-Original Article-

Two novel testis-specific long noncoding RNAs produced by 1700121C10Rik are dispensable for male fertility in mice

Chaojie LI¹⁾, Chunling SHEN¹⁾, Xuan SHANG¹⁾, Lingyun TANG¹⁾, Wenfeng XIONG¹⁾, Haoyang GE¹⁾, Hongxin ZHANG¹⁾, Shunyuan LU¹⁾, Yan SHEN¹⁾, Jinjin WANG²⁾, Jian FEI²⁾ and Zhugang WANG^{1, 2)}

¹⁾State Key Laboratory of Medical Genomics, Research Center for Experimental Medicine, Shanghai Ruijin Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China
²⁾Shanghai Research Center for Model Organisms, Shanghai 201203, China

Abstract. Testis-specific genes are prone to affect spermatogenesis or sperm fertility, and thus may play pivotal roles in male reproduction. However, whether a gene really affects male reproduction *in vivo* needs to be confirmed using a gene knockout (KO) model, a 'gold standard' method. Increasing studies have found that some of the evolutionarily conserved testisenriched genes are not essential for male fertility. In this study, we report that *1700121C10Rik*, a previously uncharacterized gene, is specifically expressed in the testis and produces two long noncoding RNAs (lncRNAs) in mouse: Transcript 1 and Transcript 2. qRT-PCR, northern blotting, and *in situ* hybridization revealed that expression of both the lncRNAs commenced at the onset of sexual maturity and was predominant in round and elongating spermatids during spermiogenesis. Moreover, we found different subcellular localization of Transcript 1 and Transcript 2 that was predominant in the cytoplasm and nucleus, respectively. *1700121C10Rik*-KO mouse model disrupting Transcript 1 and Transcript 2 expression was generated by CRISPR/Cas9 to determine their role in male reproduction. Results showed that *1700121C10Rik*-KO male mice were fully fertile with approximately standard testis size, testicular histology, sperm production, sperm morphology, sperm motility, and induction of acrosome reaction. Thus, we conclude that both the testis-specific *1700121C10Rik*-produced lncRNAs are dispensable for male fertility in mice under standard laboratory conditions.

Key words: 1700121C10Rik, Knock-out model, Long noncoding RNA, Male fertility, Testis

(J. Reprod. Dev. 66: 57–65, 2020)

n mammals, male gametes are produced in testis through spermatogenesis, which involves renewal of spermatogonia, production of round haploid spermatids, and their differentiation into elongated spermatids, prior to their release into tubule lumen. It is anticipated that approximately 1 in 25 of all mammalian genes are specifically expressed in the male germline, thus exemplifying the complexity of spermatogenesis and indicates that potential mutation in thousands of different genes can affect male fertility [1, 2].

Moreover, emerging evidence suggests that male reproduction is controlled not only by the protein-coding segments of the genome, but also by noncoding regions including loci that produce microR-NAs (miRNAs), piwi-interacting RNAs (piRNAs), circular RNAs (circRNAs), long noncoding RNAs (lncRNAs), and other noncoding genes [3–5]. lncRNAs are mRNA-like transcripts that are more than 200 nucleotides long and not translated into proteins. Many of the lncRNAs have been reported to exhibit intense tissue-specific expression patterns such as in the brain and testis, which are reported to express the leading number of lncRNAs [6]. lncRNAs were previously thought to be 'transcriptional noise' with no biological functions. However, increasing evidences indicate that lncRNAs are involved in diverse biological processes such as X chromosome inactivation, genome imprinting, and embryo development [7, 8]. Furthermore, they have been reported to play important roles in the occurrence and development of various diseases [9-11]. Currently, several IncRNAs have been identified to be involved in male reproductive processes [12]. For instance, lncRNA AK015322 has been suggested to promote proliferation of mouse spermatogonial stem cell line (C18-4) by acting as a decoy for microRNA-19b-3p [13]. Similarly, lncRNA 033862 has been shown to support survival of mouse spermatogonial stem cells by interacting with GDNF family receptor alpha 1 (Gfra1) [14]. lncRNA Mrhl has been reported to mediate meiotic commitment of mouse spermatogonial cells by regulating SRY-box transcription factor 8 (Sox8) expression [15]. As lncRNAs are the largest and most complex class of noncoding RNAs; their unique role, especially of testis-specific lncRNAs in male reproduction, remains to be investigated. They may be potential causes and/or therapeutic targets in male fertility disorders.

1700121C10Rik is a RIKEN cDNA 1700121C10 gene, which is identified by RIKEN Mouse Gene Encyclopedia Project as a novel full-length mouse cDNA in the testis of an adult mouse. Open Reading Frame (ORF) analysis have revealed that *1700121C10Rik* mRNA

Received: August 27, 2019

Accepted: November 11, 2019

Advanced Epub: December 5, 2019

^{©2020} by the Society for Reproduction and Development

Correspondence: Z Wang (e-mail: zhugangw@shsmu.edu.cn) and C Shen (e-mail: clshen@139.com)

This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial No Derivatives (by-nc-nd) License. (CC-BY-NC-ND 4.0: https://creativecommons.org/licenses/by-nc-nd/4.0/)

contains three putative ORFs, all of which can potentially encode less than 100 amino acids (aa). Micropeptides, shorter than 100 aa, have been usually reported to be encoded by small ORFs [16-18]. Small ORFs were previously ignored due to their small size and lack of protein-coding evidence. However, increasing evidences suggest that small ORF-encoded micropeptides play important roles in many fundamental biological processes including important correlations in pathogenesis [19, 20]. This has attracted considerable interest of the scientific community for the in-depth study on lncRNAs. It remains unknown whether 1700121C10Rik encodes micropeptides, and what are their biological functions. Our group focuses on dissecting the functions of novel testis-specific genes in male reproduction by establishing transgenic and KO mouse models [21, 22]. To clarify the biological functions of 1700121C10Rik in testis and its potential role in male reproduction, we characterized the expression pattern of 1700121C10Rik, identified its full-length cDNA sequences, and evaluated their protein-coding potential. Furthermore, we established 1700121C10Rik-KO mouse model and analyzed the reproductive phenotype of these mice. Results showed that 1700121C10Rik consists of three exons, and its expression leads to two transcripts that are testis-specific lncRNAs. However, 1700121C10Rik-KO male mice were found to be fully fertile, which indicated that 1700121C10Rik is dispensable for male fertility in mice.

Materials and Methods

Ethics statement

All research protocols involving animal experiments were approved by Institutional Animal Care and Use Committee of Shanghai Research Center for Model Organisms.

