as MAPK8, were downregulated (Table 5). Although EIF2AK3, an ER stress-related gene, was not

differentially expressed (Table 5), its homologous gene EIF2AK1 was upregulated (Table 5). Similarly, although ERN1 was not differentially expressed, it displayed more than 2-fold change (Table 5). Oxidative stress-related DE genes included SOD1, SOD2, NOXA1, XDH, and MAPK14 (Table 6). The expression of SOD1, SOD2, MAPK14, NFKB1, and SP1 was upregulated (Table 6), whereas that of NOX1, XDH, NFE2L1, and NCF1C was downregulated (Table 6). HOMX2 was upregulated but not HMOX1 (Table 6). Similarly, GPX3 was not differentially expressed; however, GPX1 and GPX4 were upregulated (Table 6). Further, certain apoptotic genes, such as CASP6, MCL1, NFKB1, BIRC2, AKT3, GADD45G, CAPN1, EIF2A, RAF1, MAP2K1, ENDO, and TNFRSF1A, were differentially expressed (Supplementary Table 2).

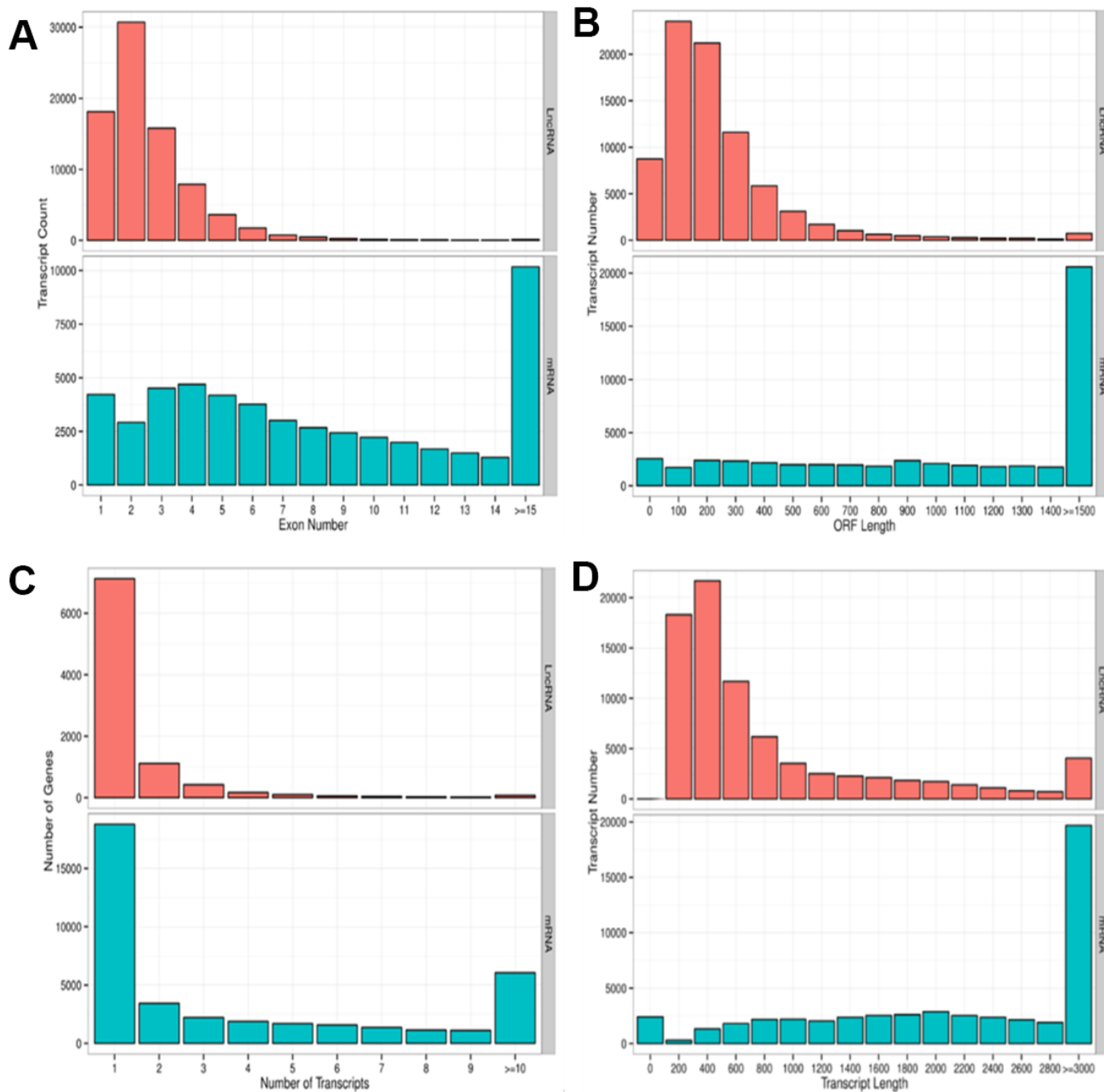
The majority of DE genes were also enriched in the identified KEGG pathways, such as Alzheimer's disease pathway, Parkinson's disease pathway, ribosome and fatty acid degradation pathway (Supplementary Table 8). Alzheimer's disease and Parkinson's disease pathways were associated with ER stress or oxidative stress-induced cell apoptosis, consistent with the previous GO terms analysis results (Supplementary Table 7). However, certain genes involved in ER stress and oxidative stress initiation in these pathways, such as APP and PRKN, were not differentially expressed in oligozoospermic individuals (Supplementary Table 2).

### Co-expression analysis of DE lncRNAs

The co-expression network analysis of lncRNAs and mRNAs in oligozoospermic individuals revealed

potential internal adjustment mechanisms. Several lncRNAs correlated with a single mRNA and vice versa. For example, 21,468 lncRNA and mRNA partial correlation pairs were found for the top 20 DE lncRNAs (Supplementary Table 9). Of these, 1,542 pairs showed a strong correlation coefficient, 8,421 pairs showed a moderate correlation coefficient, and 11,505 pairs had a weak correlation coefficient (Supplementary Table 9).

The majority of these lncRNAs had common correlated genes. For example, the upregulated lncRNA lnc-GNS-3 correlated with target genes BANF1, TXN2, PIK3C3, and KDM4b, which also strongly correlated with other upregulated lncRNAs, namely lnc-SERHL2-7, lnc-PHLDB1-1, and lnc-IL31RA-1 (Figure 6 and Supplementary Table 9). However, the majority of genes associated with XLOC\_515910 and lnc-NLRP2-2, such



**Figure 2. Parameters of lncRNA and mRNA transcripts.** (A) Distribution pattern of exon numbers of lncRNA and mRNA transcripts. (B) Distribution pattern of ORF length of lncRNA and mRNA transcripts. (C) Distribution pattern of gene numbers of lncRNA and mRNA transcripts. (D) Distribution pattern of transcript lengths of lncRNA and mRNA transcripts.



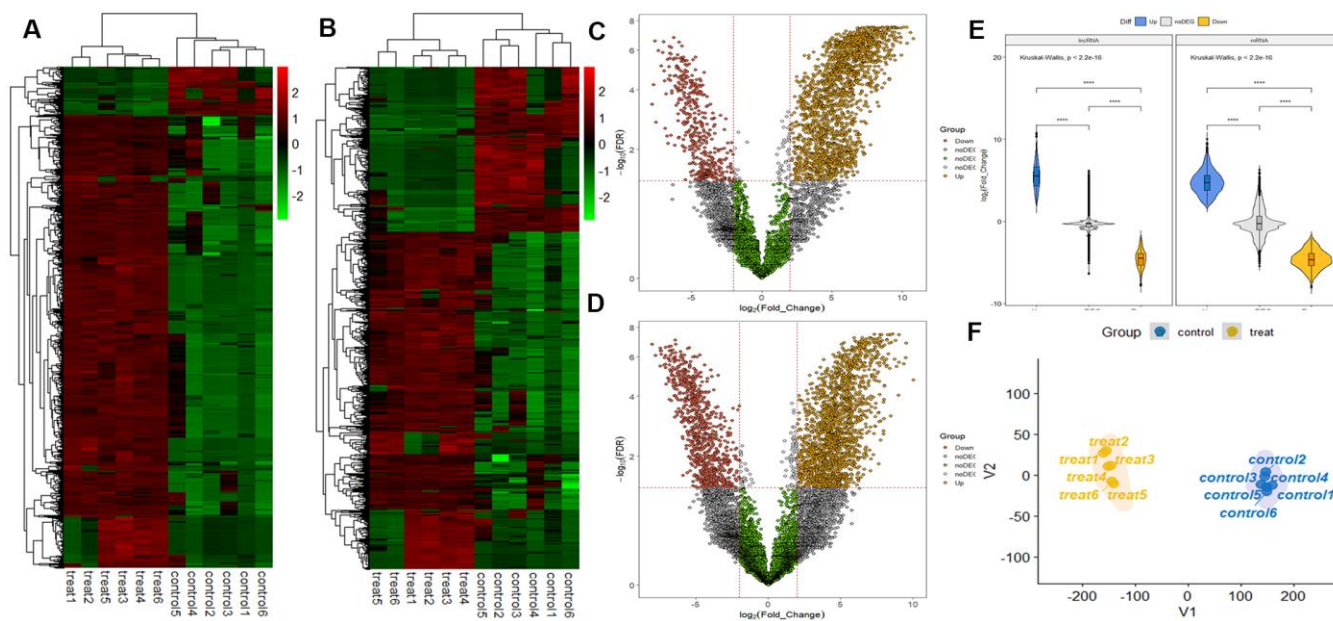
as CHCHD5 and KAT2B, only correlated with one of the top DE lncRNAs (Figure 6, Supplementary Table 9).

LncRNAs strongly correlating with EIF2A, a major ER stress gene, were lnc-FAM198B-1, lnc-C4orf21-1, lnc-SPIN3-5, and lnc-AC006050.2.1-6 (Supplementary Tables 10, 11, 22). Those with a moderate negative correlation included lnc-TMEM75-8, lnc-CYB5R2-3, XLOC\_2993599, and lnc-MPP7-8 (Table 7). LncRNAs with moderate positive correlation included lnc-TOR3A-1, lnc-STARD9-1, lnc-SLC25A47-5, and XLOC\_002818 (Table 7). Lnc-TULP4-1, lnc-PTPN2-6, FOXP1-AS1, and lnc-ZNF101-4 showed a strong positive correlation with oxidative stress-related gene SOD1 (Table 7). LncRNAs with moderate negative correlation with SOD1 included XLOC\_1174819, XLOC\_2396188, XLOC\_2837388, and XLOC\_23762 (Table 7). Those with moderate positive correlation included lnc-BCAP31-1, lnc-C2orf84-1, XLOC\_2394463, and lnc-ADAM21-4. SRP9, a gene involved in protein localization, positively correlated with lnc-IARS2-2, lnc-WDR77-2, lnc-C1QL3-1, and lnc-SLC36A1-5 (Table 7). It showed a moderate negative correlation with lnc-PPP1R3G-3, lnc-GGCT-1, XLOC\_2028132, and lnc-MPP7-8 (Table 7). Lnc-

RAD9B-1, lnc-ZNF687-1, lnc-SLC3A2-3, and lnc-MAP6-1 exhibited a moderate positive correlation with SRP9. ER stress, oxidative stress, and protein localization shared several lncRNAs (Table 7). For example, ADIRF-AS1 not only controlled EIF2A and EIF2AK1 in ER stress but was also related to NCF2 and NCF1C of oxidative stress, and SPCS1 and UBAC2 of protein transport (Supplementary Tables 12, 13, and Table 7).

