



OPEN ACCESS

EDITED BY

Ana Fabricia Braga Magalhães,
Universidade Federal dos Vales do
Jequitinhonha e Mucuri (UFVJM), Brazil

REVIEWED BY

Ivan Cunha Bustamante-Filho,
University of Vale do Taquari, Brazil
Fabieli Loise Braga Feitosa,
Federal University of Bahia (UFBA), Brazil

*CORRESPONDENCE

Mustafa Hitit
✉ muhitit@pvamu.edu;
✉ vetdrmustafahitit@gmail.com

RECEIVED 13 November 2023

ACCEPTED 08 April 2024

PUBLISHED 10 May 2024

CITATION

Hitit M, Kaya A and Memili E (2024) Sperm
long non-coding RNAs as markers for ram
fertility. *Front. Vet. Sci.* 11:1337939.
doi: 10.3389/fvets.2024.1337939

COPYRIGHT

© 2024 Hitit, Kaya and Memili. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with these
terms.

Sperm long non-coding RNAs as markers for ram fertility

Mustafa Hitit^{1,2*}, Abdullah Kaya³ and Erdogan Memili²

¹Department of Genetics, Faculty of Veterinary Medicine, Kastamonu University, Kastamonu, Türkiye, ²College of Agriculture, Food and Natural Resources, Cooperative Agricultural Research Center, Prairie View A&M University, Prairie View, TX, United States, ³Department of Animal and Dairy Sciences, College of Agricultural and Life Sciences, University of Wisconsin–Madison, Madison, WI, United States

It is critical in sheep farming to accurately estimate ram fertility for maintaining reproductive effectiveness and for production profitability. However, there is currently a lack of reliable biomarkers to estimate semen quality and ram fertility, which is hindering advances in animal science and technology. The objective of this study was to uncover long non-coding RNAs (lncRNAs) in sperm from rams with distinct fertility phenotypes. Mature rams were allocated into two groups: high and low fertility (HF; $n = 31$; $94.5 \pm 2.8\%$, LF; $n = 25$; $83.1 \pm 5.73\%$; $P = 0.028$) according to the pregnancy rates sired by the rams (average pregnancy rate; $89.4 \pm 7.2\%$). Total RNAs were isolated from sperm of the highest- and lowest-fertility rams ($n = 4$, pregnancy rate; $99.2 \pm 1.6\%$, and $73.6 \pm 4.4\%$, respectively) followed by next-generation sequencing of the transcripts. We uncovered 11,209 lncRNAs from the sperm of rams with HF and LF. In comparison to each other, there were 93 differentially expressed (DE) lncRNAs in sperm from the two distinct fertility phenotypes. Of these, 141 mRNAs were upregulated and 134 were downregulated between HF and LF, respectively. Genes commonly enriched for 9 + 2 motile cilium and sperm flagellum were *ABHD2*, *AK1*, *CABS1*, *ROPN1*, *SEPTIN2*, *SLIRP*, and *TEKT3*. Moreover, *CABS1*, *CCDC39*, *CFAP97D1*, *ROPN1*, *SLIRP*, *TEKT3*, and *TTC12* were commonly enriched in flagellated sperm motility and sperm motility. Differentially expressed mRNAs were enriched in the top 16 KEGG pathways. Targets of the differentially expressed lncRNAs elucidate functions in *cis* and *trans* manner using the genetic context of the lncRNA locus, and lncRNA sequences revealed 471 mRNAs targets of 10 lncRNAs. This study illustrates the existence of potential lncRNA biomarkers that can be implemented in analyzing the quality of ram sperm and determining the sperm fertility and is used in breeding soundness exams for precision livestock farming to ensure food security on a global scale.

KEYWORDS

ram, sperm, fertility, long non-coding RNAs, gene expression

1 Introduction

Male fertility is crucial for animal reproduction as it encompasses the successful fertilization of an egg with a viable spermatozoon, subsequently leading to embryonic and fetal development, ensuring the species' continued existence. Fertility is an economically important trait, and ram flock represents more than 50% of the genetics of every sheep farmer's flock (1). Thus, proper management of rams for maximum performance and longevity is vital to the success and reproductive efficiency of sheep farming (2). Accordingly, advancing ram fertility in the livestock system is imperative to provide animal-based food demands to feed the ever-increasing human population on a global scale.

In addition to genetic evaluation and testing, rams have been selected according to breeding soundness exam (BSE), which requires a series of examinations, including physical exam, scrotal measurement, sperm morphology, and motility. Despite massive attempts to evaluate ram fertility using the BSE, the predictability of ram fertility still awaits improvement. Non-compensable factors, such as DNA damage, RNA molecules (3), and protein markers (4, 5) in sperm fertility, are attributed to minute sperm abnormalities that cannot be determined using conventional procedures (6). Advanced omics approaches have paved the way for underlying molecular mechanisms related to spermatogenesis, fertilization, and embryogenesis (7, 8). Such methods may be used more widely in the future for producing farm animals in combination with the evaluation of semen parameters (9, 10), while also estimating accurate sperm fertility markers in farm animals (11, 12).

Spermatozoa can transmit not just the paternal DNA but also certain RNA molecules and transcription factors, presumably inherited into the oocytes during fertilization (13). Researchers have recently used cutting-edge approaches for discovering small noncoding RNAs with various nucleotide lengths and biogenesis processes, such as PIWI-interacting RNAs (piRNA), microRNAs (miRNA), tRNAs, and long noncoding RNA (lncRNA)-derived short RNAs (14–17). Sperm bearing RNA molecules are implicated in spermatogenesis and embryo development at transcriptional and posttranscriptional levels (18–20). As such, they play roles in regulating spermatogenesis and fertilization by transferring small noncoding RNA (snRNA) and lncRNA into the oocyte (3, 19, 21). In addition, paternally derived noncoding RNAs are key regulators of preimplantation embryos (22) since some of them are involved in the control of gene expression in zygotic and early embryonic development (23). Accordingly, sperm noncoding RNAs can enhance the transmission of epigenetic information to the offspring (13) because environmental and metabolic-induced modifications of sperm cells may influence the epigenetic modulation of embryo development by changing the gene expression through noncoding RNAs (24–27).

lncRNAs comprise nucleotides located in the cytoplasm and nucleus, which are transcribed by RNA polymerase II and are longer than 200 nucleotides and lack protein-coding capacity. lncRNAs, with their higher and more stable structures, modulate gene expression (*cis* and *trans* manner) at several levels, including epigenetic, transcriptional, posttranscriptional, and posttranslational, through their interactions with mRNA, proteins, and other snRNAs. They have significant functions in biological processes such as modifying the chromatin structure, activating gene expression, inhibiting gene expression, and translating mRNA molecules. lncRNAs have crucial roles in regulating the many biological processes that are highly expressed in bovine and mouse testis and mature sperm (28, 29). There is a linkage between noncoding RNA markers and male fertility, as the transcripts can be involved in the prediction of fertility (30, 31). The lncRNAs differentially expressed in distinct motility phenotypes of bovines and humans (32) imply that sperm lncRNA may possess functional roles in fertility (33). Using RNA sequencing technologies, sperm bearing RNA was found to be conserved among many species, such as stallions, goats, and boar (34–36). The unique male sperm ncRNAs with consistent fertility phenotypes can be valuable

as potent fertility biomarkers. Accordingly, this study aimed to uncover long noncoding RNA profiles from ram sperm possessing distinct fertility phenotypes.

2 Material and methods

2.1 Ram fertility assessment and experimental design

The Republic of Turkey's Ministry of Agriculture and Forestry's Institute of Bahri-Dagdaş International Agricultural Research supplied information on the fertility phenotypes of adult rams. Pregnancy rates from natural mating were used to determine the fertility of mature rams ($n = 66$) at least 3–4 years old during the breeding seasons of 2017–2018–2019. The ewes' estrus was detected using teaser rams, who were not permitted to mate, by covering the prepuce area. The teaser rams were presented to the ewes early in the morning for about 30 min. The estrus was considered to be ewes seeking, standing for teasing, and allowing mount attempts by teaser rams. A handler selected estrous ewes and brought them into an enclosure along with a randomly selected single ram for natural mating. Throughout the breeding season, estrus detection was maintained, and ewes were accepted to mate with the chosen ram at random. Ewes were regarded as pregnant if they did not return to estrus within 35 days of mating. In addition, the number of pregnant and non-pregnant ewes for each ram was confirmed by matching the mating and lambing dates according to the duration of pregnancies. The rams' fertility scores were rated based on their conception rates. The rams were split into two groups based on their fertility levels: the high-fertility (HF) group ($n = 31$; $94.5 \pm 2.8\%$) and the low-fertility (LF) group ($n = 25$; $83.1 \pm 5.73\%$). The average pregnancy rate was $89.4 \pm 7.2\%$ ($n = 66$), and the rams were divided into these groups based on their fertility levels. We conducted an *a priori* power analysis using G*Power3 (V3.1.9.4) to test the differences between two independent group means using a two-tailed test, an effect size ($d = 2.52$), and an alpha level of .05. The result showed that a total sample of eight animals with two equal-sized groups of $n = 4$ was required to achieve a power of .80 for RNA profiling. However, we excluded the rams that did not have 50 mating and were not used, so we profiled four animals for each group out of a total of 56. During the breeding season, each ram served at least 50 ewes in both groups.