Northern blotting

Total RNA (30 µg each) was isolated from mouse testes and other tissues, and subjected to northern blot analysis. DNA templates containing T7 or T3 RNA polymerase promoter site were generated by PCR reactions with specific set of primers: forward primer, 5'-TAATACGACTCA CTATAGGGAGAATCTTCCTACGTA-CTCCCCTTTAGATGATC-3' and reverse primer, 5'-AATTAA-CCCTCACTAAAGGGAGATCTAATCATTTATTATTCTCCA-GCAGTCCAAGG-3'. Further, they were used for *in vitro* transcription using MAXIscript Kit (Thermo Fisher Scientific, Rochester, NY, USA) to synthesize single-stranded digoxigenin (DIG)-labeled RNA probes according to the manufacturer's protocol. Hybridization was performed using NorthernMax-Gly Kit (Thermo Fisher Scientific) according to the manufacturer's instructions.

Rapid amplification of cDNA ends (RACE)

5'- and 3'- RACE were performed using SMARTer RACE 5'/3' Kit (Takara Bio, Dalian, China) according to the manufacturer's instructions. RNA was isolated from the testes of adult mice. Primers were designed based on the known sequence information, and their sequences are listed in Supplementary Table 1 (online only).

RT-PCR and qRT-PCR

Total RNA was extracted from mouse tissues and cells using TRIzol Plus RNA Purification Kit (Invitrogen, Carlsbad, CA, USA) and reverse transcribed into cDNA using PrimeScript RT Master Mix (Takara Bio) following manufacturer's instructions. cDNAs were amplified using specific set of primers as illustrated in Supplementary Table 1 for semi-quantitative or real-time RT-PCR. Semi-qRT-PCR products were separated by electrophoresis on 1.5% agarose gel and visualized by ethidium bromide staining. RT-PCR was performed by Mastercycler ep realplex (Eppendorf, Hamburg, Germany) using SYBR Premix Ex Taq Kit (Takara Bio). Product of interest was resolved from nonspecific amplification by melt curve analysis. Gene expression levels were normalized to β -Actin (*Actb*) and analyzed by 2^{- $\Delta\Delta$ Ct} or 2^{- Δ Ct} method.

Coding potential assessment

Coding potential of *1700121C10Rik* RNA was assessed by Coding Potential Calculator (CPC) (http://cpc.cbi.pku.edu.cn/) and Coding-Potential Assessment Tool (CPAT) (http://lilab.research.bcm.edu/ cpat/) [23, 24].

Cell transfection and western blot analysis

Three putative ORFs (sequences shown in Fig. 2B) within 1700121C10Rik RNA and TAFA chemokine like family member 2 (Tafa2) ORF were cloned respectively from mouse testis and brain cDNA library, and amplified using specific set of primers as listed in Supplementary Table 1. 1700121C10Rik and Tafa2 ORFs were cloned in frame with FLAG into pEGFP-N2 vector (BD Biosciences, San Jose, CA, USA) and pcDNA3.1 (+) vector (Invitrogen). Mouse spermatocyte cell line GC-2spd(ts) and human 293T cells were transfected with DNA constructs using Lipofectamine 3000 (Invitrogen) according to the manufacturer's instructions and harvested for 48 h later. Proteins were extracted from cell pellets using RIPA lysis buffer (Beyotime Biotechnology, Shanghai, China) with protease inhibitor cocktails (Roche, Basel, Switzerland). In addition, constructs with T7 promoter were used for in vitro transcription and translation (IVT) using TnT Quick Coupled Transcription/Translation Systems (Promega, Madison, WI, USA). Further, proteins were separated on 15% SDS-PAGE gels, and transferred onto nitrocellulose membranes. Membranes were then blocked with 5% nonfat milk in PBS for 1 h followed by incubation with primary antibodies: anti-FLAG (MBL, Woburn, MA, USA) and anti-glyceraldehyde 3-phosphate dehydrogenase (GAPDH) (Sigma-Aldrich, St. Louis, MO, USA). Membranes were incubated with secondary antibodies conjugated with IRdye 800CW (LI-COR, Lincoln, NE, USA) and then, visualized by Odyssey infrared imager (LI-COR). GAPDH was used as an internal control.

Immunofluorescence staining

Human 293T cells were plated on glass coverslips and transfected with ORF1-Flag, ORF2-Flag, ORF3-Flag, and *Tafa2* ORF-Flag vectors for 48 h. Further, cells were fixed with 4% paraformaldehyde, permeabilized with 0.1% Triton X-100, incubated with anti-FLAG, and subsequently, incubated with corresponding Alexa Fluor 568-conjugated secondary IgG antibodies (Molecular Probes, Eugene, OR, USA). Cellular nuclei were counterstained with DAPI. Slides were mounted with fluorescence mounting medium (DAKO, Glostrup, Denmark), cover slipped, and examined under fluorescence microscope (Nikon Eclipse 80i, Nikon, Tokyo, Japan).

In situ hybridization

Adult mouse testes were fixed with Bouin's buffer, embedded in paraffin, and sliced into 7 μ m thick sections. Same PCR product as from northern blotting was used for *in vitro* transcription to synthesize DIG-labeled RNA probes. Hybridization was performed according to a previously reported protocol [25]. After hybridization and washing, the sections were incubated overnight with HRP-coupled anti-DIG antibody at 4°C. The reaction with tyramide-DNP was performed with TSA Plus DNP (AP) System (PerkinElmer, Boston, MA, USA) according to the manufacturer's instructions. After incubating with BCIP/NBT substrate (Vector Laboratories, Burlingame, CA, USA), the sections were dehydrated, mounted, and then, observed under light microscope.

Preparation of subcellular fractions of testicular cells

Isolation of testicular cells was performed according to previously published protocol [26] with minor modifications. Briefly, testes were cut into 2-3 mm³ pieces after removal of the tunica albuginea. These pieces were then placed into a disposable disaggregator Medicon with 1 ml of ice-cold PBS and processed for 50 sec in Medimachine System (BD Biosciences). Cell suspension was recovered from Medicon unit using disposable syringe, and subsequently, filtered through 50 µm Filcon (BD Biosciences) and 25 µm nylon mesh. Further, cells were washed once and resuspended in ice-cold PBS. Nuclear and cytoplasmic fractions of testicular cells were separated using NE-PER Nuclear and Cytoplasmic Extraction Reagents (Thermo Fisher Scientific) according to the manufacturer's instructions. Two set of primers were used to check the quality of the two fractions as previously described [27, 28]. Gapdh (in5-ex6) was amplified to detect immature mRNA in the nucleus, and Gapdh (ex5-ex6) was used to identify mature mRNA in the cytoplasm.