To study the function of lncRNAs in gene regulation and pathogenesis of oligozoospermia, cis- and trans-predictions were performed based on co-expression network analysis. All top 20 DE lncRNAs were predicted to have trans-regulated target genes. Lnc-TICAM1-1 and lnc-PHLDB1-regulated the expression of KDM4B and TRAPPC4 with cis-interaction effect (Supplementary Table 14). Further, lnc-TICAM1-1 may be regulated by transcription factors LF-A1, p300, and EllaE-A (Supplementary Table 14).

The majority of identified lncRNAs in ER stress, oxidative stress, and protein transport exerted trans-regulation effects on target genes; however, no lncRNA displayed cis-regulation effect (Supplementary Tables 15–18, 22, and Figure 7A). LncRNAs



**Figure 3. LncRNA and mRNA profiles based on RNA sequencing data.** (A) Comparison of two-dimensional hierarchical clustering of distinguishable lncRNA expression profiles in individuals with oligozoospermia and controls. Red indicates high expression, green indicates low expression. Probes are shown in rows, and samples are shown in columns. (B) Two-dimensional hierarchical clustering of distinguishable mRNA expression profiles. (C) Volcano plot of differentially expressed lncRNAs in individuals with oligozoospermia compared with normal controls. Red points represent downregulated lncRNAs and yellow points represent upregulated lncRNAs with oligozoospermia with a greater than 2.0-fold change. (D) Volcano plot of differentially expressed mRNAs. (E) Violin plot of lncRNA and mRNA profiles in individuals with oligozoospermia. (F) t-distributed stochastic neighbor embedding (t-SNE) plot of samples based on lncRNA and mRNA profiles in individuals with oligozoospermia.

with trans-regulation effect included lnc-TSKS-1, lnc-PRR15L-2, and lnc-NT5DC2-1 (Supplementary Tables 15–19). The ER stress-related lnc-TSKS-1 may be regulated by ELLaE-A, E47, Elk-1, and other transcription factors (Supplementary Table 20 and Figure 7B). The oxidative stress-related lnc-PRR15L-2 may be correlated with transcription factors C/EBPalpha, C/EBPbeta, and wt1. L (Supplementary Table 20). Other transcription factors such as p300, c-Ets-1, and R2 may regulate lnc-NT5DC2-1 involved in protein transport (Supplementary Table 21).

To validate the reliability of the lncRNA microarray data, we selected five upregulated lncRNAs (NFKB1, XBP1, SRP9, EIF2AK1, and ATF4) that were abundantly and differentially expressed (FCs > 6.0). qRT-PCR was used to analyze the differences in their expression. The results of the qRT-PCR analysis were largely consistent with those of the microarray data (Figure 8).

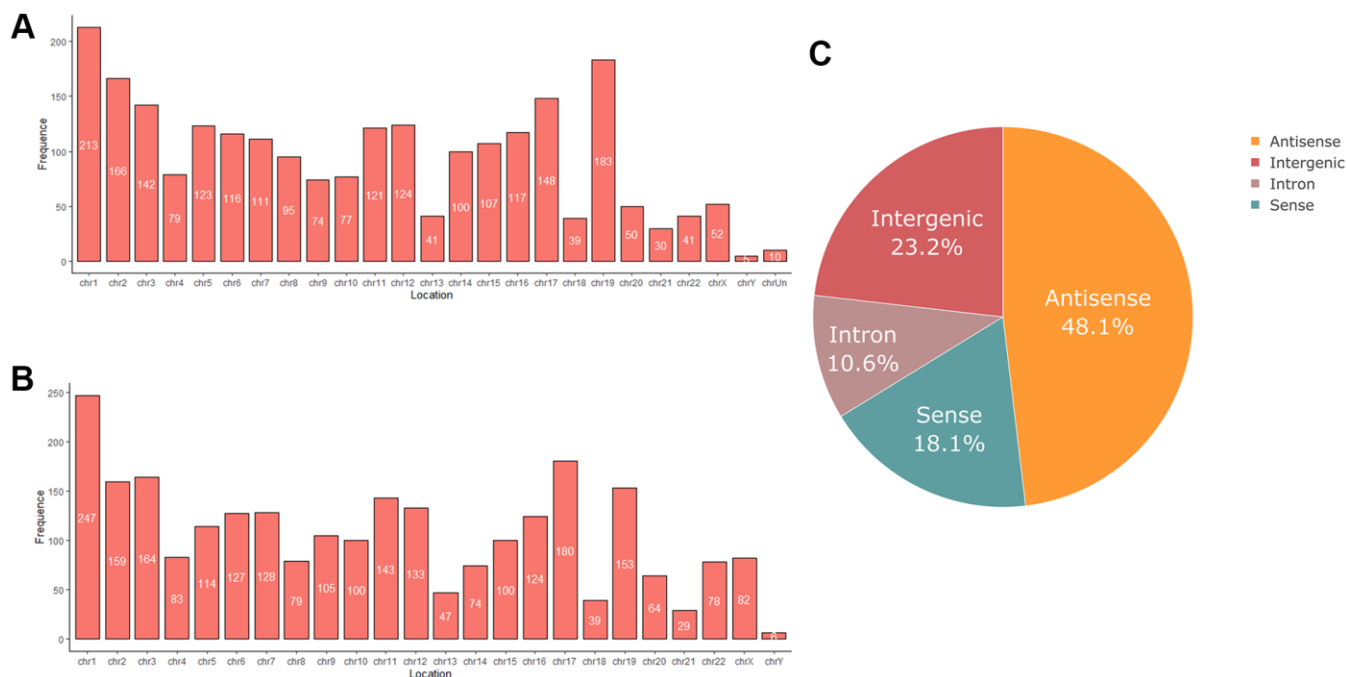
## DISCUSSION

Non-coding RNAs participate in the pathogenesis and development of several reproductive diseases. For example, certain miRNAs, such as mir-290-295 and mir-34c, are involved in spermatogenesis and oogenesis [27, 38, 39], and their dysregulation results in primordial germ cell migration disorders or

downregulated expression of Nanos2 in spermatogonial stem cells [27, 38, 39]. We first screened the whole genome expression patterns of lncRNAs and mRNAs in spermatozoa of patients with oligospermia and normal individuals. We next compared the expression profiles of lncRNAs and mRNAs between the two groups using bioinformatics and systematically analyzed the characteristics of oligozoospermia-related lncRNAs and mRNAs [40, 41].

Our results revealed that 2,364 lncRNAs and 2,599 mRNAs were abnormally expressed in patients with oligozoospermia. The proportion of downregulated lncRNAs and mRNAs was less than that of upregulated ones. Furthermore, we did not find any difference in the expression of lncRNAs by qRT-PCR and RNA-sequencing [40, 41], which could be attributed to different normalization processes used in the two methods [41–43].

The chromosomal distribution of DE lncRNAs and DE mRNAs was uneven, with the majority of them located on chromosomes 1, 17, and 19. The number of differential genes on the X chromosome was higher than that on the Y chromosome [44]. Bache conducted a cohort study of male infertility and reported 5 out of 10 disease-related autosomal bands on chromosome 1 [45], suggesting that chromosome 1 contains a domain



**Figure 4.** (A) Chromosome distribution of differentially expressed lncRNAs in oligozoospermia. (B) Chromosome distribution of differentially expressed mRNAs in oligozoospermia. (C) Pie chart of differentially expressed lncRNAs identified in different subgroup categories.

**Table 2. Top differentially expressed lncRNAs in oligozoospermia.**

<b>LncRNA ID</b>	<b>Regulation</b>	<b>Fold change</b>	<b>p-Value</b>	<b>Corr P</b>	<b>Type</b>
lnc-GNS-3	Upregulation	10.09212	1.44E-12	2.91E-08	Intergenic
XLOC_515910	Upregulation	8.866827	4.46E-12	2.91E-08	Sense
lnc-SERHL2-7	Upregulation	10.29479	5.34E-12	2.91E-08	Antisense
lnc-IL31RA-1	Upregulation	9.21277	6.47E-12	2.91E-08	Intron
lnc-NLRP2-2	Upregulation	7.527131	7.31E-12	2.91E-08	Antisense
lnc-CDK12-2	Upregulation	9.47426	7.56E-12	2.91E-08	Antisense
XLOC_1093926	Upregulation	7.485795	7.96E-12	2.91E-08	Intron
lnc-LRRC38-2	Upregulation	8.591819	8.33E-12	2.91E-08	Antisense
lnc-FOXN1-1	Upregulation	8.90089	8.60E-12	2.91E-08	Antisense
lnc-SLC46A2-1	Upregulation	9.376134	8.89E-12	2.91E-08	Antisense
lnc-CSNK1A1-7	Upregulation	8.601429	9.60E-12	2.91E-08	Antisense
lnc-PHLDB1-1	Upregulation	8.974802	9.65E-12	2.91E-08	Antisense
lnc-KB-1980E6.3.1-6	Upregulation	7.757378	1.04E-11	2.91E-08	Antisense
lnc-LAT2-2	Upregulation	7.96087	1.04E-11	2.91E-08	Intron
XLOC_2394941	Upregulation	8.509539	1.24E-11	3.12E-08	Intergenic
lnc-EEF1B2-3	Upregulation	7.892956	1.30E-11	3.12E-08	Antisense
lnc-AP3S1-15	Upregulation	9.267483	1.34E-11	3.12E-08	Antisense
lnc-SLA2-1	Upregulation	6.812729	1.68E-11	3.31E-08	Intron
lnc-RAPH1-6	Upregulation	6.822993	1.76E-11	3.31E-08	Antisense
lnc-TICAM1-1	Upregulation	8.925062	1.77E-11	3.31E-08	Antisense

**Table 3. Top differentially expressed mRNAs in oligozoospermia.**

<b>Gene ID</b>	<b>Gene symbol</b>	<b>Regulation</b>	<b>Fold change</b>	<b>p-Value</b>	<b>Corr P</b>
59348	ZNF350	Upregulation	7.336254	1.43E-11	3.16E-08
283651	HMG2P46	Upregulation	9.026764	3.58E-11	3.36E-08
146956	EME1	Upregulation	7.822658	3.68E-11	3.36E-08
164633	CABP7	Upregulation	8.243164	3.69E-11	3.36E-08
125111	GJD3	Upregulation	8.452515	3.77E-11	3.36E-08
692094	MSMP	Upregulation	8.397109	4.12E-11	3.45E-08
8896	BUD31	Upregulation	6.498333	8.85E-11	5.68E-08
148808	MFS4	Upregulation	7.073111	8.85E-11	5.68E-08
11267	SNF8	Upregulation	5.932717	9.18E-11	5.68E-08
51003	MED31	Upregulation	6.986587	9.65E-11	5.68E-08
3206	HOXA10	Upregulation	7.704079	1.03E-10	5.68E-08
51399	TRAPPC4	Upregulation	7.058648	1.43E-10	6.68E-08
100130771	EFCAB10	Upregulation	7.244548	1.48E-10	6.78E-08
340719	NANOS1	Upregulation	7.507104	1.50E-10	6.78E-08
51282	SCAND1	Upregulation	6.752215	1.54E-10	6.86E-08
29893	PSMC3IP	Upregulation	8.194087	1.81E-10	7.76E-08
444	ASPH	Downregulation	-6.434	1.89E-10	7.92E-08
529	ATP6V1E1	Upregulation	6.463281	1.96E-10	8.02E-08
6202	RPS8	Upregulation	8.205232	1.98E-10	8.04E-08
220988	HNRNPA3	Upregulation	8.786221	2.13E-10	8.43E-08