2.2 Semen collection

The Bahri-Dagdaş Research Center Ethical Committee, Turkey (Number: 22.12.2016/58), approved the animal procedures. We trained the rams to obtain sperm using an artificial vagina (AV) that enabled them to mount on teaser ewes during estrus. Rams were permitted to ejaculate into the AV upon mounting. The first three collections were discarded prior to the collection of research samples, which was followed by semen collection and processing for use in research. Fertility rates that were 1 standard deviation above or below the mean were termed outliers. Four rams with the highest fertility (pregnancy rate; $\% 99.2 \pm 1.6$) and four rams with

the lowest fertility (pregnancy rate; % 73.6 ± 4.4) were selected for lncRNA sequencing with high confidence. Subsequently, we collected about 2×10^9 /ml spermatozoa per ejaculate, and then the aliquots from each sample were adjusted to a final concentration of 10^7 /ml in straws and frozen at -80°C until lncRNA analysis.

2.3 RNA isolation

Prior to RNA isolation, we purified sperm by filtering the semen samples with a 500-mesh sieve to eliminate cell debris. Then, samples were treated with a somatic cell lysis solution (0.3% Triton X-100 and 0.1% SDS in DEPC-treated H_2O) for 30 min on ice to eradicate somatic cells, followed by microscopic analysis of non-sperm cell contamination. We isolated total RNA from the purified ram sperm ($n = 4$, for each group) samples using the SanPrep column microRNA miniprep kit (Bio Basic Inc, Canada) with slight modification using the manufacturer's protocols. We added 800 μl of a guanidine–thiocyanate lysis buffer enriched in 20 mM DL-dithiothreitol onto the pellet, and then sperm cells were homogenized by passing the samples through a 26-G needle syringe 20–25 times. After other contaminants were thoroughly removed and total RNA was attached to the membrane, an on-column DNase digestion was carried out to remove any traces of DNA contamination. We evaluated the concentration and integrity of the total RNA samples using a NanoDrop (Colibri Microvolume Spectrometer, Titertek-Berthold, Germany) and a 2100-Bioanalyzer with the RNA 2100 Nano Chip (Applied Biosystems, Carlsbad, CA, USA), respectively.

2.4 Library preparation for lncRNA sequencing

Each RNA sample was utilized to prepare 2 g of RNA for the RNA library, and ribosomal RNA was first eliminated using the Ribo-Zero™ rRNA Removal Kit from Epicenter. Then, using the NEBNext® Ultra™-Directional RNA Library Prep Kit from Illumina® (NEB, USA) in compliance with the commercial kit protocol, sequencing libraries were developed utilizing the rRNA-depleted RNA. Reverse transcription was used to create the first strand of the cDNA following fragmentation with an average length of 200 bp. The Agilent Bioanalyzer 2100 equipment was used to evaluate library quality following product purification using the AMPure XP system. As a result, libraries underwent sequencing using the Illumina NovaSeq 6000 (Illumina Inc., San Diego, CA, United States), and Novogene Corporation (Beijing, China) generated 150-bp paired-end reads.

2.5 RNA-Seq read alignment and transcript assembly

Initially, rRNA, adapter sequences, empty reads, and low-quality reads were eliminated from the raw data. All trimmed reads were confirmed to satisfy the quality threshold (Q -score; Q20 and Q30) to ensure that there was no bias in the evaluation

step toward approaches that favor maximum read. The Phred scale indicating the reliability of base-calling, with Q20 representing a base call accuracy of 99% (or a 1% chance of error) and Q30 representing a base call accuracy of 99.9% (or a 0.1% chance of error) was used as the quality score. The Ovis aries (v4.0) reference genome was indexed with Bowtie v2.0.6, and the processed paired-end reads were mapped to that genome with HISAT2 2.1.0 (37). Each sample's mapped reads were constructed using StringTie (v1.3.1) (38). Finally, Cuffcompare, a program included in Cufflinks (v2.1.1), was used to annotate the assembled transcripts.

2.6 Putative lncRNA identification

To classify newly screened lncRNAs with respect to their positional relationship with known mRNAs, putative lncRNAs were identified. To minimize the false-positive rate (FDR), assembled transcripts were classified to retrieve putative lncRNAs, such as lincRNA, antisense lncRNA, intronic lncRNA, and sense-overlapping. (A) Transcripts with a single exon were eliminated. (B) Transcripts with fewer than 200 nucleotides were eliminated. (C) Using Cufflinks v2.1.1, the annotated lncRNAs in the database were used to exclude the transcripts that overlapped with the exon region of the database annotation. (D) All transcripts with modest levels of expression [FPKM thresholds were set for the categorization of transcript expression levels; genes with very low or no expression (FPKM < 0.5), and FPKM = 0.5 was chosen as the cutoff to filter out the average read coverage per transcript which was much higher than the other transcript] were omitted (FPKM of a single exon transcript). Using three methods, namely, Coding-Non-Coding-Index (CNCI) (39), Pfam-scan (40), and coded potential calculator (CPC) (41), all estimated transcripts with coding potential were filtered out, and a set of putative lncRNAs was compiled from those with noncoding potential. Using Cuffcompare, the distinct categories of lncRNAs were obtained.

2.7 Analysis of mRNA and lncRNA expression levels

The FPKM value was used to evaluate levels of mRNA and lncRNA expression. Cuffdiff (v2.1.1) was used to determine lncRNA FPKM values. Later, the statistically significant DE genes were quantified by a \log_2 -fold change higher or equal to 2 (P -value < 0.05) or P -adjust < 0.05 (applied correction for multiple testing to the P -values to FDR), using Ballgown (42).

2.8 Target gene prediction

DE lncRNAs were selected to determine target genes. Pearson's correlation was used to assess potential coexpression between lncRNAs and mRNAs. A Pearson correlation >0.7 and a P -value of 0.05 were used to determine a positive association between an lncRNA and an mRNA. lncRNAs can act as *cis* regulators by remodeling factors onto local chromatin. We described *cis*-modulated genes as protein-coding genes that were coexpressed

TABLE 1 Summary of data production.

Sample	Raw_reads	Clean_reads	Raw_data (G)	Clean_data (G)	Error_rate (%)	Q20 (%)	Q30 (%)	GC_content (%)
LF	41,020,716	40,183,273	12.3	12.1	0.03	96.06	90.96	60.89
HF	50,087,458	49,069,550	15	14.7	0.03	95.39	89.5	62.37

Raw read statistics: the total number of reads for each file is determined, and each set of four consecutive lines represents the information for one read. Clean reads are the same as raw reads, but only the filtered reads are used to calculate the results of any further study; Clean reads (G) the sum of a sequence's length and number, expressed in giga bases; Error rate: the sequencing error rate; Q20, Q30: the proportion of all bases for which the Phred score is 20 or 30 or above, indicating base call accuracy; the proportion of guanine (G) and cytosine (C) in all bases is known as GC content.

with one dysregulated lncRNA and were within 30 kb upstream or downstream in genomic distance in the same allele. To participate in certain biological processes, key transcription factors (TFs) are regulated in a *trans* manner by unique lncRNAs. As a result, we matched these lncRNAs' coexpressed mRNAs to mRNAs that were regulatory targets of specific TFs to anticipate that these lncRNAs might be involved in pathways controlled by these TFs.

2.9 Gene ontology terms and KEGG pathway enrichment

We assessed DE mRNAs for gene ontology (GO) enrichment analysis using g:Profiler (43). GO terms possessing a corrected P -value < 0.05 were accepted as significantly enriched by DE genes. The statistical enrichment of lncRNA target genes was examined for KEGG pathway functional analysis based on the reactome pathway database by WebGestalt (WEB-based Gene Set Analysis Toolkit) with $P < 0.05$ and FDR $< 5.0\%$ (44). The following parameters were specified for the enrichment analysis. A particular organism was labeled as *Ovis aries* (sheep). Sequential GO analyses (biological process; BP, cellular component; CC, and molecular function; MF) were conducted. The g:SCS method is used to compute multiple testing corrections for P -values based on GO and pathway enrichment analysis, with the user threshold set at 0.05.

2.10 Statistical analysis

Data were analyzed through SPSS software (version 22.0). Statistical plots were generated using GraphPad Prism 9 (GraphPad Software, USA). During the experimental process, four biological replicates were included with the measurement repeated twice. LF vs. HF groups were analyzed using an independent t -test. Significance was accepted at P -value = 0.05.