Mice and genotyping

KO mice were generated by non-homologous recombination using CRISPR/Cas9. Guide RNAs and Cas9 mRNA were produced by *in vitro* transcription. Cas9 mRNA (10 ng/µl) was co-injected with two guide RNAs (10 ng/µl) into fertilized eggs of C57BL/6J mice. For further development, fertilized eggs were placed in pseudo-pregnant females. Genotypes of mutant mice were determined by PCR using following primers: P1, 5'-CAACCGACCATAGACCCAAAATA-3'; P2, 5'-TGCCCAGGACCTTAGATGTGC-3'; and P3, 5'-GGCTTCCAGTGCTTCAGACAGG-3'. The resulting bands were verified by Sanger sequencing. Additionally, semi-qRT-PCR and qRT-PCR analyses were performed to examine *1700121C10Rik* expression in mouse testes of three different genotypes.

Characterization of 1700121C10Rik-KO male mice

To evaluate the reproductive capability of male mice, four adult male mice of each genotype were caged with wild-type (wt) females at 1:2 sex ratio for a month. Any plugged female was removed from cage and replaced with new ones. Number of total females, plugged females, litters, and offspring was recorded to calculate frequency of copulatory plug (FCP) and frequency of conception (FC). For histology, testes from adult age-matched mice with different genotypes were removed and weighed. Testes and cauda epididymides were fixed in Bouin's buffer, embedded in paraffin, sliced into 5 μ M thick tissue sections, and stained with hematoxylin and eosin (H&E). Sperms were isolated from cauda epididymides of wt and *1700121C10Rik*^{-/-} mice, and spotted onto glass slides. Sperm smears were stained with H&E and Alexa Fluor 488-conjugated lectin peanut agglutinin (PNA, Molecular Probes), respectively. Motility and acrosome reaction of sperm from wt and *1700121C10Rik*^{-/-} mice were also assessed as previously described [21].

Statistical analysis

All data are represented as mean values \pm standard error of the mean (SEM) (n \geq 3), unless otherwise stated. Student's *t*-test was used for statistical analysis. P < 0.05 was considered as statistically significant.

Results

Testis-specific and age-dependent expression pattern of 1700121C10Rik in mice

Total RNAs were prepared from 12 adult mouse tissues including the heart, kidney, liver, lung, spleen, testis, epididymis, brain, stomach, intestine, ovary, and uterus. We investigated tissue expression profile of 1700121C10Rik by qRT-PCR and northern blotting. qRT-PCR results showed that 1700121C10Rik mRNA is predominantly expressed in the testis tissue. Meanwhile, northern blot analysis revealed a band of approximate 500 bp, which was exclusively expressed in the testis (Fig. 1A). It is known that testis maturation is agedependent and the first wave of spermatogenesis has been reported to take place within initial 35 days postnatal [29]. To determine the temporal and spatial expression of 1700121C10Rik in testis, we further investigated tissues from postnatal day 14 to day 90 for the expression pattern of 1700121C10Rik mRNA in developing mouse testes by qRT-PCR and northern blotting. Results revealed that 1700121C10Rik mRNA initially appeared at postnatal day 28 and was then consistently expressed throughout adulthood (Fig. 1B). This suggested that 1700121C10Rik expression in the testis corresponds with the sexual maturity of mice. Such a specific expression pattern might imply an important physiological role of 1700121C10Rik during spermiogenesis in mice.

5'- and 3'-RACE identified two transcripts that are produced by 1700121C10Rik

On searching National Center for Biotechnology Information (NCBI) and Mouse Genome Informatics (MGI) database, we obtained few incomplete transcript sequences of *1700121C10Rik*. To further acquire the full-length sequence of *1700121C10Rik* and identify additional transcripts that may exist, 5'- and 3'-RACE experiments were performed using gene-specific primers (GSPs) that were designed based on known sequence information. As a result, additional 5'-end sequences of 53 bp and 219 bp that match genomic sequences were obtained by 5'-RACE. Two transcriptional start sites were identified for *1700121C10Rik* among 20 subclones using 5'-RACE. On the other hand, 3'-RACE revealed one transcriptional termination site among eight subclones. Consequently, two different transcript variants of *1700121C10Rik* were identified in the testis of adult mice, namely, Transcript 1 and Transcript 2, respectively (Fig. 2A). It is noteworthy that Transcript 2 was found to be a novel transcript,

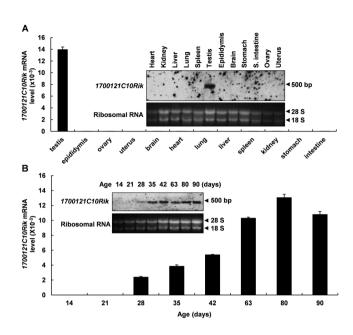


Fig. 1. Expression pattern of *1700121C10Rik* in the mouse. (A) Testisspecific expression of *1700121C10Rik* mRNA as revealed by qRT-PCR and northern blotting. (B) Developmental stagedependent expression of *1700121C10Rik* mRNA in mouse testes as examined by qRT-PCR and northern blotting. Error bars indicate standard error of the mean (SEM). Agarose gel electrophoresis images are visualized by ethidium bromide staining of RNA samples.

which has never been recorded in a public database. Consistent with our northern blot analysis, full-length cDNA sequences of Transcript 1 and Transcript 2 with polyadenylic acid [poly(A)] tail were found to be 479 bp and 486 bp, respectively (Fig. 2B). Their nucleotide sequences were submitted to GenBank database under accession numbers, MK913899 and MK913900, respectively. Subsequently, the expression patterns of the transcripts in different tissues of adult mice and developing mouse testes were detected by RT-PCR. Results showed that expression pattern of the two transcripts that are produced by *1700121C10Rik* in mice is testis-specific and age-dependent (Supplementary Fig. 1: online only).

1700121C10Rik is a non-coding gene

Although *1700121C10Rik* is annotated as a protein-coding gene in NCBI Gene Database (Gene ID: 67364), its consensus coding sequence has not been identified. We analyzed the putative ORFs within *1700121C10Rik* transcripts using ORF finder (https://www. ncbi.nlm.nih.gov/orffinder/). We found three putative ORFs in Transcript 1, and the longest ORF was observed to putatively encode 67 aa. Conversely, no ORF was found with respect to Transcript 2 (Fig. 2B). Moreover, both transcripts of *1700121C10Rik* were observed to lack a Kozak sequence for translation initiation and did not have detectable sequence conservation among other species. This suggested that *1700121C10Rik* RNAs may not be translated into proteins. Subsequently, CPC and CPAT were used to assess the coding potential of *1700121C10Rik* RNAs. Results revealed that *1700121C10Rik* may be a non-coding gene (Fig. 3A).

