Topaz1, an essential gene for murine spermatogenesis, down-regulates the expression of many testis-specific long non-coding RNAs.

Short title: Topaz1, long non-coding RNAs and spermatogenesis

Manon Chadourne[1], Elodie Poumerol[1], Luc Jouneau[1], Bruno Passet[2], Johan Castille[2], Eli Sellem[3], Eric Pailhoux[1] and Béatrice Mandon-Pépin[1]*

[1] Université Paris-Saclay, UVSQ, INRAE, BREED, 78350, Jouy-en-Josas, France; Ecole Nationale Vétérinaire d'Alfort, BREED, 94700, Maisons-Alfort, France.

[2] Université Paris Saclay, INRAE, AgroParisTech, GABI, Jouy-en-Josas, France.

[3] R&D Department, ALLICE, Paris, France.

* Corresponding author

E-mail: beatrice.mandon-pepin@inrae.fr (BMP)

Abstract

Spermatogenesis comprises a coordinated process, including meiosis, to produce fertile male gametes. Previous study reported that *Topaz1* is a germ cell specific gene highly conserved in vertebrates. *Topaz1* knockout male mice are sterile. The mutant testes lack haploid germ cells and meiosis arrests at the first-division prophase-metaphase transition. Here, in order to better characterize the testicular phenotype of *Topaz1*$^{-/-}$ mice, we used RNA-seq analyses at two different developmental stages. At postnatal days 16 (P16), 205 genes were differentially expressed genes (DEGs) in *Topaz1*$^{-/-}$ testes. They suggest stress conditions in mutant testes. At P18, the number of DEGs was increased 10-fold and 90% were down-regulated. The absence of *Topaz1* seems to disturb the expression of genes involved in microtubule and/or cilium mobility, spermatogenesis and first meiotic division during the transition from prophase to metaphase. This is consistent with the *Topaz1*$^{-/-}$ testis phenotype where microtubule networks and centrosomes are disrupted. Moreover, a quarter of P18-DEGs are long non-coding RNAs (lncRNAs). Three of them, down-regulated at P16 and P18, were studied. They are testis-specific, located in spermatocytes and their expression starts between P11 and P15. We report here the effects of the suppression of one of these lncRNAs, *4939463O16Rik*. The mouse fertility is not affected although the sperm parameters are disturbed. Transcriptome of P18-*4939463O16Rik*$^{-/-}$ testes is altered and the affected molecular pathways include microtubule-based process, regulation of cilium movement, spermatogenesis, male gamete differentiation. The absence of TOPAZ1 protein or of *4930463O16Rik* lncRNA showed the same enrichment clusters in mutant testes despite a contrasted phenotype on the male fertility.

In conclusion, *Topaz1* is an essential gene for male fertility in mice and seems to stabilize the expression of many lncRNAs. Its absence leads to meiotic arrest. Suppression of one lncRNA is dispensable for mouse fertility but is necessary during terminal differentiation of male gametes.

Author Summary

The *Topaz1* gene was initially characterized during the meiotic initiation in the sheep fetal ovary. In order to determine its function, the KO of the murine gene was performed. In this species, only males are sterile and spermatogenesis is blocked before the first meiotic division. Here, we show that cytoskeletal elements are strongly disturbed in mutant testes, a sign that these elements take an important function in spermatogenesis. While the mitotic spindle of spermatogonia is normal, the meiotic spindle of spermatocytes is a hemispindle-shaped and the homologous chromosome pairs cannot position themselves on the equatorial plate. In addition, lncRNAs represent 25% of genes whose testis expression varies significantly due to the absence of *Topaz1*. This suggests a key role of these factors in spermatogenesis. Largely testis specific, they could be involved in spermatogenesis with a more or less critical role on the mouse fertility, probably also depending of their redundancies.

Introduction

In mammals, an organism derived from two parental haploid gametes, a maternal oocyte and a paternal sperm. Meiosis is a highly specialized event that leads to the production of these haploid germ cells [1]. In female, meiosis is initiated during fetal life while male germ cells are involved in the meiosis process around puberty. In male, it takes place during the process of spermatogenesis that involves mitotic division and multiplication of spermatogonia, segregation of homologous chromosomes via meiosis and spermiogenesis of haploid germ cells. This complex process of spermatogenesis, which progresses through precisely timed and highly organized cycles, is primordial for male fertility. All these different events are highly regulated and associated with controlled expression of several testis-enriched genes. A previous study has shown the essential role of *Topaz1* during meiosis steps in male mice [2]. *Topaz1* is a highly conserved gene in vertebrates. Its expression is germ cell-specific, as demonstrated in mice and sheep [3]. Suppression of *Topaz1* in mice (*Topaz1*$^{-/-}$) results in males azoospermia. Male meiotic blockage occurs without deregulation of chromosome alignment and

TOPAZ1 is not involved in the formation of the XY body or the maintenance of MSCI. *Topaz1* depletion increases apoptosis of male pachytene cells. A chromosome misalignment at the metaphasic I plate is observed in the absence of *Topaz1* in mouse testes. This misalignment leads to an arrest at the prophase to metaphase transition during the first meiosis division [2]. Microarray-based gene expression profiling of *Topaz1*$^{-/-}$ testis revealed that TOPAZ1 influences the expression of one hundred transcripts including several long non-coding RNA (lncRNAs) and unknown genes at postnatal day 20 (P20) [2].