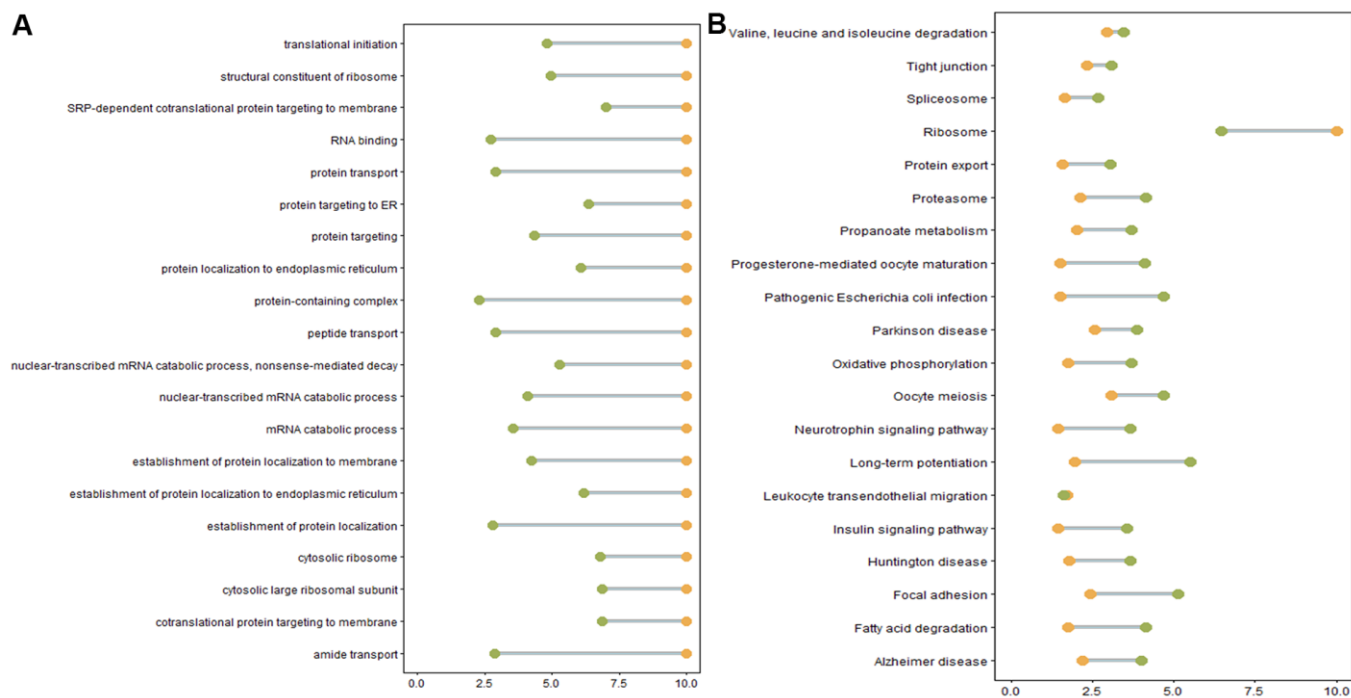


regulating the male reproductive capacity. Further, Bache reported that men with pericentric inversion of chromosome 1 were at the risk of developing infertility due to spermatogenesis dysfunction [46]. Similar findings were reported by Balasar who found a large pericentric inversion of chromosome 1 (46, XY, inv (1) (p22q32)) during a routine chromosomal analysis [46].

lncRNAs are categorized into five subgroups depending on their different transcriptional forms: sense lncRNAs, antisense lncRNAs, intronic lncRNAs, intergenic lncRNAs, and divergent lncRNAs [47]. Intergenic and antisense lncRNAs constituted 23.2% and 48.1%, respectively, i.e., almost three-quarters of the identified DE lncRNAs. Antisense lncRNAs are transcribed from the antisense strand of the protein-coding genes, whereas intergenic lncRNAs are transcribed from both strands [47, 48]. Several abnormally expressed intergenic and antisense lncRNAs identified in patients with oligozoospermia indicated that lncRNAs regulated protein-coding genes during oligozoospermia. Moreover, intergenic lncRNAs have higher tissue specificity than protein-coding genes and thus are excellent indicators of different cell subsets in the auxiliary diagnosis of oligozoospermia [47, 49].

GO and KEGG analyses revealed several genes and pathways to be enriched in protein transport and localization, such as protein targeting to ER, SRP-dependent co-translational protein targeting to the membrane, and protein localization, in the OS group. Thus, we speculated these DE genes to be associated with spermatogenesis dysfunction and used them as markers for predicting oligozoospermia and studying its molecular mechanism. However, we ignored the fact that the degree of spermatozoa apoptosis was considerably higher in patients with oligozoospermia than in normal individuals [50], whereas hyperactive regulation of protein translation, transcription, and transport could stimulate cell apoptosis. Therefore, there existed differences in protein transport and localization between patients and normal individuals.

Endoplasmic reticulum stress occurs due to the accumulation of unfolded proteins caused by impaired protein folding or mutations [51]. Abnormal activation of protein targeting processes transport excess peptides to ER, thus increasing the probability of protein unfolding [52]. Three branches of ER stress response have been identified, namely IRE1 (inositol-requiring-1), PERK (protein kinase RNA-like ER kinase), and ATF6 (activating transcription factor-6) [53–56].



**Figure 5.** (A) Enriched GO terms. Green: Log10 *p*-value from over-representation analysis (ora), orange: Log10 *p*-value from pathway-level analysis (plage). (B) Enriched KEGG pathways. Green: Log10 *p*-value from over-representation analysis (ora), orange: Log10 *p*-value from pathway-level analysis (plage).

**Table 4. Expression profile of crucial genes involved in protein transport and localization.**

Gene ID	Gene symbol	Regulation	Fold change	p-Value	Corr P
128866	CHMP4B	Upregulation	4.641415	1.26E-08	9.73E-07
29927	SEC61A1	Upregulation	4.333018	2.73E-06	7.21E-05
55176	SEC61A2	Upregulation	7.408156	3.36E-05	0.000613
23480	SEC61G	Upregulation	2.779526	0.009469	0.070513
7095	SEC62	Upregulation	4.989299	0.000227	0.003211
6449	SGTA	None	-1.45366	0.079382	0.318414
54557	SGTB	Upregulation	2.447903	0.008083	0.062471
28972	SPCS1	Upregulation	4.387165	5.97E-05	0.001012
9789	SPCS2	None	-1.45238	0.443984	0.607944
60559	SPCS3	None	3.084166	0.071509	0.301378
6726	SRP9	Upregulation	6.771347	1.96E-05	0.000387
6727	SRP14	None	3.593638	0.05458	0.251185
6728	SRP19	Upregulation	3.648171	0.006099	0.04998
6731	SRP72	None	3.660797	0.05925	0.266326
6729	SRP54	None	0.969371	0.671303	0.769497
6747	SSR3	None	3.791997	0.082548	0.320391
23471	TRAM1	None	2.082365	0.305616	0.487579
9697	TRAM2	None	-0.83275	0.089043	0.320391
6262	RYR2	Downregulation	-5.41927	0.000268	0.003698
287	ANK2	Downregulation	-7.4485	3.98E-09	4.25E-07
65992	DDRKG1	Downregulation	-3.63455	0.002235	0.022008
3638	INSIG1	Upregulation	4.908842	0.004901	0.04179
11014	KDELR2	Upregulation	3.895493	0.004696	0.040402
11079	RER1	Upregulation	4.332667	0.001388	0.014811
57142	RTN4	Upregulation	3.391149	0.021874	0.133722
202018	TAPT1	None	-2.50812	0.086066	0.320391
337867	UBAC2	None	2.42096	0.081385	0.320391

ER stress acts as a bridge between protein targeting and spermatozoa apoptosis. The enriched GO terms and KEGG pathways corresponded to negative regulation of ER stress response. Other ER stress-related GO terms, such as ER unfolded protein response (UPR), ATF6-mediated UPR, IRE1-mediated UPR, and PERK-mediated UPR were not differentially expressed. However, EIF2A, ATF4, EIF2AK1, and CEBPZ—vital genes in the PERK-EIF2 $\alpha$  ER stress pathway—were differentially expressed. Limitations of RNA-sequencing could result in the inaccurate measurement of underexpressed genes, thus eliminating or underestimating certain ER-related underexpressed genes in differential expression analysis, and consequently interfering with GO and KEGG pathway analyses [57, 58]. Moreover, the current methods of GO and KEGG pathway analyses focus on the number of significant genes but

ignore the participation of genes in different or opposite pathways.

PERK is a type-I transmembrane kinase located in the ER membrane that phosphorylates EIF2 $\alpha$  [59, 60]. Phosphorylated EIF2 $\alpha$  initiates the translation of poorly translated mRNAs such as ATF4 [59, 61]. ATF4, a transcription factor, controls redox homeostasis in the ER, and its translation results not only in amino acid biosynthesis, autophagy, protein folding, but also apoptosis via activation of the proapoptotic factor CEBPZ (growth arrest and DNA damage/CEBP homology protein, the major proapoptotic transcription factor induced by ER stress) [59, 62]. We found that EIF2A, ATF4, EIF2AK1, and CEBPZ were upregulated, indicating activation of the PERK-EIF2 $\alpha$  axis and DNA damage-induced cell death.

**Table 5. Expression profile of crucial genes involved in endoplasmic reticulum stress.**

Gene ID	Gene symbol	Regulation	Fold change	p-Value	Corr P
83939	EIF2A	Upregulation	7.368758342	1.46E-09	2.27E-07
22809	ATF5	Upregulation	6.571856	7.47E-08	3.74E-06
468	ATF4	Upregulation	3.978785	0.034773	0.19102
7494	XBP1	Upregulation	5.115628	1.66E-05	0.000336
27230	SERP1	Upregulation	7.627359469	2.28E-05	0.000439433
487	ATP2A1	Upregulation	2.138295777	2.31E-05	0.000443061
5599	MAPK8	Downregulation	-4.53268	0.001238	0.013433
10153	CEBPZ	Upregulation	9.508285	1.42E-08	1.08E-06
10018	BCL2L11	Upregulation	4.309704	0.000362	0.004786
840	CASP7	Upregulation	4.335948	0.009178	0.069068
4170	MCL1	Upregulation	5.38951	0.003147	0.029303
2081	ERN1	None	-2.00713	0.075075	0.309882
27102	EIF2AK1	Upregulation	5.956155	0.000597	0.007317
9451	EIF2AK3	None	-0.95555	0.134954	0.373771

**Table 6. Expression profile of crucial genes involved in oxidative stress.**

Gene ID	Gene symbol	Regulation	Fold change	p-Value	Corr P
6647	SOD1	Upregulation	8.027872	1.13E-06	3.44E-05
6648	SOD2	Upregulation	4.926199	9.72E-05	0.001556
6649	SOD3	None	0.094519	0.925824	0.944478
27035	NOX1	None	-0.53031	0.716694	0.809006
79400	NOX5	None	0.072599	0.957191	0.970732
10811	NOXA1	Downregulation	-1.71036	0.04837	0.23708
7498	XDH	Downregulation	-4.39865	0.001058	0.011869
1432	MAPK14	Upregulation	4.671004	2.05E-05	0.000401
4779	NFE2L1	Downregulation	-5.56491	2.03E-06	5.60E-05
4780	NFE2L2	None	0.571726	0.750761	0.834931
4790	NFKB1	Upregulation	3.770259	3.26E-05	0.000598
6667	SP1	Upregulation	6.560468	7.33E-07	2.41E-05
25828	TXN2	Upregulation	4.647904	1.06E-07	5.00E-06
2729	GCLC	Upregulation	4.867694	1.73E-07	7.47E-06
3163	HMOX2	Upregulation	4.786777	1.21E-08	9.50E-07
3162	HMOX1	None	-0.12542	0.655116	0.769497
2879	GPX4	Upregulation	5.109184	1.16E-05	0.000246
2876	GPX1	Upregulation	3.138132	8.36E-05	0.001362
2878	GPX3	None	1.266242	0.397058	0.566021
2936	GSR	Upregulation	5.151825	0.007639	0.05978
654817	NCF1C	Downregulation	-3.02159	0.001373	0.014687
653361	NCF1	None	0.043399	0.896427	0.919911
4688	NCF2	Downregulation	-3.89704	0.005773	0.047846
4784	NFIX	Upregulation	4.261149	0.00448	0.03892
2305	FOXMI	Downregulation	-4.02428	0.013965	0.09426

Genes involved in the PERK-EIF2 $\alpha$  pathway are also well-known activators of NF- $\kappa$ B signaling [51, 63], which is involved in the host's innate immunity responses and prevention of apoptosis by repressing CEBPZ expression [51, 64]. However, we observed that

both CEBPZ and NF- $\kappa$ B were upregulated, suggesting a complex mechanism of NF- $\kappa$ B-modulated gene expression and protein synthesis of CEBPZ. The IRE1-TRAF2 pathway can also activate NF- $\kappa$ B signaling, UPR-induced gene XBP1, and apoptosis-promoting

factor p53 suppressor gene JNK1 [65, 66]. We observed that JNK1 was downregulated and XBP1 was upregulated; however, ERN1, the gene encoding TRAF2 and IRE1, was not expressed.