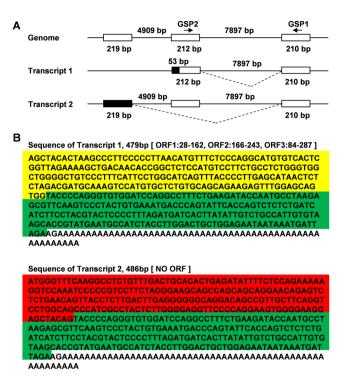
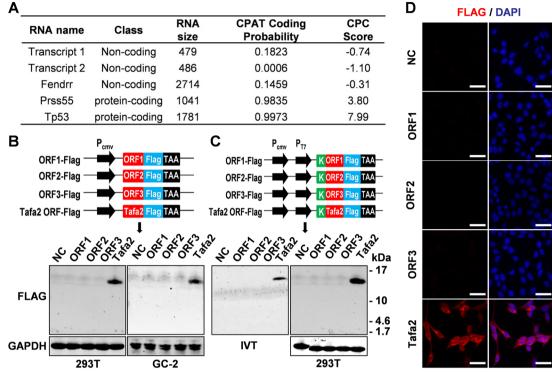


Fig. 2. Two transcript variants of 1700121C10Rik: Transcript 1 and Transcript 2, as identified by 5'- and 3'-rapid amplification of cDNA ends (RACE). (A) Schematic diagram of the genomic structure of the mouse 1700121C10Rik and associated different transcript variants. The boxes represent exons and the black region represents newly identified exon sequences. Gene-specific primers, GSP1 and GSP2 for RACE and their relative positions are indicated. (B) Full-length cDNA sequences of Transcript 1 and Transcript 2 with a polyadenylic acid [poly(A)] tail. Exon sequences are indicated as exon1 (red-shaded bases), exon2 (yellow-shaded bases) and exon3 (green-shaded bases).

To further investigate the protein-coding potential of 1700121C10Rik, three different ORFs within 1700121C10Rik RNA were cloned in frame with FLAG into pEGFP-N2 vector. DNA constructs were then transiently transfected into mouse spermatocyte cell line GC-2spd(ts) and human 293T cells. Further, we performed western blot analyses using anti-FLAG antibody. However, the peptide corresponding to the predicted protein was not found. Conversely, TAFA2 proteins encoded by a known protein-coding gene were found to be consistently produced (Fig. 3B). These results indicated that 1700121C10Rik ORFs are not translated. To further confirm the obtained data, we cloned the ORFs into an expression vector containing T7 promoter, and assessed if they could be translated in a cell-free translation system. To enhance the efficiency of translation, we also cloned Kozak sequence into these constructs, and transfected them into human 293T cells. The predicted proteins were neither detected in IVT system nor in cell lysates. By contrast, TAFA2 proteins were easily detected (Fig. 3C). Furthermore, the finding was confirmed by immunofluorescence staining (Fig. 3D). Taken together, we concluded that both the transcripts of 1700121C10Rik are lncRNAs.



Two transcripts of 1700121C10Rik as long noncoding RNAs (IncRNAs). (A) Coding probability scores of the transcripts were assessed by Fig. 3. coding-potential assessment tool (CPAT) and coding-potential calculator (CPC). Serine protease 55 (Prss55) and tumor protein 53 (Tp53) were used as positive controls while FOXF1 antisense RNA 1 (Fendrr) as a negative control. (B) Western blot analysis of human 293T cells and mouse spermatocyte cell line GC-2spd(ts) transfected with depicted constructs were detected using anti-FLAG and anti-glyceraldehyde 3-phosphate dehydrogenase (GAPDH) antibodies. GAPDH was used as loading control. (C) Western blot analysis of in vitro-translated reactions and human 293T cells transfected with depicted constructs were detected using anti-FLAG and anti-GAPDH antibodies. GAPDH was used as loading control. K, Kozak sequence. (D) The constructs shown in (C) were transfected into human 293T cells and FLAG fusion proteins were immunostained using anti-FLAG antibody. Scale bar, 20 µm; NC, negative control.

Cellular and subcellular localization of 1700121C10Rik RNAs in mouse testis

To determine the testicular cell types that express 1700121C10Rik RNAs, in situ hybridization with highly sensitive tyramide signal amplification system was performed on paraffin sections of adult mouse testes using DIG-labeled antisense and sense RNA probes (Fig. 4A). Hybridization using antisense probe demonstrated strong signals in the seminiferous epithelium whereas few signals were detected when a sense RNA probe was used as a negative control. Intense staining of 1700121C10Rik RNA was predominantly observed in round and elongating spermatids, which are located in the inner half-layer of the seminiferous tubules.

Conventionally, mRNAs are exported to the cytoplasm where they are recognized and bound by ribosomes. Conversely, lncRNAs are distributed in cytoplasm as well as nucleus, and their specific subcellular localization is often closely related to their function in cell. Thus, studying the subcellular localization of lncRNA and its dynamic changes is considered as a crucial step in elucidating the functions and mechanisms of newly discovered lncRNAs [30]. To determine the subcellular localization patterns of 1700121C10Rik RNAs, we extracted RNA from testicular cells into nuclear and cytoplasmic fractions. RT-PCR analysis revealed the different distribution pattern of 1700121C10Rik RNAs in nuclear and cytosolic components of testicular cells (Fig. 4B). Transcript 1 was found to be predominantly expressed in the cytoplasm while Transcript 2 was mainly expressed in the nucleus, which was further confirmed by qRT-PCR (Fig. 4C). Specific subcellular localization patterns of 1700121C10Rik RNAs in testicular cells suggested that 1700121C10Rik might be functioning through multiple mechanisms, which we plan to investigate in the future.

Generation of 1700121C10Rik-KO mice

The transcribed region of mouse 1700121C10Rik includes three exons spanning approximately 13.4 kb on mouse chromosome 8C5. To investigate the physiological functions of 1700121C10Rik in vivo, we deleted the entire transcribed region of 1700121C10Rik using CRISPR/Cas9 genome editing system (Fig. 5A). As a result, we successfully produced two mutant mouse lines with complete deletion of 1700121C10Rik. As we did not observe evident defects or impaired fertility in both the 1700121C10Rik-deficient mouse lines, we focused on only one of them. Deletion of the transcribed region was confirmed through PCR genotyping and Sanger sequencing (Fig. 5B-D). Further, RT-PCR and qRT-PCR analyses were performed to confirm the absence of 1700121C10Rik expression at RNA level (Fig. 5E and 5F).