3 Results

3.1 Overview of sequencing data in ram sperm between LF and HF

In the current study, we pooled two cDNA libraries out of the eight total RNAs (LF and HF, each = 4) obtained from low- and high-fertility ram sperm. The total number of raw readings obtained from all cDNA libraries was 91,108,174. Upon filtering the reads, a total of 89,252,823 clean reads were obtained. The Q20 (%) percentages were 96.06 and 95.39 for LF and HF, respectively. The

TABLE 2 A list of reads that were mapped to the reference genome.

Sample name	HF	LF
Total reads	98,139,100	80,366,546
Total mapped	82,508,774 (84.07%)	69,017,041 (85.88%)
Multiple mapped	3,526,455 (3.59%)	2,394,038 (2.98%)
Uniquely mapped	78,982,319 (80.48%)	66,623,003 (82.90%)
Read-1	40,351,963 (41.12%)	33,802,488 (42.06%)
Read-2	38,630,356 (39.36%)	32,820,515 (40.84%)
Reads map to "+"	39,407,126 (40.15%)	33,255,875 (41.38%)
Reads map to "-"	39,575,193 (40.33%)	33,367,128 (41.52%)
Non-splice reads	77,745,931 (79.22%)	65,903,933 (82.00%)
Splice reads	1,236,388 (1.26%)	719,070 (0.89%)

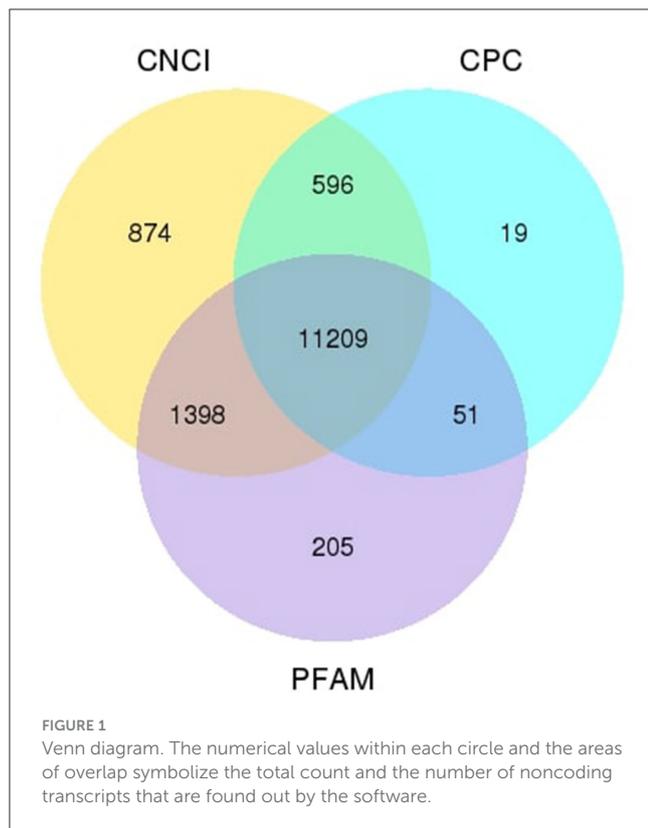
Total reads: the number of reads after data filtration (clean data); Total mapped represents the total quantity of mappable reads. If a suitable reference genome is available and no contamination occurs during the experimental procedure, the percentage will typically exceed 70%. Numerous mappings: the number of sequences mapped to multiple reference sequence positions; specifically mapped: the number of reads that map to a particular position in the reference sequences. Map to "+"; Reads that map to "-": the number of reads that are mapped to the minus strand. Quantity of reads mapped to two exons; also known as junction reads. Similarly, non-splice reads are those that are completely mapped to a single exon. The ratio of splice readings to total read length is proportional.

TABLE 3 A list of reads that were mapped to the reference genome.

Classification of mapped reads (LF)	Classification of mapped reads (HF)
Others (49,390,745 [78.4%])	Others (57,960,150 [77.95%])
Protein_coding (12,785,367 [20.3%])	Protein_coding (15,352,804 [20.65%])
lncRNA (697,609 [1.1%])	lncRNA (855,897 [1.15%])
pseudogene (125,827 [0.2%])	pseudogene (137,580 [0.2%])
rRNA (10,395 [0.02%])	rRNA (11,500 [0.02%])
miRNA (5,685 [0.01%])	miRNA (7,407 [0.01%])
misc_RNA (3,588 [0.0%])	misc_RNA (13,778 [0.0%])
ribozyme (2,495 [0.0%])	ribozyme (7,883 [0.0%])

Distribution of mapped reads in known types of RNAs for low fertility (LF) and high fertility (HF).

Q30 (%) percentages were shown to be 90.96 and 89.50 for LF and HF, respectively. The percentages of GC content (%) were shown to be LF, 60.89, and HF, 62.37 (Table 1). Furthermore, after aligning the clean reads with the ovine reference genome through the TopHat2 algorithm, we discovered that the total mapped reads or fragments referring to all samples exceeded 75% and were mapped in the reference genome (Table 2).



3.2 Chromosome read distribution and known RNA types

The distribution intensity of the total mapped reads was split up and computed for both the plus and minus strands within each chromosome (Supplementary Table 1). The distributions of reads in the known RNA types for LF and HF are depicted in Table 3. The corresponding data related to these distributions can be found in Supplementary Table 2.

3.3 Novel lncRNA identification in ram sperm between LF and HF

We applied the lncRNA filtering method to determine the novel lncRNA candidates in ram sperm. The specific filtering technique is outlined in Supplementary Table 3. Using this lncRNA filtering process, we identified a total of 14,352 lncRNAs, of which 11,209 were novel candidates, the data for which are supplied in Supplementary Table 4. We utilized three distinct software packages (CNCI, CPC, and Pfam) to estimate the potential protein-coding ability of the transcripts (Figure 1).

3.4 Identification and characteristics of lncRNAs between LF and HF in ram sperm

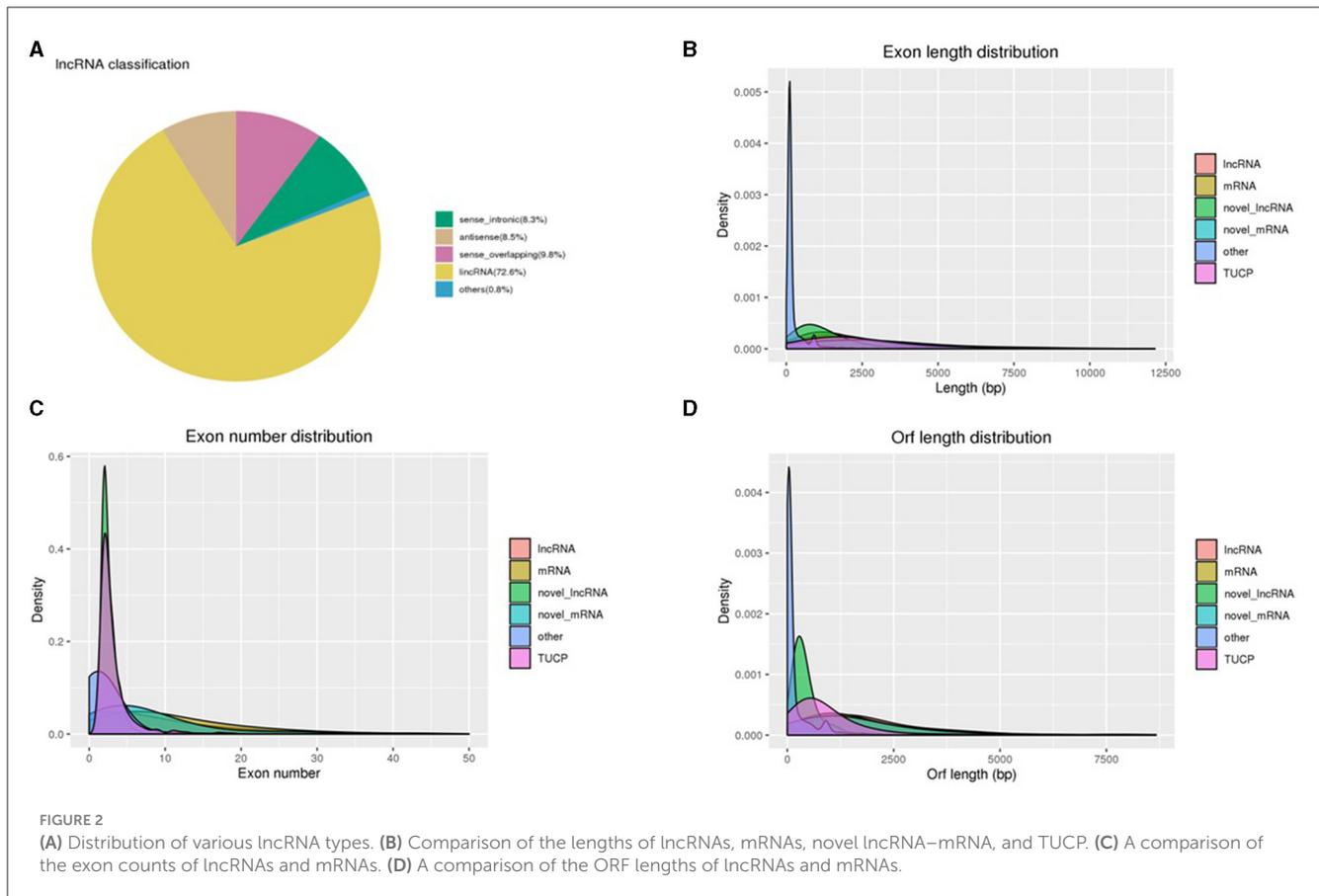
The primary characteristics of the lncRNAs' biotype distribution and length, exon intensity, and ORF length were