We also identified several pathways related to oxidative stress and neuronal apoptosis. Oxidative stress has been implicated in several neurodegenerative diseases, such as Alzheimer’s disease and Parkinson’s disease. Moreover, ER stress is known to trigger oxidative stress [67]. Thus, ER stress-triggered oxidative stress and consequent apoptosis in sperm cells could be responsible for the manifestations of oligozoospermia. Studies have reported a direct link between protein unfolding and the production of reactive oxygen species (ROS) [67–70]. Protein disulfide isomerase (PDI), endoplasmic reticulum oxidoreductin (ERO1), nuclear factor erythroid 2-related factors (Nrf), and NADPH oxidases (NOXs) are major enzymatic components of ROS production during UPR [67–70]. Furthermore, CEBPZ induces ERO1 and PDI gene family to activate NOXs, consequently enhancing cytosolic ROS production [71, 72] and depleting antioxidants such as glutathione (GSH). We found that both ERO1- $\alpha$  and ERO1- $\beta$ , and PDI genes, such as ERP29, PDIA2, PDIA5, TMX2, and TXNDC12, were upregulated in patients with oligozoospermia. Although we did not observe the expression of NOX family genes, the downstream genes of oxidative stress, including MAPK14, SP1, GPX1, and NOX5, were activated. These findings indicated that the PERK-EIF2 $\alpha$  pathway

triggered ER stress and interfered with ROS generation via the PDI-ERO1 cycle. In addition, PERK phosphorylated and activated transcription factor Nrf2, translocating it to the nucleus to activate the antioxidant response elements and maintain GSH levels [73]. The Nrf2 gene was not differentially expressed in patients with oligozoospermia, suggesting the sperm antioxidant stress response system could be blocked by unfolded proteins-induced pro-survival and pro-apoptotic signaling events.

The ER stress-induced pre-emptive quality control (ERpQC) can selectively degrade ER-targeting proteins by inhibiting their translocation via rerouting or degradation, thus limiting further protein loading into the stressed ER [51, 74–76]. During ER stress, the DERLIN family proteins, the ribosome nascent chain, and signal recognition particle complex (RNC-SRP) are rerouted from the ER translocation pathway to the cytosolic degradation pathway [52, 74–76]. Furthermore, the E3 ligase ubiquitinates the fully translated proteins and effectively transports them to the proteasome by p97 and Bag6 chaperones for degradation by the ER-associated degradation system (ERAD) [77–79]. Several proteins involved in the organization of the retro-translocon channel, such as Sec61, Derlin family proteins, p97, and HRD1 E3 ligase, have been described [74–79]. However, genes encoding these proteins were not expressed in patients with oligozoospermia under PERK-EIF2 $\alpha$  pathway-induced ER stress. This finding suggests that

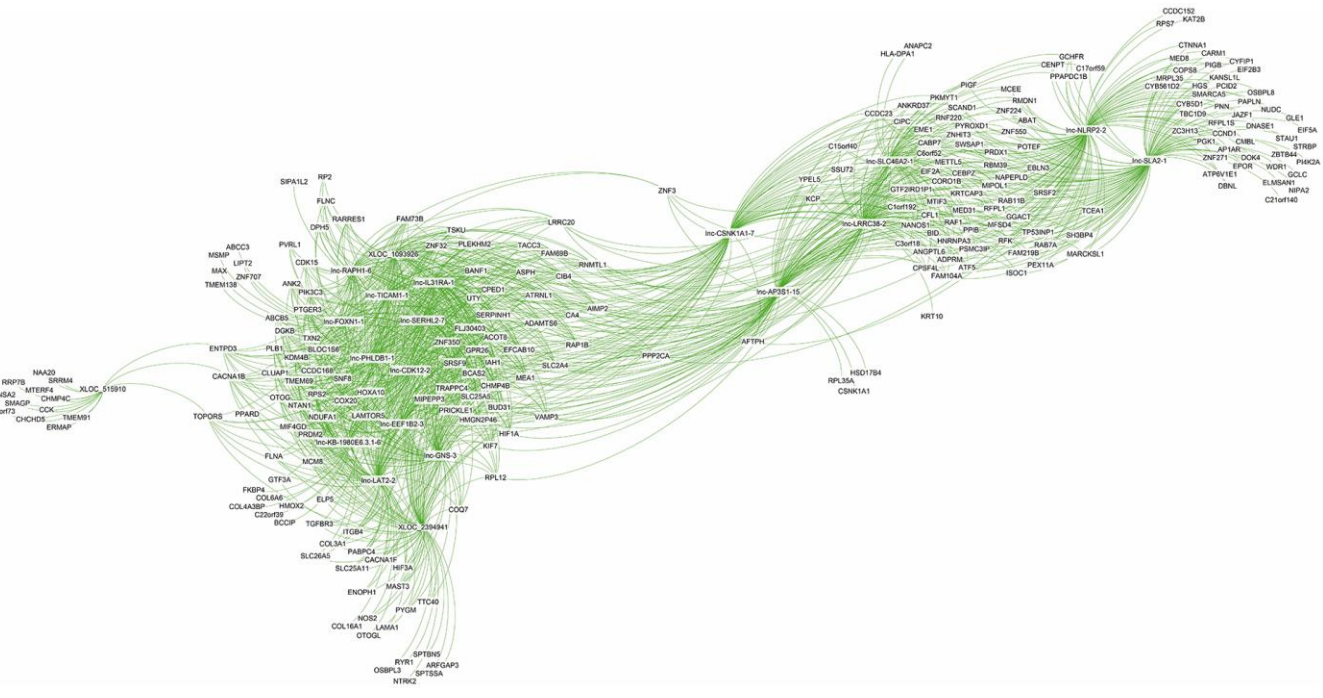


Figure 6. Co-expression subnetwork of the top 20 differentially expressed lncRNAs.

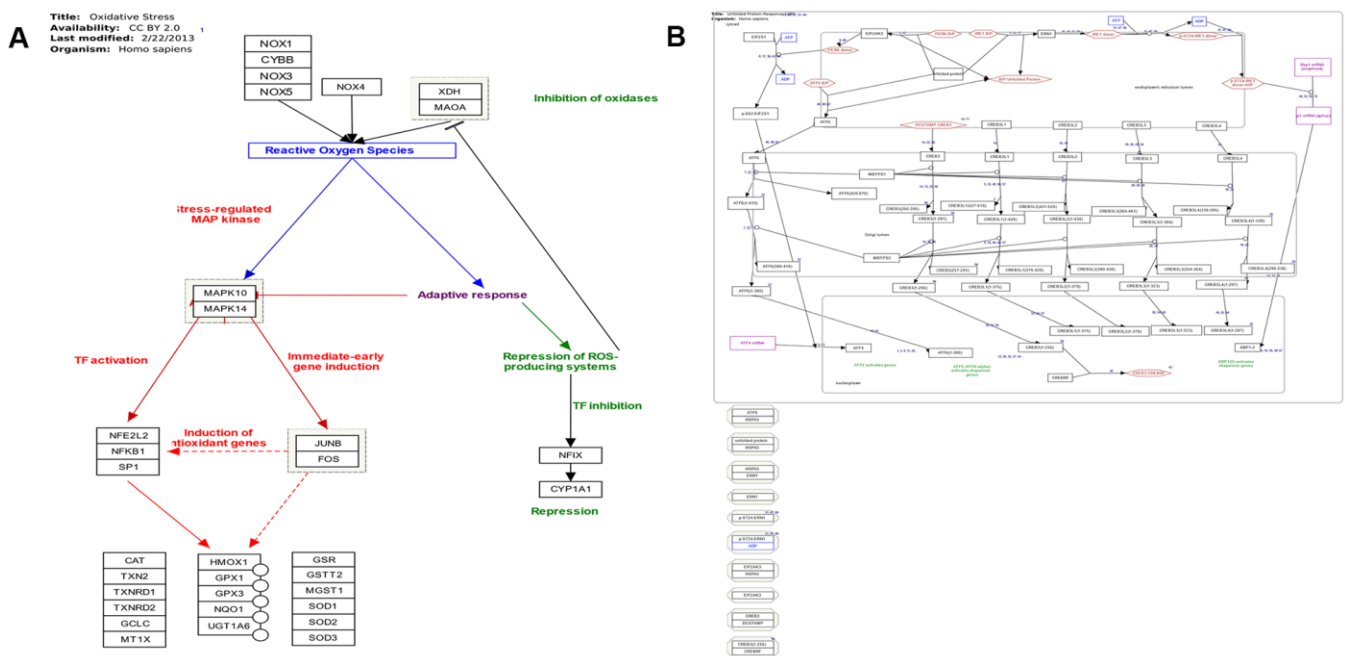
**Table 7. Expression profile of crucial genes involved in spermatogenesis.**

Gene ID	Gene symbol	Regulation	Fold change	p-Value	Corr P
3207	HOXA11	Upregulation	2.630147	0.006865	0.055022
3235	HOXD9	Downregulation	-2.95139	0.034474	0.18968
3973	LHCGR	None	-2.01129	0.0785	0.317254
4214	MAP3K1	Downregulation	-3.50326	0.001029	0.011594
6666	SOX12	Downregulation	-2.88106	0.011734	0.082572
6660	SOX5	Downregulation	-2.5398	0.077814	0.316286
64321	SOX17	Downregulation	-2.39059	0.037155	0.200292
374955	SPATA21	Downregulation	-6.26095	1.07E-06	3.32E-05
65244	SPATS2	Downregulation	-4.60831	0.00102163	0.011516328
402381	SOHLH1	Downregulation	-3.55715	0.066831	0.289663
117	ADCYAP1R1	Downregulation	-3.10056	0.010437	0.075885
7301	TYRO3	Downregulation	-3.14978	0.00333	0.030655
558	AXL	Upregulation	2.947083	5.41E-05	0.000929
6677	SPAM1	Downregulation	-2.01802	0.001119	0.012398
57504	MTA3	Downregulation	-4.65991	1.65E-07	7.20E-06

the ERpQC system could become dysfunctional to protect the sperms under ER stress, resulting in conformational diseases due to accumulation of unfolded proteins.