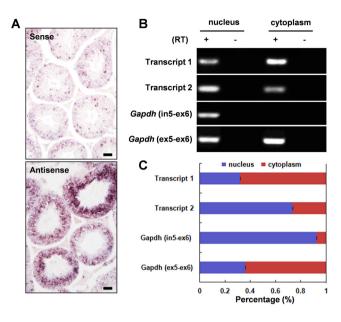


Fig. 4. Cellular and subcellular localization of *1700121C10Rik* RNAs in mouse testis. (A) Localization of *1700121C10Rik* RNAs in adult mouse testis as examined by *in situ* hybridization using DIG-labeled antisense RNA or sense RNA probe. Scale bar, 25 μm. (B) Quantity of each transcript in the nucleus and cytoplasm of testicular cells was examined by RT-PCR using oligo(dT) primer with reverse transcriptase (RT+) or without (RT-). *Gapdh* (in5-ex6) was used as a control for cytoplasmic expression. (C) qRT-PCR analyses corresponding to the quantity of each transcript in the nucleus and cytoplasm of testicular cells.

1700121C10Rik deficiency does not impair male fertility

The above data suggested that 1700121C10Rik may play important roles in regulating male reproduction in mouse. This impelled us to further investigate the effect of 1700121C10Rik deficiency on male fertility. We observed no differences in the ability of male mice of any genotype to plug wt females. Moreover, plugged females mated with 1700121C10Rik-KO males were found to be pregnant. There were no significant differences found in the number of offspring, which were produced by female mice that mated with adult males of three different genotypes (Table 1). Furthermore, on reaching sexual maturity, wt, 1700121C10Rik+/-, and 1700121C10Rik-/- littermates were found to be similar with respect to testicular size (Fig. 6A) and weight (Fig. 6B). H&E staining results revealed no differences in gross morphology and histology of testis, cauda epididymidis, and sperm among these groups (Fig. 6C and 6D). In addition, no differences in the integrity of acrosome and timely occurrence of acrosome reaction of the sperm from wt and mutant mice were detected (Fig. 6E and 6F). Finally, the motility of sperm, which was isolated from the cauda epididymis was evaluated by Computer-Assisted Semen Analysis (CASA) and no significant differences in the movement characteristics were detected (Supplementary Table 2: online only). These results demonstrated that 1700121C10Rik is not required for male fertility in mice.

Discussion

Spermiogenesis is a complex cellular differentiation process that occurs in the late stage of spermatogenesis [31]. During spermiogenesis, round spermatids have been shown to undergo a series of

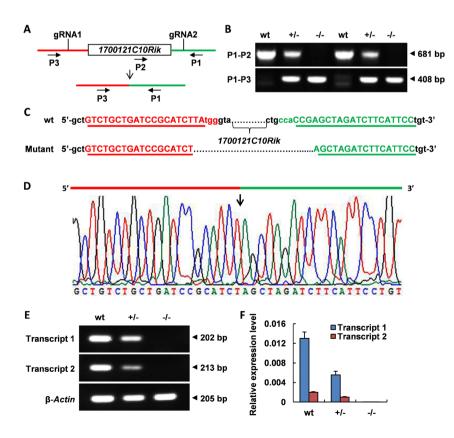


Fig. 5. Generation of 1700121C10Rik-KO mice. (A) Schematic diagram of the targeting strategy of 1700121C10Rik using CRISPR/ Cas9. Guide RNAs (gRNA1 and gRNA2) were designed to completely delete the transcribed region of 1700121C10Rik. P1-P3 indicate genotyping primers and their relative positions. (B) Identification of mice genotypes by PCR. (C) Mutant mouse line was identified and its sequence is compared to wild-type (wt) sequence as shown. gRNA1 sequence (red capitalized bases), gRNA2 sequence (green capitalized bases), and their respective protospacer adjacent motif (PAM) sequence (3 colored lowercase bases) are indicated. (D) Sequencing analysis of the mutant mouse line. Expression of 1700121C10Rik in the testes of wt, heterozygous (+/-), and homozygous (-/-) mice were examined by RT-PCR (E) and qRT-PCR (F) analyses. β -Actin was used as an internal control. Error bars indicate SEM.

*			•			
Male mice	Female mice	Plugs	Litters	Offspring (M/F)	FCP (%)	FC (%)
Wild-type $(n = 4)$	38	28	17	112 (55/57)	73.7	60.7
$1700121C10Rik^{+/-} (n = 4)$	41	33	22	163 (80/83)	80.5	66.7
$1700121C10Rik^{-/-} (n = 4)$	36	28	21	135 (53/82)	77.8	75.0

Table 1. Disruption in 1700121C10Rik does not affect male fertility

FCP, frequency of copulatory plug; FC, frequency of conception; M, male; F, female. FCP (%) = (plugs/females) \times 100; FC (%) = (litters/plugs) \times 100.

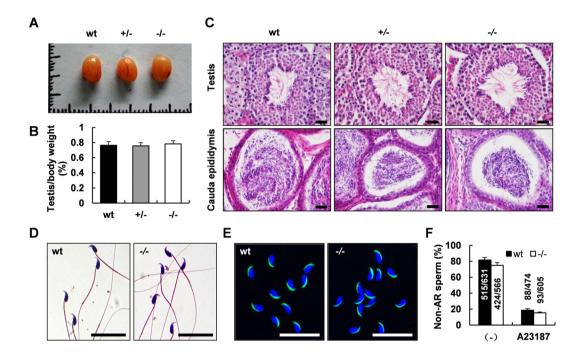


Fig. 6. Assessment of testis size, testicular histology, sperm production, sperm morphology, and acrosome reaction. (A) Representative images of testes from mice of three different genotypes. (B) Testis/body weight ratio of mice of three genotypes (n = 5, each genotype). Error bars indicate SEM. (C) Hematoxylin & eosin (H&E) staining of testis and cauda epididymis of mice corresponding to three different genotypes. Scale bar, 20 µm. (D)–(E) Sperm morphology was evaluated by H&E staining (D) and PNA staining (E). Scale bar, 20 µm. (F) Percentage of non-acrosome-reacted (non-AR) sperm before and after A23187 induction were analyzed. Error bars indicate SEM.