examined as a consequence of the sequence analysis. Using Cufflinks, we divided lncRNAs into three groups: long intervening noncoding RNA (lincRNA), antisense lncRNA, and intronic lncRNA. Of these, 8.3%, 8.5%, 9.8%, 72.6%, and 0.8% of the attained lncRNAs were sense intronic, antisense, sense overlapping, lincRNA, and others, respectively (Figure 2A). Using this classification scheme, we revealed that the majority of the sperm-specific lncRNAs (72.6%) were lincRNAs. The exon lengths of the obtained lncRNAs were between 126 and 20,040 bp, and approximately 37% of the lncRNAs were demonstrated to be intense in 946–1,356 bp length (Figure 2B), with 1,542 bp being the median value. The exon length of the obtained mRNAs ranged between 75 and 26,726 bp, with around 10% of the mRNAs being intense at 835–1,025 bp (Figure 2B). Furthermore, lncRNAs have been shown to be more abundant in the first and fourth exons, whereas mRNAs were found to be more abundant in the first and seventh exons (Figure 2C). The length of the ORF of the lncRNAs was between 54 and 20,037 bp, with a median value of 1,461 bp (Figure 2D). For mRNAs, the ORF length was reported as 3–26,460, and the median value was 1,332.

The expressions associated with transcripts of lncRNA, mRNA, novel lncRNAs, and TUCP were evaluated using Cuffdiff. The most common approach for predicting gene expression levels, FPKM, is based on the effects of gene length on sequence depth and fragments described in the RNA sequence. The FPKM values associated with samples were calculated (Supplementary Table 5). Upon comparing FPKM values between LF and HF, no significant change was observed (Figures 3A, B). Moreover, the expression levels related to novel lncRNAs were the highest among the transcripts. TUCP and novel mRNA transcripts were greater than those of lncRNA and mRNAs, while lncRNA transcripts were similar to those of mRNA (Figures 3C, D).

3.5 lncRNAs and mRNAs that are differentially expressed in LF and HF

Cuffdiff was employed to identify differentially expressed lncRNAs, mRNAs, and TUCPs. As a result, in HF and LF ram sperm, 93 lncRNAs and 275 mRNAs were reported to be differentially expressed (DE). We discovered 49 lncRNAs that are significantly upregulated and 44 that are downregulated between LF and HF groups. In addition, we discovered that 141 mRNAs were upregulated, whereas 134 were downregulated. Volcano plots show the up and downregulation highlights (Figures 4A, B; Supplementary Table 6). We also examined the expression patterns of DE lncRNAs and mRNAs using hierarchical clustering analysis, which serves as an additional approach for shedding light on differentially expressed genes by grouping genes with comparable expression patterns. The transcript FPKMs were utilized for hierarchical clustering, with discrete colors indicating the direction of the expression level. The clustering of genes on the left was caused by similar expressions (fold change >2, $P < 0.05$) and LF and HF ram sperm, whereas the expression change from blue to red indicated them as increasingly increased (Figures 4C, D).



3.6 DE mRNA functional annotation and KEGG pathways analysis

Following a study at the transcriptome level, 275 mRNAs were selected for enrichment analysis due to their differential expression levels (141 upregulated and 134 downregulated). The top eight GO annotations terms in biological process (BP) were significantly enriched in the DE mRNAs, namely, the biological process (GO:0008150), cellular process (GO:0009987), regulation of cellular process (GO:0050794), organelle organization (GO:0006996), flagellated sperm motility (GO:0030317), sperm motility (GO:0097722), and cilium movement involved in cell motility (GO:0060294). In addition, in cellular component, intracellular anatomical structure (GO:0005622), sperm flagellum (GO:0036126), and motile cilium (GO:0031514) were among the top GO terms and binding (GO:0005488), protein binding (GO:0005515), enzyme binding (GO:0019899), and protein C-terminus binding (GO:0008022) in molecular function (MF; Figure 5A). The detailed definition of the GO terms is presented in Supplementary Figure 1. According to the statistics of the pathway enrichment, the top 16 KEGG pathways, including complex I biogenesis, the citric acid (TCA) cycle, respiratory electron transport, respiratory electron transport, mitochondrial translation, and gene silencing by RNA, had higher concentrations of DE mRNAs. *ABHD2*, *AK1*, *CABS1*, *ROPN1*, *SEPTIN2*, *SLIRP*, and *TEKT3* genes were commonly enriched for 9 + 2 motile cilium, sperm flagellum, and motile cilium. Moreover, *CABS1*, *CCDC39*,

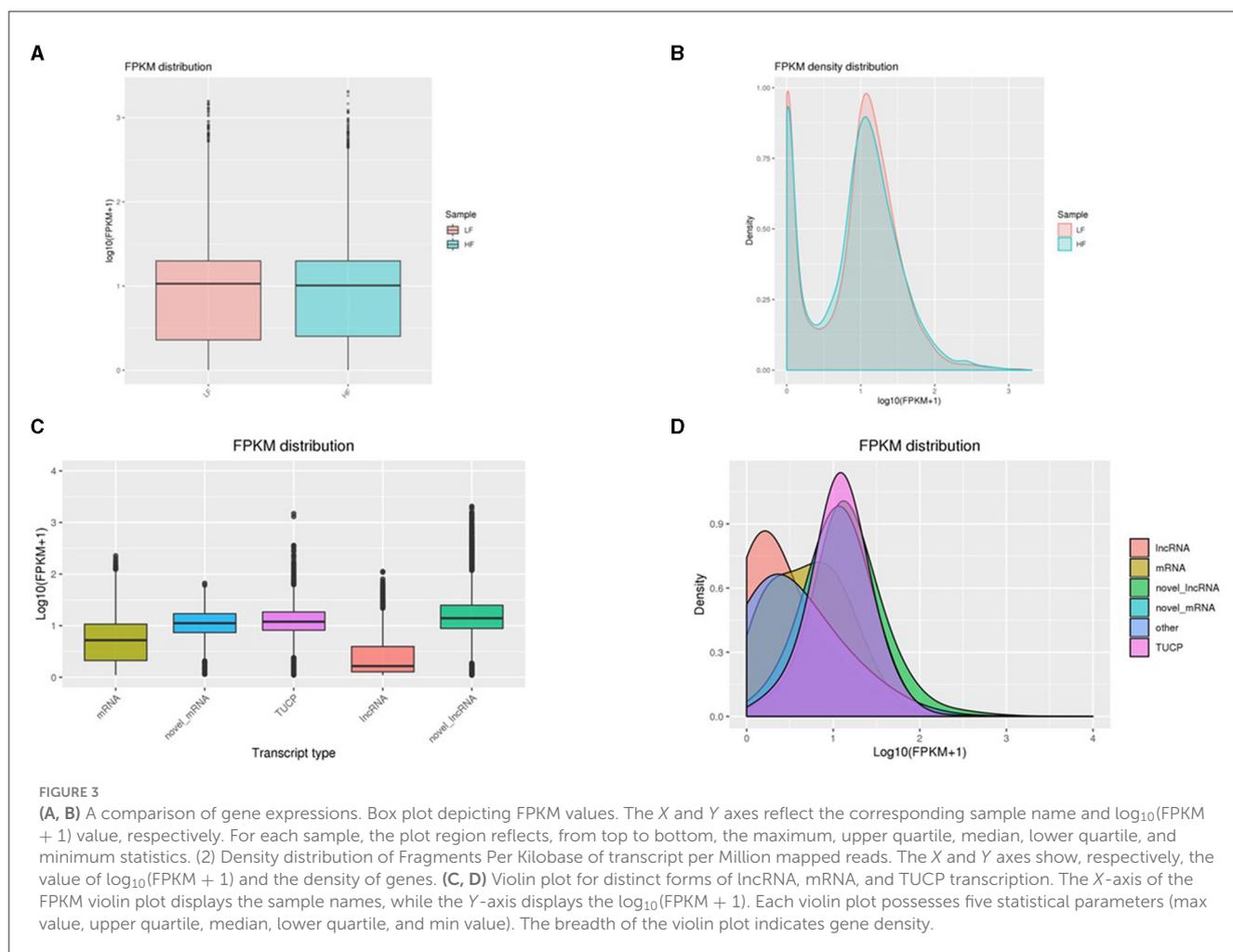
CFAP97D1, *ROPN1*, *SLIRP*, *TEKT3*, and *TTC12* were commonly enriched in flagellated sperm motility, sperm motility, and cilium movement involved in cell motility (Supplementary Figure 1).