GO and KEGG pathway analyses indicated that DE genes, such as SRP9, EIF2A, CEBPZ, EIF2AK3, and NFKB1, were mainly enriched in protein targeting, ER stress, and oxidative stress. LncRNAs are transcribed

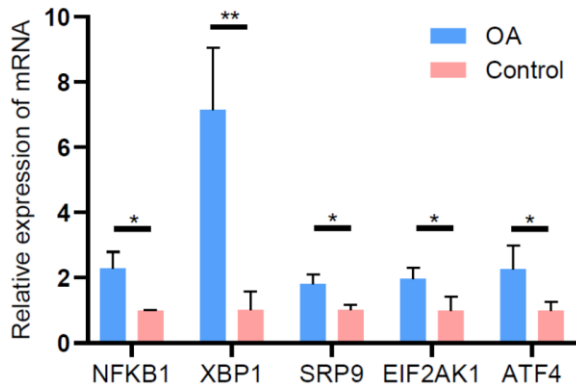
along with the co-expressed genes they regulate [47]. Therefore, the functions of lncRNAs in oligozoospermia could be studied by the function of related target genes. Their cis or trans regulation by lncRNAs, such as EIF2A trans-regulator lnc-FAM198B-1, could regulate spermatogenesis and sperm apoptosis-associated pathways and can be considered as potential diagnostic or therapeutic targets.



**Figure 7. (A)** Pathway map of oxidative stress (adapted from Wikipathway database). **(B)** Pathway map of endoplasmic reticulum stress and unfolded protein response (adapted from Wikipathway database).



Some of the previously reported genes related to spermatogenesis disorders, such as the SOX gene family, SPATA gene family, and TAM receptor genes, were differentially expressed in our study. Consistent

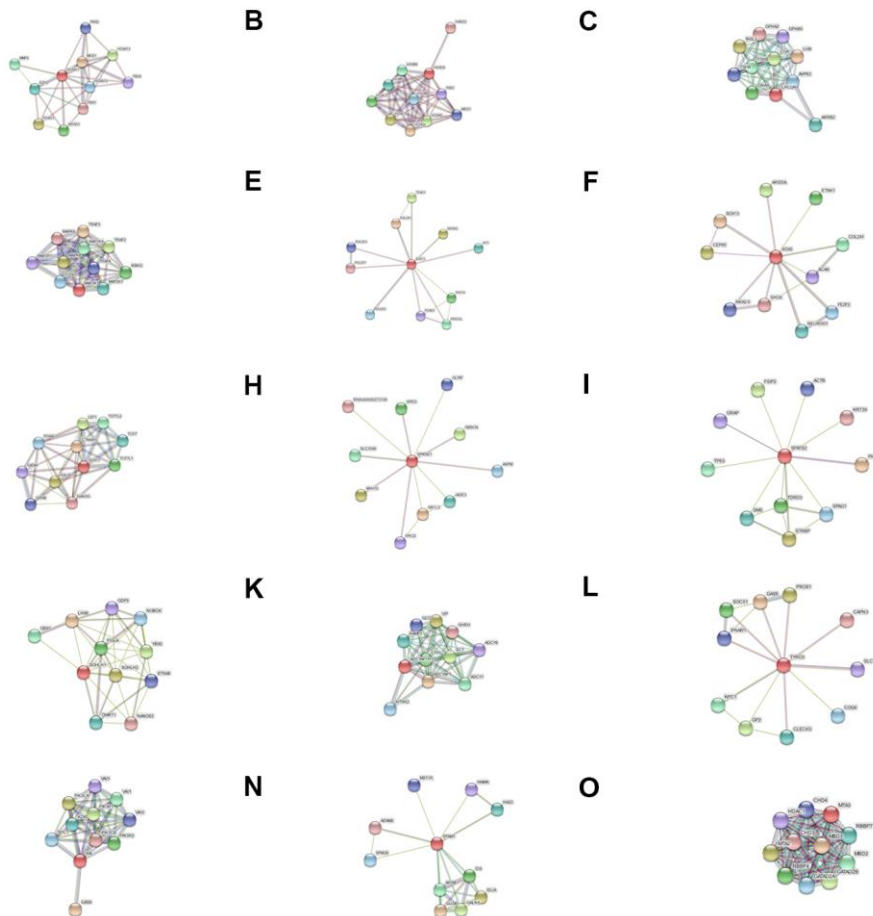


**Figure 8.** Validation of differentially expressed lncRNAs by qRT-PCR.

with the findings of a study by Jiang who reported that the Sox gene family maintains adult male fertility [80], our study found that reduced expression of SOX genes blocked the initiation or progression of spermatogenesis and affected cell maturation. Furthermore, SPATA family genes function in spermatogenesis, sperm maturation, and fertilization. Hypermethylation in the promoter regions of SPATA family genes correlates with oligozoospermia and male infertility [81]. Although SPATS2 and SPATA21 were downregulated in patients with oligozoospermia in our study, altered methylation patterns were not reported. Lu reported that mice knockout for genes (TYRO3, AXL, and MER) were deficient in sperm maturation [82].

## CONCLUSIONS

We identified key lncRNAs, pathways, and gene sub-networks using RNA sequencing datasets that are associated with the pathogenesis of oligozoospermia (Figure 9). The identified lncRNAs and pathways could serve as effective therapeutic targets for male infertility.



**Figure 9.** Protein-protein (PPI) interaction networks for (A) HOXA11, (B) HOXD9, (C) LHCGR, (D) MAP3K1, (E) SOX12, (F) SOX5, (G) SOX17, (H) SPATA21, (I) SPATS2, (J) SOHLH1, (K) ADCYAP1R1, (L) TYRO3, (M) AXL, (N) SPAM1, and (O) MTA3.

## MATERIALS AND METHODS

### Sperm collection and experimental design

All 12 human sperm samples were obtained from Peking University International Hospital from March 2016 to December 2016. This study was approved by the Ethics Committees of Peking University International Hospital and was conducted in accordance with the Declaration of Helsinki. Participants provided their informed consent in writing. Participants were selected by excluding those with abnormal karyotype and Y chromosome microdeletion. The clinical data of all participants are shown in Table 1.

### Collection, analysis, and purification of sperm samples

All human sperm samples were collected and analyzed according to the WHO guidelines using a Makler R chamber (Sefi Laboratories, Tel Aviv, Israel) [17]. Normozoospermic samples were defined as those with concentrations  $\geq 15 \times 10^6/\text{mL}$ . Oligozoospermic samples had a concentration of  $< 15 \times 10^6/\text{mL}$ . All sperm samples were cryopreserved at 1:1 dilution (Irvine Scientific, Santa Ana, CA, USA) in 2 mL cryogenic vials to check the yolk buffer and equilibrated to room temperature for 15 min. Next, all sperm samples were suspended 10 cm above liquid nitrogen for 1 to 2 h and finally transferred to liquid nitrogen. All samples were thawed as follows: cryotubes were transferred to a 37° C water bath and incubated for 5 min. The cell suspension was centrifuged at 500 g for 5 min and the supernatant was discarded. Sperms were purified and the RNA was extracted immediately. The samples were stored at -80° C until further processing.

### RNA extraction and quality control

The RNA was extracted from frozen human sperm samples using an RNA Extraction Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol. The RNA was quantified by NanoDrop ND-1000, and the RNA quality was assessed on Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). The RNA integrity was assessed by running the sample on a standard denaturing agarose gel.

### Library preparation and lncRNA sequencing

Total RNA was isolated using the Trizol reagent (Invitrogen, CA, USA). The RNA quantity and purity were detected by Bioanalyzer 2100 and RNA 6000 Nano LabChip Kit (Agilent Technologies). High-quality RNA (10 µg; RNA integrity number > 7) per sample was obtained by removing ribosomal RNA using the

Ribo-Zero Gold rRNA removal kit (Illumina, San Diego, USA) according to manufacturer's instructions. Approximately 1.5 µg of purified RNA per sample was used to construct the sequencing libraries using the NEBNext Ultra Directional RNA Library Prep Kit from Illumina (NEB, Ipswich, MA, USA) and index codes were added to the corresponding samples. Briefly, the purified RNA was fragmented into small pieces using divalent cations at elevated temperatures. The first-strand cDNA was synthesized using cleaved RNA fragments as templates, random hexamer primers, and M-MuLV reverse transcriptase. Subsequently, the second-strand cDNA was synthesized using Polymerase I and RNase H with reaction buffer (dTTPs were replaced by dUTPs). Exonuclease/polymerase was used to degrade the remaining overhangs and create blunt ends. After adenylation of the 3' ends of DNA fragments, NEBNext Adaptor with a hairpin loop structure was ligated for hybridization. The 300-bp DNA fragments were selected and purified with AMPure XP system (Beckman Coulter, Beverly, MA, USA). Paired-end sequencing was performed on an Illumina HiSeq X-Ten (Novogene Corporation, Beijing, China).

### Identification of lncRNAs and mRNAs

To obtain high-quality lncRNA and mRNA reads, the adaptor sequences and low-quality sequences (N bases > 10% or base quality < 10; greater than 50%) were removed. The remaining reads were aligned to the UCSC (human genome, hg19) using HISAT (v0.1.6). The mapped reads were assembled by StringTie (v1.0.4) separately. Next, all assembled transcripts from samples were merged to reconstruct a comprehensive transcriptome with Cuffcompare (v2.1.1). To obtain highly reliable novel lncRNAs, we first removed the transcripts with a length of less than 200 bp. Transcripts with peak expression of less than 2.0 across all samples and present only in one sample were considered background and removed [18]. The transcripts overlapping with known RNAs (lncRNAs and mRNAs) on the same strand were removed. The remaining transcripts encoding any conserved protein domains were excluded by Pfam database alignment with HMMER. Finally, the transcripts predicted by the lncRNA prediction software (CPC v0.9-r2 and CNCI v2.1) were used for further analysis.

### Analysis of DE lncRNAs and DE mRNAs

HTseq (v0.6.1) feature extraction software was used to calculate the number of readings of each gene (and lncRNA). The R software Limma package was used to normalize the expression data. Genes with low expression were eliminated by filterByExpr function in

edgeR package with parameters: min.count = 0, min.total.count = 15, large.n = 10, min.prop = 0.7. Raw counts were converted to log<sub>2</sub> count per million (CPM) reads. Next, we checked the outliers of the standardized data, calculated the count outside the threshold using the mean and standard deviation method ( $\times \pm 2$  SD) as the outliers, and replaced the count with the predicted value of the trimmed average of all samples. This method is similar to replacing outliers in the DESeq2 package using Cook's distances. Subsequently, the limFit function of the Gaussian regression model in the Limma package was used for model fitting, and DE genes (and lncRNAs) in RNA sequencing data were identified using the empirical Bayes estimation. The regression model was adjusted for age and BMI that could interfere with the analysis. The corrected *p*-value (corr) was calculated based on the Benjamini–Hochberg controlled error detection rate (FDR) [1]. The thresholds of DE genes and lncRNAs were corrected as *p*-value < 0.05 and fold change  $\geq 2$ , respectively.

### GO and KEGG enrichment analyses

To study the biological effects of DE lncRNAs in sperm samples, over-representation analysis method (ora) of R software gene ontology for RNA-seq package (goseq, v1.38.0) and the pathway level analysis method (plage) of gene set variation analysis package (gsva, v1.34.0) were used to determine GO terms and KEGG pathways. Based on GO and KEGG databases, these functions mapped the target genes to the corresponding pathways or biochemical processes. The *p*-value was calculated to indicate the significance of gene enrichment. For both GO and pathway analyses, a *p*-value < 0.05 was considered significant.