dynamic morphological changes such as nuclear condensation and elongation, formation of flagella and acrosome, reorganization of organelles, cytoplasmic removal, and eventually, differentiation into spermatozoa [32]. Approximately half of the testis-specific genes have been reported to be exclusively expressed during spermiogenesis [33]. Many of these genes have been inactivated in mice to generate KO mouse model, and some of these gene-KO mice have been observed to display male infertility due to nonfunctional sperm with no or different degrees of structural abnormalities [34, 35]. In the present study, we cloned and characterized a novel gene, 1700121C10Rik, from adult mouse testes, which consists of three exons and produces two transcripts. Among the 12 mouse tissues examined, 1700121C10Rik was found to be expressed exclusively in the testis (Fig. 1A). Furthermore, it was initially detected at postnatal day 28 and was found to persist at high levels throughout adulthood (Fig. 1B). Specific types of spermatogenic cells have been shown to produce at different developmental stages of mouse testis [29]. The

first wave of spermatogenesis is known to take place within initial 35 days postnatal, and during the first 21 days after birth, spermatogonia and spermatocytes are known to be the main cell types in the mouse testis. Following, spermatids have been reported to become the main cell types in the later stages. The temporal expression pattern of *1700121C10Rik* as determined from the obtained results suggested that it is expressed in spermatids; however, not in spermatocytes or spermatogonia. Moreover, *in situ* hybridization revealed that positive expression was mainly observed at the luminal side of the seminiferous tubules (Fig. 4A), where round and elongating spermatids are reported to be predominantly distributed [36]. Although some spermatocytes or spermatogonia appeared stained, on co-analyzing qRT-PCR and northern blotting, we hypothesized that the signals in the spermatocytes or spermatogonia might be non-specific or affected by the signals from spermatids.

Further experiments revealed that the two *1700121C10Rik*produced transcripts belong to the class of lncRNAs. Interestingly, they were found to have different subcellular localizations in the testicular cells as validated by semi-qRT-PCR (Fig. 4B) and qRT-PCR (Fig. 4C). Previously, it has been reported that majority of lncRNAs are expressed at a lower level than protein-coding genes. Moreover, these RNAs generally display poor primary sequence conservation over evolution [37, 38], which is consistent with the lack of orthologs of 1700121C10Rik among other species. In contrast to mRNAs, IncRNAs are found to be widely distributed in many kinds of cell compartments and organelles, and their specific subcellular localization is often reported to be closely related to their functions in cell [30]. Different subcellular localizations of the two lncRNAs may imply their functional differences in the testicular cells. They may be serving as important regulators in gene expression networks by controlling chromosome architecture and transcription in the nucleus; and modulating mRNA stability, translation, and post-translational modifications in the cytoplasm as suggested by other researchers [39].

Recent genome-wide studies have indicated that lncRNAs are key regulators of cellular processes and development [40, 41]. Although the studies have been generally performed in cell lines with lncRNA overexpression or knockdown, phenotypes at the cellular and organismal levels are frequently discrepant. For example, loss-of-function studies of Malat1 or Neat1 in mouse have revealed subtle or undetectable phenotypes, despite the fact that these lncRNAs have been suggested to be important at the cellular level [42, 43]. Therefore, further in vivo analyses of lncRNAs are necessary to reveal their mechanism of action. We investigated the potential role of 1700121C10Rik in vivo using a global 1700121C10Rik-KO mouse model. Surprisingly, our study revealed that loss of 1700121C10Rik has no apparent effect on the reproductive capability of male mice (Table 1). Although the offspring of 1700121C10Rik^{-/-} males seemed to be more biased to female progeny, no statistically significant difference was observed among the three genotypes (P > 0.05, data not shown). Furthermore, we found that testicular histology, sperm production, sperm morphology, sperm motility, and acrosome reaction of sperm from $1700121C10Rik^{-/-}$ mice were not significantly different from those in wt mice. One possibility to explain these observations is that other lncRNAs with redundant functions to 1700121C10Rik may compensate for 1700121C10Rik loss. Previous studies have found that deletion of testis-enriched genes can be dispensable for male fertility due to potential redundancy [44-49]. Thus, our results are not completely surprising. Recently, Wen et al. used their optimized CRISPR system to delete 105 testis-specific Drosophila lncRNAs [50]. They found that 33 lncRNA knockouts exhibited developmental defects in late spermatogenesis, resulting in low or no male fertility, whereas the remaining exhibited normal male fertility. Therefore, male germ cells may have compensatory mechanisms for some critical mRNAs and lncRNAs to progress through spermatogenesis [51]. In addition, an emerging function for lncRNAs is their contribution to various genetic programmes that enable response to different environmental conditions [52-54]. It is noteworthy to mention that we analyzed the phenotypes under only standard laboratory conditions; and thus, we cannot exclude the possibility that 1700121C10Rik may be functional in some stimulated conditions.

In conclusion, we identified two novel lncRNAs that are produced from 1700121C10Rik locus. They are predominantly expressed in spermatids and may play diverse functions based on their different subcellular localizations. However, *1700121C10Rik* deficiency was not found to impair male fertility in mice, at least under standard laboratory conditions. Our results indicate that not all lncRNAs which are expressed exclusively in the testes are essential for male fertility in mammals. It is vital to publicize such important data, which would save valuable resources by preventing the same KO recapitulation and help researchers to focus on genes that are indispensable for male reproduction. Furthermore, *1700121C10Rik* locus itself serves as a site for insertion of transgenes for future studies in the field of germ cell development and testis function.

Acknowledgments

This work was supported by National Natural Science Foundation of China (Grant No. 81430028 and No. 81671538), Science and Technology Commission of Shanghai Municipality (Grant No. 16DZ2280800 and No. 18ZR1423500), and Shanghai Municipal Health Commission (Grant No. 201640279 and No. 20174Y0120).