3.7 Target genes of cis- or trans-regulated by lncRNAs

To better understand how lncRNAs act in both *cis* and *trans* manner, we predicted their targets. There were 471 mRNAs that were discovered as targets of 10 lncRNAs (five upregulated and five downregulated), using 30 kb as the cutoff (Supplementary Figure 2). According to the GO enrichment study results, 55 significant GO terms were found (corrected *P*-value 0.05). In MF, the top five GO keywords were nucleoside phosphate binding, protein binding, small-molecule binding, binding, and nucleotide binding (Supplementary Figure 3). In BP, negative regulation of the cellular metabolic process, system development, response to chemical stimuli, organonitrogen compound metabolic process, and multicellular organism development were among the top five GO terms and cytoplasm, cell junction, cytosol, nucleoplasm, and synapse in CC (Figure 5B).

4 Discussion

The quality of sperm on a cellular level alone is no longer considered to be a reliable predictor of male fertility in selecting

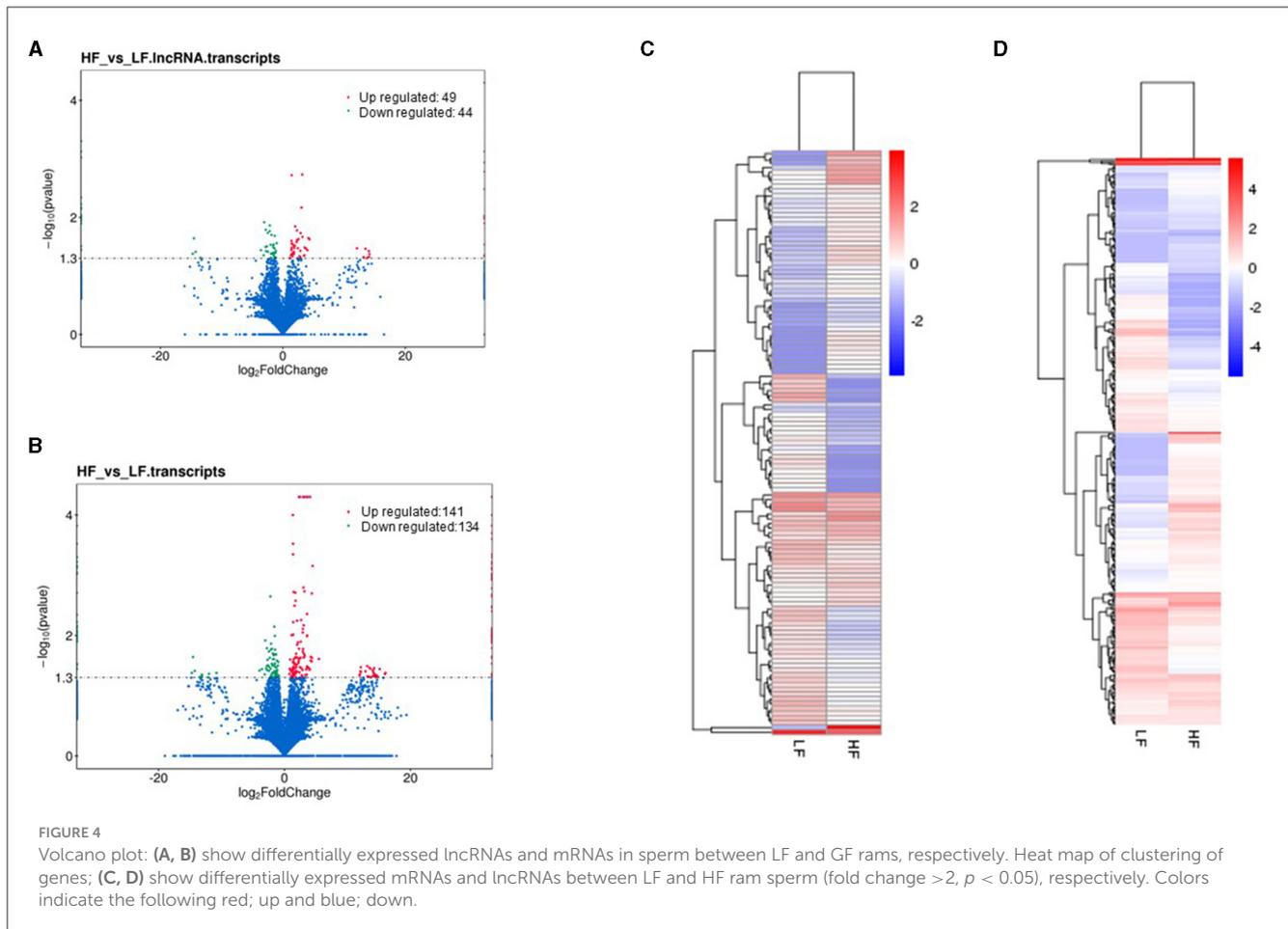


superior male prospects in livestock as a source of frozen semen (12, 31). Methods for analyzing the sperm transcriptome, such as measuring messenger RNA levels, have been linked to increased fertility (45). Analysis of the sperm transcriptome has become the primary tool for predicting male fertility potential in the livestock business, due in part to the widespread application of this technology in livestock growth (45, 46). Accordingly, this study demonstrated long noncoding RNA profiles in sperm from rams with HF and LF phenotypes to uncover potential lncRNAs associated with fertility.

The comparative analysis across multiple species indicates a considerable gap in our understanding of the functional roles of lncRNAs within sperm cells despite their prevalence in male germ cell development. Previous studies on cattle, boar, sheep, mice, and humans have collectively identified substantial numbers of potential lncRNAs in sperm cells (46–48). However, the functional annotation and investigation of these lncRNAs have remained limited. We obtained a total of 91,108,174 cDNA libraries, of which 89,252,823 were clean reads from ram spermatozoa. This provides a robust dataset for comparative analysis, especially when considering the consistency with similar studies conducted, such as on human, turkey, and boar (79.8, 84, and about 65.5 million, respectively) spermatozoa (49, 50). Furthermore, the range of

unique mapped reads, ranging from 66,623,003 to 78,982,319, indicated a strong alignment of the sequenced data to the reference genome of sheep. The high percentage of over 80% uniquely mapped reads demonstrated a significant mapping consistency comparable to that observed in bull and stallion spermatozoa (32, 36). This consistency across species further confirms the reliability and quality of the obtained sequence data for ram spermatozoa, supporting the validity of results.

In this study on ram fertility, we identified 11,209 sperm lncRNAs, a subset comprising 93 differentially expressed lncRNAs associated with LF and HF phenotypes in ram sperm. Comparatively, in mouse mature spermatozoa, 4,088 novel lncRNA transcripts were identified out of 20,907 known lncRNA transcripts (29), demonstrating the complexity and diversity of these transcripts in different species' sperm cells. Similarly, in human sperm, 27,472 novel lncRNAs were discovered (29), with 19 differential expressions of lncRNA out of 11,561 lncRNA transcripts in mature bull spermatozoa (32). Analysis of goat spermatozoa revealed 655 lncRNA transcripts relevant to spermatogenesis from sequencing of the cDNA library, of which 1% annotated to lncRNAs was similar to ram spermatozoa with a classification of 1.1% (51). We also showed that the type of lncRNA seemed to be closer between ram and goat spermatozoa



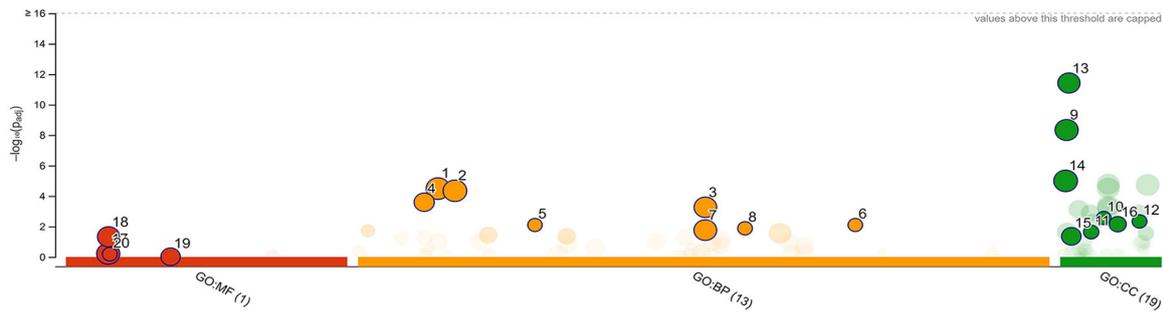
(34), both revealing a 72.6 and 76.64% annotation to lncRNAs, respectively, which highlights the relevance of these findings across species and their potential implications for understanding male fertility mechanisms.