### Co-expression network analysis

Compared with the traditional method of constructing a co-expression network using Pearson's correlation coefficient, the Gaussian graphical model (GGM) and weighted correlation network analysis (WGCNA) efficiently removed the pseudo correlation effect on the network structure. We used the R software high-throughput Gauss map model package (Statistical Inference of Large-Scale Gaussian Graphical Model [SILGGM]) to construct the co-expression network of mRNAs and lncRNAs. The identified co-expressed mRNAs and lncRNAs (*p*-value < 0.05) consisted of strong co-expression pairs with a partial correlation coefficient  $1 \geq r > 0.95$  or  $-1 \leq r < -0.95$ , moderate co-expression pairs with a partial correlation coefficient  $0.95 \geq r > 0.80$  or  $-0.95 \leq r < -0.80$ , and weak co-expression pairs with a partial correlation coefficient  $0.8 \geq r$  or  $-0.8 \leq r$ . Further, we classified the co-expressed lncRNAs with a partial correlation coefficient > 0.95 or

< -0.95 through cis-trans analysis. lncRNAs having flanked genes within 300 kb were considered as cis-acting lncRNAs, and those with lncRNAs-gene pair sequence homology (blastn, E-value < 1.0E-10 and identity > 99 and matched length  $\geq 20$  bp) were considered as trans-acting lncRNAs.

### QPCR analysis

Isolated RNA was reverse-transcribed to cDNA using a Reverse Transcription Kit (Takara, Dalian, China). The qRT-PCR analyses were performed using a StepOnePlus RT-PCR Instrument with Power SYBR Green (Takara, Dalian, China). The qRT-PCR conditions were as follows: 95° C for 2 minutes, followed by 40 cycles of 95° C for 15 seconds and 60° C for 30 seconds. All experiments were performed and analyzed in triplicate. The primers used in this study were listed in Table 1. Then, lncRNA expression levels were normalized to GAPDH and calculated using the 2<sup>-ΔΔCt</sup> method.

### Statistical analysis

Data are expressed as mean  $\pm$  standard error of the mean (SEM). The *p*-values < 0.05 were considered significant. Graphs were plotted and analyzed using GraphPad Prism version 6.0 (GraphPad Software, La Jolla, CA, USA).

### Availability of data and materials

Data generated during the study and presented in the manuscript are available from the corresponding author upon request.

### Ethics approval and consent to participate

All experiments involving human sperm samples in this study were approved by the Ethics Committees of Peking University International Hospital. The participants provided their informed consent in writing.

### Editorial note

\*This corresponding author has a verified history of publications using a personal email address for correspondence.

### Abbreviations

RNA-seq: RNA sequencing; ncRNA: non-coding RNA; lncRNAs: long non-coding RNAs; miRNAs: microRNAs; piRNA: PIWI-interacting RNA; 3' UTR: 3' untranslated region; GO: gene ontology; KEGG:

Kyoto Encyclopedia of Genes and Genomes; DE: differentially expressed; NS: normozoospermic sample; OS: oligozoospermic sample; SRP9: signal-recognition particle 9; BMI: body mass index; DFI: DNA fragmentation index; FSH: follicle-stimulating hormone; LH: luteinizing hormone; T: testosterone; BP: biological process; CC: cell component; MF: molecular function.

## AUTHOR CONTRIBUTIONS

Li Tian, Ya-Peng Wang, and Shou-Long Deng conceived the study and designed the experiments. Tie-Cheng Sun and Kun Yu performed the experiments. Yi Zhang and Tie-Cheng Sun analyzed the data. Tie-Cheng Sun, Shan-Jie Zhou, Hong Yu, and Yi Zhang contributed reagents/materials/analysis tools. Tie-Cheng Sun, Yi Zhang, and Kun Yu wrote the manuscript.

## ACKNOWLEDGMENTS

We are grateful to Xin-Rong He, Xiao-Yuan Wu, and Jia-Long Liang for their technical assistance.

## CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

## FUNDING

This study was supported by the Peking Postdoctoral Research Fund (EE2019-50), Peking University International Hospital Research Funds (No. YN2019QN13), and the National Natural Science Foundation of China (No. 81601275).

## REFERENCES

1. Agarwal A, Mulgund A, Hamada A, Chyatte MR. A unique view on male infertility around the globe. *Reprod Biol Endocrinol*. 2015; 13:37. <https://doi.org/10.1186/s12958-015-0032-1> PMID:[25928197](https://pubmed.ncbi.nlm.nih.gov/25928197/)
2. Chen H, Pu XY, Zhang RP, A ZC. Association of common SNP rs1136410 in PARP1 gene with the susceptibility to male infertility with oligospermia. *J Assist Reprod Genet*. 2014; 31:1391–95. <https://doi.org/10.1007/s10815-014-0311-3> PMID:[25106941](https://pubmed.ncbi.nlm.nih.gov/25106941/)
3. Gu A, Ji G, Shi X, Long Y, Xia Y, Song L, Wang S, Wang X. Genetic variants in Piwi-interacting RNA pathway genes confer susceptibility to spermatogenic failure in a Chinese population. *Hum Reprod*. 2010; 25:2955–61. <https://doi.org/10.1093/humrep/deq274>

- PMID:[20940137](https://pubmed.ncbi.nlm.nih.gov/20940137/)
4. Montjean D, De La Grange P, Gentien D, Rapinat A, Belloc S, Cohen-Bacrie P, Menezo Y, Benkhalifa M. Sperm transcriptome profiling in oligozoospermia. *J Assist Reprod Genet*. 2012; 29:3–10. <https://doi.org/10.1007/s10815-011-9644-3> PMID:[21989496](https://pubmed.ncbi.nlm.nih.gov/21989496/)
5. Zou T, Liu X, Ding S, Xing J. Evaluation of sperm mitochondrial function using rh123/PI dual fluorescent staining in asthenospermia and oligoasthenozoospermia. *J Biomed Res*. 2010; 24:404–10. [https://doi.org/10.1016/S1674-8301\(10\)60054-1](https://doi.org/10.1016/S1674-8301(10)60054-1) PMID:[23554656](https://pubmed.ncbi.nlm.nih.gov/23554656/)
6. Zhang WD, Zhang Z, Jia LT, Zhang LL, Fu T, Li YS, Wang P, Sun L, Shi Y, Zhang HZ. Oxygen free radicals and mitochondrial signaling in oligospermia and asthenospermia. *Mol Med Rep*. 2014; 10:1875–80. <https://doi.org/10.3892/mmr.2014.2428> PMID:[25109708](https://pubmed.ncbi.nlm.nih.gov/25109708/)
7. McLachlan RI. Approach to the patient with oligozoospermia. *J Clin Endocrinol Metab*. 2013; 98:873–80. <https://doi.org/10.1210/jc.2012-3650> PMID:[23472228](https://pubmed.ncbi.nlm.nih.gov/23472228/)
8. Liu DY, Baker HW. Defective sperm-zona pellucida interaction: a major cause of failure of fertilization in clinical *in-vitro* fertilization. *Hum Reprod*. 2000; 15:702–08. <https://doi.org/10.1093/humrep/15.3.702> PMID:[10686223](https://pubmed.ncbi.nlm.nih.gov/10686223/)
9. Balkan M, Tekes S, Gedik A. Cytogenetic and Y chromosome microdeletion screening studies in infertile males with oligozoospermia and azoospermia in southeast Turkey. *J Assist Reprod Genet*. 2008; 25:559–65. <https://doi.org/10.1007/s10815-008-9272-8> PMID:[18953646](https://pubmed.ncbi.nlm.nih.gov/18953646/)
10. Zhang X, Chen F, Huang Z. Apoptosis induced by acrylamide is suppressed in a 21.5% fat diet through caspase-3-independent pathway in mice testis. *Toxicol Mech Methods*. 2009; 19:219–24. <https://doi.org/10.1080/15376510802499048> PMID:[19750022](https://pubmed.ncbi.nlm.nih.gov/19750022/)
11. Kajihara T, Okagaki R, Ishihara O. LPS-induced transient testicular dysfunction accompanied by apoptosis of testicular germ cells in mice. *Med Mol Morphol*. 2006; 39:203–08. <https://doi.org/10.1007/s00795-006-0334-7> PMID:[17187183](https://pubmed.ncbi.nlm.nih.gov/17187183/)
12. McSwiggin HM, O'Doherty AM. Epigenetic reprogramming during spermatogenesis and male factor infertility. *Reproduction*. 2018; 156:R9–R21.



- <https://doi.org/10.1530/REP-18-0009> PMID:[29717022](https://pubmed.ncbi.nlm.nih.gov/29717022/)
13. O'Doherty AM, McGettigan PA. Epigenetic processes in the male germline. *Reprod Fertil Dev.* 2015; 27:725–38.  
<https://doi.org/10.1071/RD14167>  
PMID:[25200708](https://pubmed.ncbi.nlm.nih.gov/25200708/)
  14. Laurentino S, Beygo J, Nordhoff V, Kliesch S, Wistuba J, Borgmann J, Buiting K, Horsthemke B, Gromoll J. Epigenetic germline mosaicism in infertile men. *Hum Mol Genet.* 2015; 24:1295–304.  
<https://doi.org/10.1093/hmg/ddu540> PMID:[25336341](https://pubmed.ncbi.nlm.nih.gov/25336341/)
  15. Stuppia L, Franzago M, Ballerini P, Gatta V, Antonucci I. Epigenetics and male reproduction: the consequences of paternal lifestyle on fertility, embryo development, and children lifetime health. *Clin Epigenetics.* 2015; 7:120.  
<https://doi.org/10.1186/s13148-015-0155-4>  
PMID:[26566402](https://pubmed.ncbi.nlm.nih.gov/26566402/)
  16. Urdinguio RG, Bayón GF, Dmitrijeva M, Toraño EG, Bravo C, Fraga MF, Bassas L, Larriba S, Fernández AF. Aberrant DNA methylation patterns of spermatozoa in men with unexplained infertility. *Hum Reprod.* 2015; 30:1014–28.  
<https://doi.org/10.1093/humrep/dev053>  
PMID:[25753583](https://pubmed.ncbi.nlm.nih.gov/25753583/)
  17. Schagdarsurengin U, Steger K. Epigenetics in male reproduction: effect of paternal diet on sperm quality and offspring health. *Nat Rev Urol.* 2016; 13:584–95.  
<https://doi.org/10.1038/nrurol.2016.157>  
PMID:[27578043](https://pubmed.ncbi.nlm.nih.gov/27578043/)
  18. Kropp J, Carrillo JA, Namous H, Daniels A, Salih SM, Song J, Khatib H. Male fertility status is associated with DNA methylation signatures in sperm and transcriptomic profiles of bovine preimplantation embryos. *BMC Genomics.* 2017; 18:280.  
<https://doi.org/10.1186/s12864-017-3673-y>  
PMID:[28381255](https://pubmed.ncbi.nlm.nih.gov/28381255/)
  19. Laqqan M, Solomayer EF, Hammadeh M. Aberrations in sperm DNA methylation patterns are associated with abnormalities in semen parameters of subfertile males. *Reprod Biol.* 2017; 17:246–51.  
<https://doi.org/10.1016/j.repbio.2017.05.010>  
PMID:[28566207](https://pubmed.ncbi.nlm.nih.gov/28566207/)
  20. Nasri F, Gharesi-Fard B, Namavar Jahromi B, Farazi-Fard MA, Banaei M, Davari M, Ebrahimi S, Anvar Z. Sperm DNA methylation of H19 imprinted gene and male infertility. *Andrologia.* 2017; 49:e12766.  
<https://doi.org/10.1111/and.12766>  
PMID:[28295500](https://pubmed.ncbi.nlm.nih.gov/28295500/)
  21. Holliday R. The inheritance of epigenetic defects. *Science.* 1987; 238:163–70.  
<https://doi.org/10.1126/science.3310230>  
PMID:[3310230](https://pubmed.ncbi.nlm.nih.gov/3310230/)
  22. Stewart KR, Veselovska L, Kelsey G. Establishment and functions of DNA methylation in the germline. *Epigenomics.* 2016; 8:1399–413.  
<https://doi.org/10.2217/epi-2016-0056>  
PMID:[27659720](https://pubmed.ncbi.nlm.nih.gov/27659720/)
  23. Luk AC, Chan WY, Rennert OM, Lee TL. Long noncoding RNAs in spermatogenesis: insights from recent high-throughput transcriptome studies. *Reproduction.* 2014; 147:R131–41.  
<https://doi.org/10.1530/REP-13-0594> PMID:[24713396](https://pubmed.ncbi.nlm.nih.gov/24713396/)
  24. Berghoff EG, Clark MF, Chen S, Cajigas I, Leib DE, Kohtz JD. Evt2 (Dlx6as) lncRNA regulates ultraconserved enhancer methylation and the differential transcriptional control of adjacent genes. *Development.* 2013; 140:4407–16.  
<https://doi.org/10.1242/dev.099390>  
PMID:[24089468](https://pubmed.ncbi.nlm.nih.gov/24089468/)
  25. Yap KL, Li S, Muñoz-Cabello AM, Raguz S, Zeng L, Mujtaba S, Gil J, Walsh MJ, Zhou MM. Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. *Mol Cell.* 2010; 38:662–74.  
<https://doi.org/10.1016/j.molcel.2010.03.021>  
PMID:[20541999](https://pubmed.ncbi.nlm.nih.gov/20541999/)
  26. Klattenhoff CA, Scheuermann JC, Surface LE, Bradley RK, Fields PA, Steinhauser ML, Ding H, Butty VL, Torrey L, Haas S, Abo R, Tabebordbar M, Lee RT, et al. Braveheart, a long noncoding RNA required for cardiovascular lineage commitment. *Cell.* 2013; 152:570–83.  
<https://doi.org/10.1016/j.cell.2013.01.003>  
PMID:[23352431](https://pubmed.ncbi.nlm.nih.gov/23352431/)
  27. Robles V, Valcarce DG, Riesco MF. Non-coding RNA regulation in reproduction: their potential use as biomarkers. *Noncoding RNA Res.* 2019; 4:54–62.  
<https://doi.org/10.1016/j.ncrna.2019.04.001>  
PMID:[31193491](https://pubmed.ncbi.nlm.nih.gov/31193491/)
  28. Zhang J, Liu W, Jin Y, Jia P, Jia K, Yi M. MiR-202-5p is a novel germ plasm-specific microRNA in zebrafish. *Sci Rep.* 2017; 7:7055.  
<https://doi.org/10.1038/s41598-017-07675-x>  
PMID:[28765643](https://pubmed.ncbi.nlm.nih.gov/28765643/)
  29. von Meyenn F, Berrens RV, Andrews S, Santos F, Collier AJ, Krueger F, Osorno R, Dean W, Rugg-Gunn PJ, Reik W. Comparative principles of DNA methylation reprogramming during human and mouse *in vitro* primordial germ cell specification. *Dev Cell.* 2016; 39:104–15.  
<https://doi.org/10.1016/j.devcel.2016.09.015>  
PMID:[27728778](https://pubmed.ncbi.nlm.nih.gov/27728778/)