References

- Schultz N, Hamra FK, Garbers DL. A multitude of genes expressed solely in meiotic or postmeiotic spermatogenic cells offers a myriad of contraceptive targets. *Proc Natl Acad Sci USA* 2003; 100: 12201–12206. [Medline] [CrossRef]
- Matzuk MM, Lamb DJ. The biology of infertility: research advances and clinical challenges. Nat Med 2008; 14: 1197–1213. [Medline] [CrossRef]
- Yao C, Liu Y, Sun M, Niu M, Yuan Q, Hai Y, Guo Y, Chen Z, Hou J, Liu Y, He Z. MicroRNAs and DNA methylation as epigenetic regulators of mitosis, meiosis and spermiogenesis. *Reproduction* 2015; 150: R25–R34. [Medline] [CrossRef]
- Ro S, Park C, Song R, Nguyen D, Jin J, Sanders KM, McCarrey JR, Yan W. Cloning and expression profiling of testis-expressed piRNA-like RNAs. *RNA* 2007; 13: 1693–1702. [Medline] [CrossRef]
- Bie B, Wang Y, Li L, Fang H, Liu L, Sun J. Noncoding RNAs: Potential players in the self-renewal of mammalian spermatogonial stem cells. *Mol Reprod Dev* 2018; 85: 720–728. [Medline] [CrossRef]
- Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, Rinn JL. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev* 2011; 25: 1915–1927. [Medline] [CrossRef]
- Wilusz JE, Sunwoo H, Spector DL. Long noncoding RNAs: functional surprises from the RNA world. *Genes Dev* 2009; 23: 1494–1504. [Medline] [CrossRef]
- Fatica A, Bozzoni I. Long non-coding RNAs: new players in cell differentiation and development. Nat Rev Genet 2014; 15: 7–21. [Medline] [CrossRef]
- Costa FF. Non-coding RNAs: new players in eukaryotic biology. *Gene* 2005; 357: 83–94. [Medline] [CrossRef]
- Takahashi H, Carninci P. Widespread genome transcription: new possibilities for RNA therapies. *Biochem Biophys Res Commun* 2014; 452: 294–301. [Medline] [CrossRef]
- Fatima R, Akhade VS, Pal D, Rao SM. Long noncoding RNAs in development and cancer: potential biomarkers and therapeutic targets. *Mol Cell Ther* 2015; 3: 5. [Medline] [CrossRef]
- Luk AC, Chan WY, Rennert OM, Lee TL. Long noncoding RNAs in spermatogenesis: insights from recent high-throughput transcriptome studies. *Reproduction* 2014; 147: R131–R141. [Medline] [CrossRef]
- Hu K, Zhang J, Liang M. LncRNA AK015322 promotes proliferation of spermatogonial stem cell C18-4 by acting as a decoy for microRNA-19b-3p. *In Vitro Cell Dev Biol Anim* 2017; 53: 277–284. [Medline] [CrossRef]
- Li L, Wang M, Wang M, Wu X, Geng L, Xue Y, Wei X, Jia Y, Wu X. A long noncoding RNA interacts with Gfra1 and maintains survival of mouse spermatogonial stem cells. *Cell Death Dis* 2016; 7: e2140. [Medline] [CrossRef]
- Kataruka S, Akhade VS, Kayyar B, Rao MRS. Mrhl long noncoding RNA mediates meiotic commitment of mouse spermatogonial cells by regulating Sox8 expression. *Mol Cell Biol* 2017; 37: 37. [Medline] [CrossRef]
- Anderson DM, Anderson KM, Chang CL, Makarewich CA, Nelson BR, McAnally JR, Kasaragod P, Shelton JM, Liou J, Bassel-Duby R, Olson EN. A micropeptide encoded by a putative long noncoding RNA regulates muscle performance. *Cell* 2015;

160: 595-606. [Medline] [CrossRef]

- Huang JZ, Chen M, Chen D, Gao XC, Zhu S, Huang H, Hu M, Zhu H, Yan GR. A Peptide Encoded by a Putative lncRNA HOXB-AS3 Suppresses Colon Cancer Growth. *Mol Cell* 2017; 68: 171–184.e6. [Medline] [CrossRef]
- Nelson BR, Makarewich CA, Anderson DM, Winders BR, Troupes CD, Wu F, Reese AL, McAnally JR, Chen X, Kavalali ET, Cannon SC, Houser SR, Bassel-Duby R, Olson EN. A peptide encoded by a transcript annotated as long noncoding RNA enhances SERCA activity in muscle. *Science* 2016; 351: 271–275. [Medline] [CrossRef]
- Yeasmin F, Yada T, Akimitsu N. Micropeptides encoded in transcripts previously identified as long noncoding RNAs: A new chapter in transcriptomics and proteomics. *Front Genet* 2018; 9: 144. [Medline] [CrossRef]
- Makarewich CA, Olson EN. Mining for micropeptides. Trends Cell Biol 2017; 27: 685–696. [Medline] [CrossRef]
- Shen C, Kuang Y, Liu J, Feng J, Chen X, Wu W, Chi J, Tang L, Wang Y, Fei J, Wang Z. Prss37 is required for male fertility in the mouse. *Biol Reprod* 2013; 88: 123. [Medline] [CrossRef]
- Shang X, Shen C, Liu J, Tang L, Zhang H, Wang Y, Wu W, Chi J, Zhuang H, Fei J, Wang Z. Serine protease PRSS55 is crucial for male mouse fertility via affecting sperm migration and sperm-egg binding. *Cell Mol Life Sci* 2018; 75: 4371–4384. [Medline] [CrossRef]
- Kong L, Zhang Y, Ye ZQ, Liu XQ, Zhao SQ, Wei L, Gao G. CPC: assess the proteincoding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Res* 2007; 35: W345-9. [Medline] [CrossRef]
- Wang L, Park HJ, Dasari S, Wang S, Kocher JP, Li W. CPAT: Coding-Potential Assessment Tool using an alignment-free logistic regression model. *Nucleic Acids Res* 2013; 41: e74. [Medline] [CrossRef]
- Braissant O, Wahli W. A simplified in situ hybridization protocol using non-radioactively labeled probes to detect abundant and rare mRNAs on tissue sections. *Biochemica (Indianap, Ind)* 1998; 1: 1–8.
- Rodríguez-Casuriaga R, Geisinger A, López-Carro B, Porro V, Wettstein R, Folle GA. Ultra-fast and optimized method for the preparation of rodent testicular cells for flow cytometric analysis. *Biol Proced Online* 2009; 11: 184–195. [Medline] [CrossRef]
- Kurihara M, Shiraishi A, Satake H, Kimura AP. A conserved noncoding sequence can function as a spermatocyte-specific enhancer and a bidirectional promoter for a ubiquitously expressed gene and a testis-specific long noncoding RNA. J Mol Biol 2014; 426: 3069–3093. [Medline] [CrossRef]
- Yoneda R, Satoh Y, Yoshida I, Kawamura S, Kotani T, Kimura AP. A genomic region transcribed into a long noncoding RNA interacts with the Prss42/Tessp-2 promoter in spermatocytes during mouse spermatogenesis, and its flanking sequences can function as enhancers. *Mol Reprod Dev* 2016; 83: 541–557. [Medline] [CrossRef]
- Zindy F, den Besten W, Chen B, Rehg JE, Latres E, Barbacid M, Pollard JW, Sherr CJ, Cohen PE, Roussel MF. Control of spermatogenesis in mice by the cyclin D-dependent kinase inhibitors p18^{Ink4c} and p19^{Ink4d}. *Mol Cell Biol* 2001; 21: 3244–3255. [Medline] [CrossRef]
- Carlevaro-Fita J, Johnson R. Global positioning system: understanding long noncoding RNAs through subcellular localization. *Mol Cell* 2019; 73: 869–883. [Medline] [Cross-Ref]
- Nishimura H, L'Hernault SW. Spermatogenesis. Curr Biol 2017; 27: R988–R994. [Medline] [CrossRef]
- 32. Fawcett DW. The mammalian spermatozoon. Dev Biol 1975; 44: 394–436. [Medline] [CrossRef]
- Yan W. Male infertility caused by spermiogenic defects: lessons from gene knockouts. Mol Cell Endocrinol 2009; 306: 24–32. [Medline] [CrossRef]
- Kang-Decker N, Mantchev GT, Juneja SC, McNiven MA, van Deursen JM. Lack of acrosome formation in Hrb-deficient mice. *Science* 2001; 294: 1531–1533. [Medline] [CrossRef]
- Xiong W, Wang Z, Shen C. An update of the regulatory factors of sperm migration from the uterus into the oviduct by genetically manipulated mice. *Mol Reprod Dev* 2019; 86: 935–955. [Medline] [CrossRef]
- Kotaja N, Kimmins S, Brancorsini S, Hentsch D, Vonesch JL, Davidson I, Parvinen M, Sassone-Corsi P. Preparation, isolation and characterization of stage-specific sper-