The GO and KEGG investigations were carried out for associated genes of DE mRNAs and lncRNAs to completely examine the functional roles of mRNAs and lncRNAs in ram sperm fertility. Candidate genes have been identified using this bioinformatics approach, and they were related to male reproductive biology. Our findings demonstrated that 275 mRNA transcripts were enriched for the biological process, cellular component, and molecular function GO characterizations. It has become apparent that several metabolic pathways and regulatory mechanisms have crucial roles in fertility, as related to sperm motility. We demonstrated that *ABHD2*, *AK1*, *CABS1*, *ROPNI*, *SEPTIN2*, *SLIRP*, and *TEKT3* genes were commonly enriched for the 9 + 2 motile cilium, sperm flagellum, and motile cilium. Of these, *ABHD2*, an isolated molecule from sperm tails, is needed to activate sperm (52). *AK1* is an enzyme that is often responsible for cellular energy balance. It is detected in the flagella of murine and bovine sperm, which suggests that it is involved in sperm motility and is also demonstrated to be directly associated with bull fertility (53–55). Moreover, *CABS1*, *CCDC39*, *CFAP97D1*, *ROPNI*, *SLIRP*, *TEKT3*, and *TTC12* were commonly enriched in flagellated sperm motility, sperm motility, and cilium

movement involved in cell motility. As a Ca²⁺ storage protein in mature sperm, *CABS1* is a crucial factor in the regulation of calcium signaling and has been shown to preserve sperm flagella structure (56, 57). *ROPNI*, implicated in fibrous sheath integrity and sperm motility, is engaged in PKA-dependent signaling for spermatozoa capacitation; therefore, mutation and lack of its expression in murine sperm cells cause impaired fertility (58, 59). We showed that differentially abundant protein profiles of sperm from rams with contrasting fertility phenotypes were associated with metabolic energy generation by sperm cells along with the motility signaling pathway (4). Our results are consistent with results from the previous study that mRNA–lncRNA interaction seemed to regulate signaling pathways for functional motility.

The primary potential roles of lncRNAs are to control the expression of neighboring protein-coding genes through both *cis* and *trans* manners, integrating transcriptional coactivation or repression. Consequently, we conducted an in-depth analysis and determined the mRNAs situated within the 30 kilobase (kb) threshold upstream and downstream regions of the differentially expressed lncRNAs. We then employed GO and KEGG analyses on the target genes to determine the undertaken lncRNAs. We found that lncRNAs with different levels of expression are involved in a number of important biological processes. We demonstrated that the identified lncRNA TCONS_00136350 may regulate the differentially expressed coding gene *ADAM* metalloproteinase

A



ID	Source	Term ID	Term Name	Padj (query...)
1	GO:BP	GO:0008150	biological_process	3.254 × 10 ⁻⁵
2	GO:BP	GO:0009987	cellular process	4.535 × 10 ⁻⁵
3	GO:BP	GO:0050794	regulation of cellular process	5.474 × 10 ⁻⁴
4	GO:BP	GO:0006996	organelle organization	2.551 × 10 ⁻⁴
5	GO:BP	GO:0030317	flagellated sperm motility	7.971 × 10 ⁻³
6	GO:BP	GO:0097722	sperm motility	7.971 × 10 ⁻³
7	GO:BP	GO:0050789	regulation of biological process	1.730 × 10 ⁻²
8	GO:BP	GO:0060294	cilium movement involved in cell motility	1.311 × 10 ⁻²
9	GO:CC	GO:0005622	intracellular anatomical structure	4.664 × 10 ⁻⁹
10	GO:CC	GO:0036126	sperm flagellum	2.743 × 10 ⁻³
11	GO:CC	GO:0031514	motile cilium	2.320 × 10 ⁻²
12	GO:CC	GO:0097729	9+2 motile cilium	4.618 × 10 ⁻³
13	GO:CC	GO:0005737	cytoplasm	3.801 × 10 ⁻¹²
14	GO:CC	GO:0005575	cellular_component	1.016 × 10 ⁻⁵
15	GO:CC	GO:0005856	cytoskeleton	4.426 × 10 ⁻²
16	GO:CC	GO:0048471	perinuclear region of cytoplasm	7.027 × 10 ⁻³
17	GO:MF	GO:0005488	binding	6.164 × 10 ⁻¹
18	GO:MF	GO:0005515	protein binding	4.730 × 10 ⁻²
19	GO:MF	GO:0019899	enzyme binding	9.666 × 10 ⁻¹
20	GO:MF	GO:0008022	protein C-terminus binding	6.585 × 10 ⁻¹

B

■ FDR ≤ 0.05 ■ FDR > 0.05

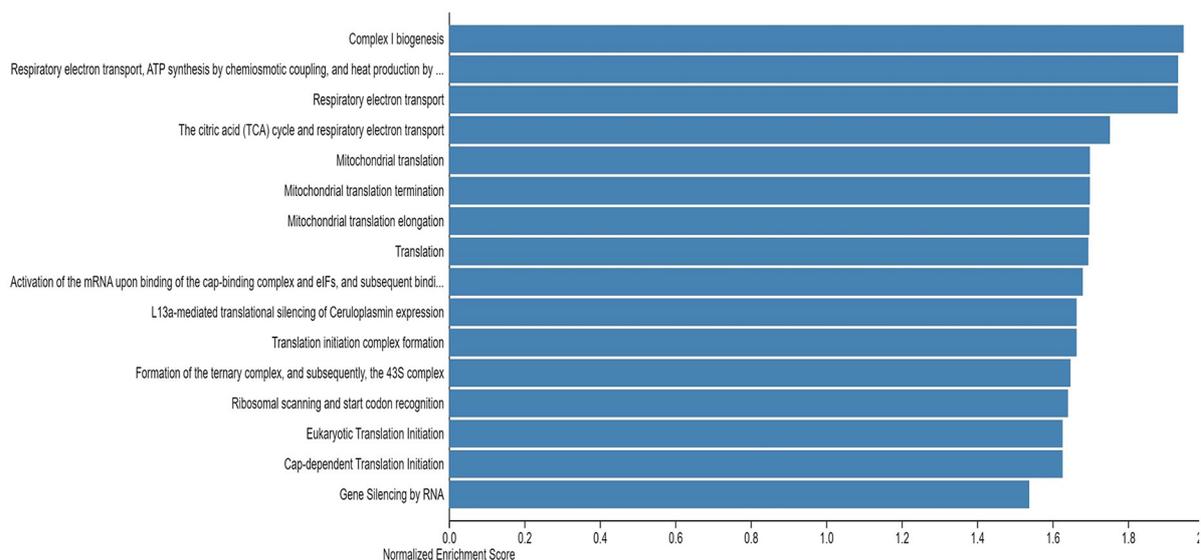


FIGURE 5

(A) Functional enrichment analysis of DE mRNAs in LF and HF sperm of ram. Significantly enriched target gene terms are shown. GO keywords are represented by red color codes, molecular functions by orange ones, biological processes by green ones, and cellular components by green ones (CC). **(B)** Functional enrichment analysis of target genes of *cis-* or *trans-*regulated by lncRNAs. Illustrated are normalized enrichment scores for specific sets of target genes.

domain 32 (*ADAM32*). *ADAM32* is a member of the disintegrin family of membrane-anchored proteins and is detected on the surface of mature sperm (60). *ADAM32* has been demonstrated to be implicated in sperm–egg plasma membrane interaction, and (61) its expression levels are correlated with the quality of human sperm (62). Although *ADAM32* is not required for normal male fertility (63), we showed that it was identified in the top 10 differentially abundant proteins in ram sperm (4). We also discovered that differentially expressed lncRNA TCONS_00035618 modulated enolase 1 (*ENO1*), the protein expression of which was higher in increased fertility of ram sperm (4). *ENO1* plays an important role in the process of metabolism of monocarboxylic acids, and it is a component in the pathway that leads to glycolysis and gluconeogenesis (KEGG:00010) associated with bull fertility (64), and its lower expression levels lead to low motility (65). In line with our study and of others, lncRNAs are potentially important for the coactivation/repression of target genes that regulate sperm fertility and motility.

5 Conclusions

Even though several potential lncRNAs have already been identified through cutting-edge technology and accessible lncRNA annotation tools, functional annotation of lncRNAs in sperm biology is still in its infancy and holds great promise. In this study, differentially expressed lncRNAs in ram sperm were ascertained along with their associations with low vs. high fertility phenotypes. These findings are important because they help advance both the fundamental science of mammalian male gamete biology and applied science that may provide practical value for potential male fertility markers.

Data availability statement

The datasets generated and/or analyzed during the current study are available in the EMBL-EBI HYPERLINK <http://www.ebi.ac.uk/biostudies/studies/S-BSST1231> accessible at this accession number “S-BSST1231”. The data encompass all raw relevant data.

Ethics statement

The animal study was approved by Bahri-Dagdas, Research Center Ethical Committee, Turkey (Number: 22.12.2016/58). The study was conducted in accordance with the local legislation and institutional requirements.