30. Lü M, Tian H, Cao YX, He X, Chen L, Song X, Ping P, Huang H, Sun F. Downregulation of miR-320a/383-sponge-like long non-coding RNA NLC1-C (narcolepsy candidate-region 1 genes) is associated with male infertility and promotes testicular embryonal carcinoma cell proliferation. *Cell Death Dis.* 2015; 6:e1960.  
<https://doi.org/10.1038/cddis.2015.267>  
PMID:[26539909](https://pubmed.ncbi.nlm.nih.gov/26539909/)
31. Ferlita A, Battaglia R, Andronico F, Caruso S, Cianci A, Purrello M, Pietro CD. Non-coding RNAs in endometrial physiopathology. *Int J Mol Sci.* 2018; 19:2120.  
<https://doi.org/10.3390/ijms19072120>  
PMID:[30037059](https://pubmed.ncbi.nlm.nih.gov/30037059/)
32. Panoutsopoulou K, Avgeris M, Scorilas A. miRNA and long non-coding RNA: molecular function and clinical value in breast and ovarian cancers. *Expert Rev Mol Diagn.* 2018; 18:963–79.  
<https://doi.org/10.1080/14737159.2018.1538794>  
PMID:[30338716](https://pubmed.ncbi.nlm.nih.gov/30338716/)
33. Diez-Fraile A, Lammens T, Tilleman K, Witkowski W, Verhasselt B, De Sutter P, Benoit Y, Espeel M, D'Herde K. Age-associated differential microRNA levels in human follicular fluid reveal pathways potentially determining fertility and success of *in vitro* fertilization. *Hum Fertil (Camb).* 2014; 17:90–98.  
<https://doi.org/10.3109/14647273.2014.897006>  
PMID:[24684237](https://pubmed.ncbi.nlm.nih.gov/24684237/)
34. Taylor DH, Chu ET, Spektor R, Soloway PD. Long non-coding RNA regulation of reproduction and development. *Mol Reprod Dev.* 2015; 82:932–56.  
<https://doi.org/10.1002/mrd.22581> PMID:[26517592](https://pubmed.ncbi.nlm.nih.gov/26517592/)
35. Zhang C, Gao L, Xu EY. LncRNA, a new component of expanding RNA-protein regulatory network important for animal sperm development. *Semin Cell Dev Biol.* 2016; 59:110–17.  
<https://doi.org/10.1016/j.semcd.2016.06.013>  
PMID:[27345292](https://pubmed.ncbi.nlm.nih.gov/27345292/)
36. Magnúsdóttir E, Surani MA. How to make a primordial germ cell. *Development.* 2014; 141:245–52.  
<https://doi.org/10.1242/dev.098269> PMID:[24381195](https://pubmed.ncbi.nlm.nih.gov/24381195/)
37. Wang X, Yang C, Guo F, Zhang Y, Ju Z, Jiang Q, Zhao X, Liu Y, Zhao H, Wang J, Sun Y, Wang C, Zhu H, Huang J. 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.  
<https://doi.org/10.1038/s41598-018-38462-x>  
PMID:[30765858](https://pubmed.ncbi.nlm.nih.gov/30765858/)
38. Medeiros LA, Dennis LM, Gill ME, Houbaviv H, Markoulaki S, Fu D, White AC, Kirak O, Sharp PA, Page DC, Jaenisch R. Mir-290-295 deficiency in mice results in partially penetrant embryonic lethality and germ cell defects. *Proc Natl Acad Sci U S A.* 2011; 108:14163–68.  
<https://doi.org/10.1073/pnas.1111241108>  
PMID:[21844366](https://pubmed.ncbi.nlm.nih.gov/21844366/)
39. Takada S, Berezikov E, Choi YL, Yamashita Y, Mano H. Potential role of miR-29b in modulation of Dnmt3a and Dnmt3b expression in primordial germ cells of female mouse embryos. *RNA.* 2009; 15:1507–14.  
<https://doi.org/10.1261/rna.1418309>  
PMID:[19509302](https://pubmed.ncbi.nlm.nih.gov/19509302/)
40. Dallas PB, Gottardo NG, Firth MJ, Beesley AH, Hoffmann K, Terry PA, Freitas JR, Boag JM, Cummings AJ, Kees UR. Gene expression levels assessed by oligonucleotide microarray analysis and quantitative real-time RT-PCR -- how well do they correlate? *BMC Genomics.* 2005; 6:59.  
<https://doi.org/10.1186/1471-2164-6-59>  
PMID:[15854232](https://pubmed.ncbi.nlm.nih.gov/15854232/)
41. Arikawa E, Sun Y, Wang J, Zhou Q, Ning B, Dial SL, Guo L, Yang J. Cross-platform comparison of SYBR green real-time PCR with TaqMan PCR, microarrays and other gene expression measurement technologies evaluated in the MicroArray quality control (MAQC) study. *BMC Genomics.* 2008; 9:328.  
<https://doi.org/10.1186/1471-2164-9-328>  
PMID:[18620571](https://pubmed.ncbi.nlm.nih.gov/18620571/)
42. Morey JS, Ryan JC, Van Dolah FM. Microarray validation: factors influencing correlation between oligonucleotide microarrays and real-time PCR. *Biol Proced Online.* 2006; 8:175–93.  
<https://doi.org/10.1251/bpo126> PMID:[17242735](https://pubmed.ncbi.nlm.nih.gov/17242735/)
43. Alvarez ML, Doné SC. SYBR® Green and TaqMan® quantitative PCR arrays: expression profile of genes relevant to a pathway or a disease state. *Methods Mol Biol.* 2014; 1182:321–59.  
[https://doi.org/10.1007/978-1-4939-1062-5\\_27](https://doi.org/10.1007/978-1-4939-1062-5_27)  
PMID:[25055922](https://pubmed.ncbi.nlm.nih.gov/25055922/)
44. Zhou X, Li Y, Wang W, Wang S, Hou J, Zhang A, Lv B, Gao C, Yan Z, Pang D, Lu K, Ahmad NH, Wang L, et al. Regulation of Hippo/YAP signaling and Esophageal Squamous Carcinoma progression by an E3 ubiquitin ligase PARK2. *Theranostics.* 2020; 10:9443–57.  
<https://doi.org/10.7150/thno.46078> PMID:[32863938](https://pubmed.ncbi.nlm.nih.gov/32863938/)
45. Bache I, Assche EV, Cingoz S, Bugge M, Tümer Z, Hjorth M, Lundsteen C, Lespinasse J, Winther K, Niebuhr A, Kalscheuer V, Liebaers I, Bonduelle M, et al. An excess of chromosome 1 breakpoints in male infertility. *Eur J Hum Genet.* 2004; 12:993–1000.  
<https://doi.org/10.1038/sj.ejhg.5201263>  
PMID:[15367911](https://pubmed.ncbi.nlm.nih.gov/15367911/)
46. Balasar Ö, Zamani AG, Balasar M, Acar H. Male infertility associated with de novo pericentric inversion of chromosome 1. *Turk J Urol.* 2017; 43:560–62.