matogenic cells for cellular and molecular analysis. Nat Methods 2004; 1: 249-254. [Medline] [CrossRef]

- Ponting CP, Oliver PL, Reik W. Evolution and functions of long noncoding RNAs. Cell 2009; 136: 629–641. [Medline] [CrossRef]
- Ulitsky I, Bartel DP. lincRNAs: genomics, evolution, and mechanisms. *Cell* 2013; 154: 26–46. [Medline] [CrossRef]
- Yao RW, Wang Y, Chen LL. Cellular functions of long noncoding RNAs. Nat Cell Biol 2019; 21: 542–551. [Medline] [CrossRef]
- Batista PJ, Chang HY. Long noncoding RNAs: cellular address codes in development and disease. *Cell* 2013; 152: 1298–1307. [Medline] [CrossRef]
- Kopp F, Mendell JT. Functional classification and experimental dissection of long noncoding RNAs. *Cell* 2018; 172: 393–407. [Medline] [CrossRef]
- Zhang B, Arun G, Mao YS, Lazar Z, Hung G, Bhattacharjee G, Xiao X, Booth CJ, Wu J, Zhang C, Spector DL. The lncRNA Malat1 is dispensable for mouse development but its transcription plays a cis-regulatory role in the adult. *Cell Reports* 2012; 2: 111–123. [Medline] [CrossRef]
- Nakagawa S, Naganuma T, Shioi G, Hirose T. Paraspeckles are subpopulation-specific nuclear bodies that are not essential in mice. J Cell Biol 2011; 193: 31–39. [Medline] [CrossRef]
- Zhou J, Leu NA, Eckardt S, McLaughlin KJ, Wang PJ. STK31/TDRD8, a germ cell-specific factor, is dispensable for reproduction in mice. *PLoS One* 2014; 9: e89471. [Medline] [CrossRef]
- 45. Miyata H, Castaneda JM, Fujihara Y, Yu Z, Archambeault DR, Isotani A, Kiyozumi D, Kriseman ML, Mashiko D, Matsumura T, Matzuk RM, Mori M, Noda T, Oji A, Okabe M, Prunskaite-Hyyrylainen R, Ramirez-Solis R, Satouh Y, Zhang Q, Ikawa M, Matzuk MM. Genome engineering uncovers 54 evolutionarily conserved and testisenriched genes that are not required for male fertility in mice. *Proc Natl Acad Sci USA* 2016; **113**: 7704–7710. [Medline] [CrossRef]
- Nalam RL, Lin YN, Matzuk MM. Testicular cell adhesion molecule 1 (TCAM1) is not essential for fertility. *Mol Cell Endocrinol* 2010; 315: 246–253. [Medline] [CrossRef]
- Yuan S, Swiggin HM, Zheng H, Yan W. A testis-specific gene, Ubqlnl, is dispensable for mouse embryonic development and spermatogenesis. *Mol Reprod Dev* 2015; 82: 408–409. [Medline] [CrossRef]
- Yuan S, Qin W, Riordan CR, McSwiggin H, Zheng H, Yan W. Ubqln3, a testis-specific gene, is dispensable for embryonic development and spermatogenesis in mice. *Mol Re*prod Dev 2015; 82: 266–267. [Medline] [CrossRef]
- 49. Lu Y, Oura S, Matsumura T, Oji A, Sakurai N, Fujihara Y, Shimada K, Miyata H, Tobita T, Noda T, Castaneda JM, Kiyozumi D, Zhang Q, Larasati T, Young SAM, Kodani M, Huddleston CA, Robertson MJ, Coarfa C, Isotani A, Aitken RJ, Okabe M, Matzuk MM, Garcia TX, Ikawa M. CRISPR/Cas9-mediated genome editing reveals 30 testis-enriched genes dispensable for male fertility in mice. *Biol Reprod* 2019; 101: 501–511. [CrossRef]
- Wen K, Yang L, Xiong T, Di C, Ma D, Wu M, Xue Z, Zhang X, Long L, Zhang W, Zhang J, Bi X, Dai J, Zhang Q, Lu ZJ, Gao G. Critical roles of long noncoding RNAs in Drosophila spermatogenesis. *Genome Res* 2016; 26: 1233–1244. [Medline] [CrossRef]
- Wichman L, Somasundaram S, Breindel C, Valerio DM, McCarrey JR, Hodges CA, Khalil AM. Dynamic expression of long noncoding RNAs reveals their potential roles in spermatogenesis and fertility. *Biol Reprod* 2017; 97: 313–323. [Medline] [CrossRef]
- Carrieri C, Cimatti L, Biagioli M, Beugnet A, Zucchelli S, Fedele S, Pesce E, Ferrer I, Collavin L, Santoro C, Forrest AR, Carninci P, Biffo S, Stupka E, Gustincich S. Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat. *Nature* 2012; 491: 454–457. [Medline] [CrossRef]
- 53. Adriaens C, Standaert L, Barra J, Latil M, Verfaillie A, Kalev P, Boeckx B, Wijnhoven PW, Radaelli E, Vermi W, Leucci E, Lapouge G, Beck B, van den Oord J, Nakagawa S, Hirose T, Sablina AA, Lambrechts D, Aerts S, Blanpain C, Marine JC. p53 induces formation of NEAT1 lncRNA-containing paraspeckles that modulate replication stress response and chemosensitivity. *Nat Med* 2016; 22: 861–868. [Medline] [CrossRef]
- Zeng Q, Wang Q, Chen X, Xia K, Tang J, Zhou X, Cheng Y, Chen Y, Huang L, Xiang H, Cao K, Zhou J. Analysis of lncRNAs expression in UVB-induced stress responses of melanocytes. J Dermatol Sci 2016; 81: 53–60. [Medline] [CrossRef]