Author contributions

MH: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization,

Writing – original draft, Writing – review & editing. AK: Conceptualization, Formal analysis, Investigation, Methodology, Resources, Supervision, Validation, Writing – original draft, Writing – review & editing. EM: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This study was funded by the Kastamonu University, Scientific Research Coordination Unit with a project number of KÜ-BAP01/2020-63. Animal materials provided were by the Scientific and Technological Research Council of Turkey (TUBITAK) with project no: (117O992) grant to AK.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fvets.2024.1337939/full#supplementary-material>

SUPPLEMENTARY FIGURE 1

Distribution of lncRNA on chromosomes. (A) Distribution of lncRNAs in low fertility (LF) ram group. (B) Distribution of lncRNAs in high fertility (HF) ram group.

SUPPLEMENTARY FIGURE 2

Identification pipeline for lncRNAs. Each step is documented in detail in the Methods section.

SUPPLEMENTARY FIGURE 3

Functional enrichment analysis of target genes of Cis- or Trans-regulated by lncRNAs. Illustrated are normalized enrichment scores for specific sets of target genes. Significantly enriched target gene terms were shown. GO keywords are represented by red color codes, molecular functions by orange ones, biological processes by green ones, and cellular components by green ones (CC).

References

1. MacLaren APC. Ram fertility in South-West Scotland. *Br Vet J.* (1988) 144:45–54. doi: 10.1016/0007-1935(88)90151-0
2. Pardos L, Rubio MTM, Fantova E. The diversity of sheep production systems in Aragón (Spain): characterisation and typification of meat sheep farms. *Spanish J Agric Res.* (2008) 6:497–507. doi: 10.5424/sjar/2008064-344
3. Sendler E, Johnson GD, Mao S, Goodrich RJ, Diamond MP, Hauser R, et al. Stability, delivery and functions of human sperm RNAs at fertilization. *Nucleic Acids Res.* (2013) 41:4104–17. doi: 10.1093/nar/gk-t132
4. Hitit M, Özbek M, Ayaz-Guner S, Guner H, Oztug M, Bodu M, et al. Proteomic fertility markers in ram sperm. *Anim Reprod Sci.* (2021) 235:106882. doi: 10.1016/j.anireprosci.2021.106882
5. Kaya A, Dogan S, Peter V, Kutchy NA, Ross P, Topper E, et al. Sperm proteins ODF2 and PAWP as markers of fertility in breeding bulls. *Cell Tissue Res.* (2021) 387:159. doi: 10.1007/s00441-021-03529-1
6. Ugur MR, Saber Abdelrahman A, Evans HC, Gilmore AA, Hitit M, Arifantini RI, et al. Advances in Cryopreservation of Bull Sperm. *Front Vet Sci.* (2019) 6:268. doi: 10.3389/fvets.2019.00268
7. Long JA. The ‘omics’ revolution of genomic, transcriptomic, proteomic and metabolomic tools to predict male reproductive traits that impact fertility in livestock and poultry. *Anim Reprod Sci.* (2020) 220:106354. doi: 10.1016/j.anireprosci.2020.106354
8. Carrell DT, Aston KI, Oliva R, Emery BR, De Jonge CJ. The “omics” of human male infertility: integrating big data in a systems biology approach. *Cell Tissue Res.* (2016) 363:295–312. doi: 10.1007/s00441-015-2320-7
9. de Azevedo Viana AG, Ribeiro IM, Rodrigues Carvalho RP, Memili E, Moura AA, Machado-Neves M. Functional attributes of seminal proteins in bull fertility: a systematic review. *Reproduction.* (2021) 161:459–75. doi: 10.1530/REP-20-0392
10. Ugur MR, Dinh T, Hitit M, Kaya A, Topper E, Didion B, et al. Amino acids of seminal plasma associated with freezability of bull sperm. *Front Cell Dev Biol.* (2020) 7:347. doi: 10.3389/fcell.2019.00347
11. Parisi AM, Thompson SK, Kaya A, Memili E. Molecular, cellular, and physiological determinants of bull fertility. *Turkish J Vet Anim Sci.* (2014) 38:8. doi: 10.3906/vet-1404-76
12. Özbek M, Hitit M, Kaya A, Jousan FD, Memili E. Sperm functional genome associated with bull fertility. *Front Vet Sci.* (2021) 8:610888. doi: 10.3389/fvets.2021.610888
13. Chen Q, Yan W, Duan E. Epigenetic inheritance of acquired traits through sperm RNAs and sperm RNA modifications. *Nat Rev Genet.* (2016) 17:733–43. doi: 10.1038/nrg.2016.106
14. Feugang JM, Rodriguez-Osorio N, Kaya A, Wang H, Page G, Ostermeier GC, et al. Transcriptome analysis of bull spermatozoa implications for male fertility. *Reprod Biomed Online.* (2010) 21:312–24. doi: 10.1016/j.rbmo.2010.06.022
15. Krawetz SA, Kruger A, Lalancette C, Tagett R, Anton E, Draghici S, et al. survey of small RNAs in human sperm. *Hum Reprod.* (2011) 26:3401–12. doi: 10.1093/humrep/der329
16. Jodar M, Anton E. Small RNAs present in semen and their role in reproduction. In: Inorajadas JA, Gosálvez J, editors. *The Omics Revolution and its Impact on Human Reproductive Medicine.* Cambridge, MA: Academic Press (2018). p. 109–23. doi: 10.1016/B978-0-12-812571-7.00008-3
17. Peng H, Shi J, Zhang Y, Zhang H, Liao S, Li W, et al. A novel class of tRNA-derived small RNAs extremely enriched in mature mouse sperm. *Cell Res.* (2012) 22:1609–12. doi: 10.1038/cr.2012.141
18. Hayashi K, Lopes SMCS, Tang F, Surani MA. Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states. *Cell Stem Cell.* (2008) 3:391–401. doi: 10.1016/j.stem.2008.07.027
19. Luo LF, Hou CC, Yang WX. Small non-coding RNAs and their associated proteins in spermatogenesis. *Gene.* (2016) 578:141–57. doi: 10.1016/j.gene.2015.12.020
20. Watanabe T, Cheng EC, Zhong M, Lin H. Retrotransposons and pseudogenes regulate mRNAs and lncRNAs via the piRNA pathway in the germline. *Genome Res.* (2015) 25:368–80. doi: 10.1101/gr.180802.114
21. García-López J, Alonso L, Cárdenas DB, Artaza-Alvarez H, De Dios Hourcade J, Martínez S, et al. Diversity and functional convergence of small noncoding RNAs in male germ cell differentiation and fertilization. *RNA.* (2015) 21:946–62. doi: 10.1261/rna.048215.114
22. Liu WM, Pang RTK, Chiu PCN, Wong BPC, Lao K, Lee KF, et al. Sperm-borne microRNA-34c is required for the first cleavage division in mouse. *Proc Natl Acad Sci U S A.* (2012) 109:490–4. doi: 10.1073/pnas.1110368109
23. Yuan S, Schuster A, Tang C, Yu T, Ortogero N, Bao J, et al. Sperm-borne miRNAs and endo-siRNAs are important for fertilization and preimplantation embryonic development. *Dev.* (2016) 143:635–47. doi: 10.1242/dev.131755
24. Sharma U. Paternal contributions to offspring health: role of sperm small rnas in intergenerational transmission of epigenetic information. *Front Cell Dev Biol.* (2019) 7:215. doi: 10.3389/fcell.2019.00215
25. Gapp K, Jawaid A, Sarkies P, Bohacek J, Pelczar P, Prados J, et al. Implication of sperm RNAs in transgenerational inheritance of the effects of early trauma in mice. *Nat Neurosci.* (2014) 17:667–9. doi: 10.1038/nn.3695
26. Rodgers AB, Morgan CP, Leu NA, Bale TL. Transgenerational epigenetic programming via sperm microRNA recapitulates effects of paternal stress. *Proc Natl Acad Sci U S A.* (2015) 112:13699–704. doi: 10.1073/pnas.1508347112
27. Skinner MK, Ben Maamar M, Sadler-Riggleman I, Beck D, Nilsson E, McBirney M, et al. Alterations in sperm DNA methylation, non-coding RNA and histone retention associate with DDT-induced epigenetic transgenerational inheritance of disease. *Epigenet Chromatin.* (2018) 11:8. doi: 10.1186/s13072-018-0178-0
28. Gao Y, Li S, Lai Z, Zhou Z, Wu F, Huang Y, et al. Analysis of long non-coding RNA and mRNA expression profiling in immature and mature bovine (*Bos taurus*) testes. *Front Genet.* (2019) 10:646. doi: 10.3389/fgene.2019.00646
29. Zhang X, Gao F, Fu J, Zhang P, Wang Y, Zeng X. Systematic identification and characterization of long non-coding RNAs in mouse mature sperm. *PLoS ONE.* (2017) 12:e0173402. doi: 10.1371/journal.pone.0173402
30. Menezes ESB, Badial PR, El Debaky H, Husna AU, Ugur MR, Kaya A, et al. Sperm miR-15a and miR-29b are associated with bull fertility. *Andrologia.* (2020) 52:e13412. doi: 10.1111/and.13412
31. Kaya A, Memili E. Sperm macromolecules associated with bull fertility. *Anim Reprod Sci.* (2016) 169:88–94. doi: 10.1016/j.anireprosci.2016.02.015
32. Wang X, Yang C, Guo F, Zhang Y, Ju Z, Jiang Q, et al. 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. doi: 10.1038/s41598-018-38462-x
33. Dai YB, Lin Y, Song N, Sun F. LncRNA4667 is dispensable for spermatogenesis and fertility in mice. *Reprod Dev Med.* (2019) 3:18. doi: 10.4103/2096-2924.255985
34. Chen H, Miao X, Xu J, Pu L, Li L, Han Y, et al. Alterations of mRNA and lncRNA profiles associated with the extracellular matrix and spermatogenesis in goats. *Anim Biosci.* (2022) 35:544–55. doi: 10.5713/ab.21.0259
35. Fraser L, Pauksztó L, Mańkowska A, Brym P, Gilun P, Jastrzębski JP, et al. Regulatory potential of long non-coding RNAs (lncRNAs) in boar spermatozoa with good and poor freezability. *Life.* (2020) 10:1–26. doi: 10.3390/life10110300
36. Das PJ, McCarthy F, Vishnoi M, Paria N, Gresham C, Li G, et al. Stallion sperm transcriptome comprises functionally coherent coding and regulatory RNAs as revealed by microarray analysis and RNA-seq. *PLoS ONE.* (2013) 8:e56535. doi: 10.1371/journal.pone.0056535
37. Kim D, Langmead B, Salzberg SL. HISAT a fast spliced aligner with low memory requirements. *Nat Methods.* (2015) 12:357–60. doi: 10.1038/nmeth.3317
38. Perteau M, Kim D, Perteau GM, Leek JT, Salzberg SL. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat Protoc.* (2016) 11:1650–67. doi: 10.1038/nprot.2016.095
39. Sun L, Luo H, Bu D, Zhao G, Yu K, Zhang C, et al. Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts. *Nucleic Acids Res.* (2013) 41:e166. doi: 10.1093/nar/gkt646
40. Finn RD, Bateman A, Clements J, Coghill P, Eberhardt RY, Eddy SR, et al. Pfam: the protein families database. *Nucleic Acids Res.* (2014) 42:D222–30. doi: 10.1093/nar/gkt1223
41. Kong L, Zhang Y, Ye ZQ, Liu XQ, Zhao SQ, Wei L, et al. CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Res.* (2007) 35:W345–9. doi: 10.1093/nar/gkm391
42. Frazee AC, Perteau G, Jaffe AE, Langmead B, Salzberg SL, Leek JT. Ballgown bridges the gap between transcriptome assembly and expression analysis. *Nat Biotechnol.* (2015) 33:243–246. doi: 10.1038/nbt.3172
43. Young MD, Wakefield MJ, Smyth GK, Oshlack A. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* (2010) 11:R14. doi: 10.1186/gb-2010-11-2-r14
44. Liao Y, Wang J, Jaehnig EJ, Shi Z, Zhang B. WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Res.* (2019) 47:W199–205. doi: 10.1093/nar/gkz401
45. Indriastuti R, Pardede BP, Gunawan A, Ulum MF, Arifantini RI, Purwantara B. Sperm transcriptome analysis accurately reveals male fertility potential in livestock. *Animals.* (2022) 12:2955. doi: 10.3390/ani12212955