- <https://doi.org/10.5152/tud.2017.79346>  
PMID:29201526
47. Wu C, Liu H, Zhang F, Shao W, Yang L, Ning Y, Wang S, Zhao G, Lee BJ, Lammi M, Guo X. Long noncoding RNA expression profile reveals lncRNAs signature associated with extracellular matrix degradation in kashin-beck disease. *Sci Rep.* 2017; 7:17553.  
<https://doi.org/10.1038/s41598-017-17875-0>  
PMID:29242531
48. Ulitsky I, Bartel DP. lincRNAs: genomics, evolution, and mechanisms. *Cell.* 2013; 154:26–46.  
<https://doi.org/10.1016/j.cell.2013.06.020>  
PMID:23827673
49. Amin V, Harris RA, Onuchic V, Jackson AR, Charnecki T, Paithankar S, Lakshmi Subramanian S, Riehle K, Coarfa C, Milosavljevic A. Epigenomic footprints across 111 reference epigenomes reveal tissue-specific epigenetic regulation of lincRNAs. *Nat Commun.* 2015; 6:6370.  
<https://doi.org/10.1038/ncomms7370> PMID:25691256
50. Rispler MJ, Lue YH, Sinha-Hikim AP, Leung A, Wang C, Swerdloff RS. Frequency of apoptosis in spontaneous and hormonally induced oligospermia in human ejaculated spermatazoa. *Fertil Steril.* 2002; 78:S265.  
[https://doi.org/10.1016/S0015-0282\(02\)04172-9](https://doi.org/10.1016/S0015-0282(02)04172-9)
51. Salminen A, Kauppinen A, Suuronen T, Kaarniranta K, Ojala J. ER stress in Alzheimer's disease: a novel neuronal trigger for inflammation and Alzheimer's pathology. *J Neuroinflammation.* 2009; 6:41.  
<https://doi.org/10.1186/1742-2094-6-41>  
PMID:20035627
52. Kadowaki H, Nishitoh H. Endoplasmic reticulum quality control by garbage disposal. *FEBS J.* 2019; 286:232–40.  
<https://doi.org/10.1111/febs.14589> PMID:29923316
53. Schröder M, Kaufman RJ. The mammalian unfolded protein response. *Annu Rev Biochem.* 2005; 74:739–89.  
<https://doi.org/10.1146/annurev.biochem.73.011303.074134> PMID:15952902
54. Bernales S, Papa FR, Walter P. Intracellular signaling by the unfolded protein response. *Annu Rev Cell Dev Biol.* 2006; 22:487–508.  
<https://doi.org/10.1146/annurev.cellbio.21.122303.120200> PMID:16822172
55. Ron D, Walter P. Signal integration in the endoplasmic reticulum unfolded protein response. *Nat Rev Mol Cell Biol.* 2007; 8:519–29.  
<https://doi.org/10.1038/nrm2199> PMID:17565364
56. Kim I, Xu W, Reed JC. Cell death and endoplasmic reticulum stress: disease relevance and therapeutic opportunities. *Nat Rev Drug Discov.* 2008; 7:1013–30.  
<https://doi.org/10.1038/nrd2755>  
PMID:19043451
57. Kanitz A, Gypas F, Gruber AJ, Gruber AR, Martin G, Zavolan M. Comparative assessment of methods for the computational inference of transcript isoform abundance from RNA-seq data. *Genome Biol.* 2015; 16:150.  
<https://doi.org/10.1186/s13059-015-0702-5>  
PMID:26201343
58. Rozpedek W, Pytel D, Mucha B, Leszczynska H, Diehl JA, Majsterek I. The Role of the PERK/eIF2 $\alpha$ /ATF4/CHOP Signaling Pathway in Tumor Progression During Endoplasmic Reticulum Stress. *Curr Mol Med.* 2016; 16:533–44.  
<https://doi.org/10.2174/1566524016666160523143937> PMID:27211800
59. Gerakis Y, Hetz C. Emerging roles of ER stress in the etiology and pathogenesis of Alzheimer's disease. *FEBS J.* 2018; 285:995–1011.  
<https://doi.org/10.1111/febs.14332>  
PMID:29148236
60. Harding HP, Zhang Y, Ron D. Protein translation and folding are coupled by an endoplasmic-reticulum-resident kinase. *Nature.* 1999; 397:271–74.  
<https://doi.org/10.1038/16729> PMID:9930704
61. Spriggs KA, Bushell M, Willis AE. Translational regulation of gene expression during conditions of cell stress. *Mol Cell.* 2010; 40:228–37.  
<https://doi.org/10.1016/j.molcel.2010.09.028>  
PMID:20965418
62. Vattem KM, Wek RC. Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells. *Proc Natl Acad Sci USA.* 2004; 101:11269–74.  
<https://doi.org/10.1073/pnas.0400541101>  
PMID:15277680
63. Deng J, Lu PD, Zhang Y, Scheuner D, Kaufman RJ, Sonenberg N, Harding HP, Ron D. Translational repression mediates activation of nuclear factor kappa B by phosphorylated translation initiation factor 2. *Mol Cell Biol.* 2004; 24:10161–68.  
<https://doi.org/10.1128/MCB.24.23.10161-10168.2004> PMID:15542827
64. Nozaki S, Sledge GW Jr, Nakshatri H. Repression of GADD153/CHOP by NF-kappaB: a possible cellular defense against endoplasmic reticulum stress-induced cell death. *Oncogene.* 2001; 20:2178–85.  
<https://doi.org/10.1038/sj.onc.1204292>  
PMID:11360202
65. Urano F, Wang X, Bertolotti A, Zhang Y, Chung P, Harding HP, Ron D. Coupling of stress in the ER to activation of JNK protein kinases by transmembrane protein kinase IRE1. *Science.* 2000; 287:664–66.

<https://doi.org/10.1126/science.287.5453.664>

PMID:[10650002](#)

66. Hu P, Han Z, Couvillon AD, Kaufman RJ, Exton JH. Autocrine tumor necrosis factor alpha links endoplasmic reticulum stress to the membrane death receptor pathway through IRE1alpha-mediated NF-kappaB activation and down-regulation of TRAF2 expression. *Mol Cell Biol.* 2006; 26:3071–84.  
<https://doi.org/10.1128/MCB.26.8.3071-3084.2006>  
PMID:[16581782](#)
67. Cao SS, Kaufman RJ. Endoplasmic reticulum stress and oxidative stress in cell fate decision and human disease. *Antioxid Redox Signal.* 2014; 21:396–413.  
<https://doi.org/10.1089/ars.2014.5851>  
PMID:[24702237](#)
68. Ferreira E, Baldeiras I, Ferreira IL, Costa RO, Rego AC, Pereira CF, Oliveira CR. Mitochondrial- and endoplasmic reticulum-associated oxidative stress in Alzheimer's disease: from pathogenesis to biomarkers. *Int J Cell Biol.* 2012; 2012:735206.  
<https://doi.org/10.1155/2012/735206>  
PMID:[22701485](#)
69. Anelli T, Alessio M, Mezghrani A, Simmen T, Talamo F, Bachi A, Sitia R. ERp44, a novel endoplasmic reticulum folding assistant of the thioredoxin family. *EMBO J.* 2002; 21:835–44.  
<https://doi.org/10.1093/emboj/21.4.835>  
PMID:[11847130](#)
70. Otsu M, Bertoli G, Fagioli C, Guerini-Rocco E, Nerini-Molteni S, Ruffato E, Sitia R. Dynamic retention of Ero1alpha and Ero1beta in the endoplasmic reticulum by interactions with PDI and ERp44. *Antioxid Redox Signal.* 2006; 8:274–82.  
<https://doi.org/10.1089/ars.2006.8.274>  
PMID:[16677073](#)
71. Li G, Scull C, Ozcan L, Tabas I. NADPH oxidase links endoplasmic reticulum stress, oxidative stress, and PKR activation to induce apoptosis. *J Cell Biol.* 2010; 191:1113–25.  
<https://doi.org/10.1083/jcb.201006121>  
PMID:[21135141](#)
72. de la Monte SM, Wands JR. Molecular indices of oxidative stress and mitochondrial dysfunction occur early and often progress with severity of Alzheimer's disease. *J Alzheimers Dis.* 2006; 9:167–81.  
<https://doi.org/10.3233/jad-2006-9209>  
PMID:[16873964](#)
73. Cullinan SB, Zhang D, Hannink M, Arvisais E, Kaufman RJ, Diehl JA. Nrf2 is a direct PERK substrate and effector of PERK-dependent cell survival. *Mol Cell Biol.* 2003; 23:7198–209.  
<https://doi.org/10.1128/mcb.23.20.7198-7209.2003>  
PMID:[14517290](#)
74. Oyadomari S, Yun C, Fisher EA, Kreglinger N, Kreibich G, Oyadomari M, Harding HP, Goodman AG, Harant H, Garrison JL, Taunton J, Katze MG, Ron D. Cotranslocational degradation protects the stressed endoplasmic reticulum from protein overload. *Cell.* 2006; 126:727–39.  
<https://doi.org/10.1016/j.cell.2006.06.051>  
PMID:[16923392](#)
75. Orsi A, Fioriti L, Chiesa R, Sitia R. Conditions of endoplasmic reticulum stress favor the accumulation of cytosolic prion protein. *J Biol Chem.* 2006; 281:30431–38.  
<https://doi.org/10.1074/jbc.M605320200>  
PMID:[16908519](#)
76. Kang SW, Rane NS, Kim SJ, Garrison JL, Taunton J, Hegde RS. Substrate-specific translocational attenuation during ER stress defines a pre-emptive quality control pathway. *Cell.* 2006; 127:999–1013.  
<https://doi.org/10.1016/j.cell.2006.10.032>  
PMID:[17129784](#)
77. Kadowaki H, Nagai A, Maruyama T, Takami Y, Satrimafitrah P, Kato H, Honda A, Hatta T, Natsume T, Sato T, Kai H, Ichijo H, Nishitoh H. Pre-emptive quality control protects the ER from protein overload via the proximity of ERAD components and SRP. *Cell Rep.* 2015; 13:944–56.  
<https://doi.org/10.1016/j.celrep.2015.09.047>  
PMID:[26565908](#)
78. Braunstein I, Zach L, Allan S, Kalies KU, Stanhill A. Proteasomal degradation of preemptive quality control (pQC) substrates is mediated by an AIRAPL-p97 complex. *Mol Biol Cell.* 2015; 26:3719–27.  
<https://doi.org/10.1091/mbc.E15-02-0085>  
PMID:[26337389](#)
79. Kadowaki H, Satrimafitrah P, Takami Y, Nishitoh H. Molecular mechanism of ER stress-induced pre-emptive quality control involving association of the translocon, Derlin-1, and HRD1. *Sci Rep.* 2018; 8:7317.  
<https://doi.org/10.1038/s41598-018-25724-x>  
PMID:[29743537](#)
80. Jiang T, Hou CC, She ZY, Yang WX. The SOX gene family: function and regulation in testis determination and male fertility maintenance. *Mol Biol Rep.* 2013; 40:2187–94.  
<https://doi.org/10.1007/s11033-012-2279-3>  
PMID:[23184044](#)
81. Sujit KM, Singh V, Trivedi S, Singh K, Gupta G, Rajender S. Increased DNA methylation in the spermatogenesis-associated (SPATA) genes correlates with infertility. *Andrology.* 2020; 8:602–09.  
<https://doi.org/10.1111/andr.12742> PMID:[31838782](#)

82. Lu Q, Gore M, Zhang Q, Camenisch T, Boast S, Casagrande F, Lai C, Skinner MK, Klein R, Matsushima GK, Earp HS, Goff SP, Lemke G. Tyro-3 family receptors are essential regulators of mammalian spermatogenesis. *Nature*. 1999; 398:723–28.  
<https://doi.org/10.1038/19554> PMID:[10227296](https://pubmed.ncbi.nlm.nih.gov/10227296/)

## SUPPLEMENTARY MATERIALS

### Supplementary Tables

Please browse Full Text version to see the data of Supplementary Tables 1–5, 7–22.

**Supplementary Table 1. DEG\_Incrna.**

**Supplementary Table 2. DEG\_mrna.**

**Supplementary Table 3. chromosome\_Inc\_condition.**

**Supplementary Table 4. chromosome\_mrna\_condition.**

**Supplementary Table 5. Group\_Inc\_condition.**

**Supplementary Table 6. Group\_Inc\_condition.**

<b>Group_Inc</b>	<b>Percentage</b>
Antisense Down	0.025381
Antisense Up	0.455584
Intergenic Down	0.098562
Intergenic Up	0.133249
Intron Down	0.021574
Intron Up	0.084179
Sense Down	0.050761
Sense Up	0.130711

**Supplementary Table 7. go\_org\_con.**

**Supplementary Table 8. KEGG\_org\_con.**

**Supplementary Table 9. TOP\_correlation\_network.**

**Supplementary Table 10. SR\_correlation\_network.**

**Supplementary Table 11. ER\_correlation\_network.**

**Supplementary Table 12. OX\_correlation\_network.**

**Supplementary Table 13. OT\_correlation\_network.**

**Supplementary Table 14. cis\_trans\_top.**



**Supplementary Table 15. cis\_trans\_SR.**

**Supplementary Table 16. cis\_trans\_ER.**

**Supplementary Table 17. cis\_trans\_OX.**

**Supplementary Table 18. cis\_trans\_OT.**

**Supplementary Table 19. All Sample.count.condition.**

**Supplementary Table 20. TF factor.**

**Supplementary Table 21. Other TF.**

**Supplementary Table 22. Associated lncRNAs for crucial genes in endoplasmic reticulum stress, oxidative stress, protein transport and spermatogenesis process.**