46. Selvaraju S, Parthipan S, Somashekar L, Kolte AP, Krishnan Binsila B, et al. Occurrence and functional significance of the transcriptome in bovine (*Bos taurus*) spermatozoa. *Sci Rep.* (2017) 7:2392. doi: 10.1038/srep42392
47. Yang H, Wang F, Li F, Ren C, Pang J, Wan Y, et al. Comprehensive analysis of long noncoding RNA and mRNA expression patterns in sheep testicular maturation. *Biol Reprod.* (2018) 99:650–61. doi: 10.1093/biolre/i0y088
48. Gòdia M, Estill M, Castelló A, Balasch S, Rodríguez-Gil JE, Krawetz SA, et al. RNA-seq analysis to describe the boar sperm transcriptome and its seasonal changes. *Front Genet.* (2019) 10:299. doi: 10.3389/fgene.2019.00299
49. Fraser L, Brym P, Pareek CS, Mogielnicka-Brzozowska M, Paukszto L, Jastrzebski JP, et al. Transcriptome analysis of boar spermatozoa with different freezability using RNA-Seq. *Theriogenology.* (2020) 142:400–13. doi: 10.1016/j.theriogenology.2019.11.001
50. Jastrzebski JP, Lipka A, Majewska M, Makowczenko KG, Paukszto L, Bukowska J, et al. *In silico* Identification of lncRNAs regulating sperm motility in the Turkey (*Meleagris gallopavo* L.). *Int J Mol Sci.* (2022) 23:7642. doi: 10.3390/ijms23147642
51. Sahoo B, Gupta MK. Transcriptome analysis reveals spermatogenesis-related CircRNAs and lncRNAs in goat spermatozoa. *Biochem Genet.* (2023) 1–23. doi: 10.1007/s10528-023-10520-8
52. Miller MR, Mannowetz N, Iavarone AT, Safavi R, Gracheva EO, Smith JF, et al. Unconventional endocannabinoid signaling governs sperm activation via the sex hormone progesterone. *Science.* (2016) 352:555–9. doi: 10.1126/science.aad6887
53. Kasimanickam V, Kasimanickam R, Arangasamy A, Saberivand A, Stevenson JS, Kastelic JP. Association between mRNA abundance of functional sperm function proteins and fertility of Holstein bulls. *Theriogenology.* (2012) 78:2007–19.e2. doi: 10.1016/j.theriogenology.2012.07.016
54. Schoff PK, Cheatham J, Lardy HA. Adenylate kinase activity in ejaculated bovine sperm flagella. *J Biol Chem.* (1989) 264:6086–91. doi: 10.1016/S0021-9258(18)83316-6
55. Cao W, Haig-Ladewig L, Gerton GL, Moss SB. Adenylate kinases 1 and 2 are part of the accessory structures in the mouse sperm flagellum. *Biol Reprod.* (2006) 75:492–500. doi: 10.1095/biolreprod.106.053512
56. Kawashima A, Osman BAH, Takashima M, Kikuchi A, Kohchi S, Satoh E, et al. CABS1 is a novel calcium-binding protein specifically expressed in elongate spermatids of mice. *Biol Reprod.* (2009) 80:1293–304. doi: 10.1095/biolreprod.108.073866
57. Zhang X, Zhou W, Zhang P, Gao F, Zhao X, Shum WW, et al. Cabs1 maintains structural integrity of mouse sperm flagella during epididymal transit of sperm. *Int J Mol Sci.* (2021) 22:652. doi: 10.3390/ijms22020652
58. Fiedler SE, Carr DW. ROPN1, a Protein kinase A-like (R2D2) protein, regulates sperm motility. *Biol Reprod.* (2011) 85:804. doi: 10.1093/biolreprod/85.s1.804
59. Fiedler SE, Dudiki T, Vijayaraghavan S, Carr DW. Loss of R2D2 proteins ROPN1 and ROPN1L causes defects in murine sperm motility, phosphorylation, and fibrous sheath integrity. *Biol Reprod.* (2013) 41:88. doi: 10.1095/biolreprod.112.105262
60. Choi I, Woo JM, Hong S, Jung YK, Kim DH, Cho C. Identification and characterization of ADAM32 with testis-predominant gene expression. *Gene.* (2003) 304:151–62. doi: 10.1016/S0378-1119(02)01202-7
61. Civetta A. Positive selection within sperm-egg adhesion domains of fertilinn ADAM gene with a potential role in fertilization. *Mol Biol Evol.* (2003) 20:21–9. doi: 10.1093/molbev/msg002
62. Nethererton JK, Hetherington L, Ogle RA, Velkov T, Baker MA. Proteomic analysis of good-and poor-quality human sperm demonstrates that several proteins are routinely aberrantly regulated. *Biol Reprod.* (2018) 99:395–408. doi: 10.1093/biolre/iox166
63. Lee S, Hong SH, Cho C. Normal fertility in male mice lacking ADAM32 with testis-specific expression. *Reprod Biol.* (2020) 20:589–94. doi: 10.1016/j.repbio.2020.09.001
64. Park YJ, Pang WK, Ryu DY, Song WH, Rahman MS, Pang MG. Optimized combination of multiple biomarkers to improve diagnostic accuracy in male fertility. *Theriogenology.* (2019) 139:106–12. doi: 10.1016/j.theriogenology.2019.07.029
65. Li H, He Y, Yan J, Zhao Q, Di C, Zhang H. Comparative proteomics reveals the underlying toxicological mechanism of low sperm motility induced by iron ion radiation in mice. *Reprod Toxicol.* (2016) 65:148–58. doi: 10.1016/j.reprotox.2016.07.014