Since the discovery of the maternal *H19* lncRNA gene [4] and of the *Xist* gene [5] that regulate the structure of chromosomes and mediating gene repression during X chromosome inactivation, the interest in studying the role of non-coding RNAs (ncRNAs) greatly increased. Non-coding RNAs are present in many organisms, from bacteria to humans, where only 1.2% of the human genome codes for functional proteins [6–8]. They are divided into two groups according to their length. The small non-coding RNAs (sncRNAs) group contains transcripts smaller than 200 nucleotides (nt). They refer to microRNAs (miRNAs, 20-25 nt), small interfering RNAs (siRNAs), PIWI-interacting RNAs (piRNAs, 26-31 nt) and circular RNAs (cricRNA) and are essential for several functions such as the regulation of gene expression and genome protection (Ref in [9]). The fundamental role of sncRNAs in spermatogenesis is well described in recent years [10–12]. The second group, the lncRNAs, contains transcripts longer than 200 nt without significant open reading frame. Advances in high-throughput sequencing allowed the identification of new transcripts, including lncRNAs. Most of them are transcribed by the RNA polymerase II and possess a 5' cap and polyadenylated tail (ref in [13]). They are classified according to their length, location in the genome, (surrounding regulatory elements for example) or functions.

While much remains to be discovered about the functions of ncRNAs and their molecular interactions, accumulative evidences suggest that ncRNAs participate in various biological processes such as cell differentiation, development, proliferation, apoptosis and cancers.

Several studies pointed out that testes have a very high proportion of lncRNA compared to other organs [14,15]. However, this high testicular expression is only observed in the adult organ, as the level

of lncRNAs in the developing testis is comparable to somatic organs [15]. In *Drosophila*, Wen *et al*. produced mutant fly lines by deleting 105 testis-specific lncRNAs and demonstrated the essential role of 33 of them for spermatogenesis and/or male fertility [16]. In mice, some testis-expressed lncRNAs were functionally characterized during spermatogenesis. For examples, *Mrhl* RNA repressed *Wnt* signaling in the Gc1-Spg spermatogonial cell line suggesting a role in spermatocyte differentiation [17]. The expression of the Testis-specific X-linked gene was specific of and highly abundant in pachytene-stage spermatocytes and could regulate germ cells progression in meiosis [18]. Lastly, it has been shown that the *Dmrt1*-related gene negatively regulates *Dmrt1* (doublesex and mab-3 related transcription factor 1) and that this regulation could be involved in the switching between mitosis and meiosis in male germ cells [19].

Following a previous study, which presented microarray comparative analyses of wild-type and *Topaz1*$^{-/-}$ testis RNAs at P15 and P20 [2], we performed a deep sequencing by bulk RNA-seq of these testes collected at P16 and P18 in order to refine the developmental stages showing transcriptional differences between the two mouse lines. Since the proportion of deregulated lncRNAs represented about a quarter of the differentially expressed genes (DEGs), we studied the testicular localization of three of them and created a mouse line deleted of one of them (*4930463O16Rik*). These knockout mice showed normal fertility in both sexes, but male mutants produced half as much sperm than wild-type controls.

Results

Topaz1 mutant testes have a deregulated transcriptome as early as P16.

To expand the previous microarray comparative analyses of wild-type and mutant testis RNA performed at P15 and P20, during the first wave of spermatogenesis [2], transcriptomic analyses by RNA-seq were performed on WT and *Topaz1*$^{-/-}$ mouse testes at two developmental stages: P16 and P18. The P16 stage was chosen because these previous microarray analyses revealed that only the

126 Topaz1 gene was expressed differently at P15, its expression starting from 5 d*pp*. This means that the
127 time when TOPAZ1 should have an effect is just after P15. Furthermore, whereas at P15, seminiferous
128 tubules contain spermatocytes that have advanced to mid and late-pachytene, at P16, they contain
129 spermatocyte cells that progressed from the end-pachytene to early diplotene of meiosis I. At P20, the
130 first round spermatids appear, whereas at P18, late-pachytene spermatocytes are abundant and the
131 very first spermatocytes II appear [20]. Thus, the two P16 and P18 stages chosen in this study
132 surrounded as close as possible the time lapse just before and after the first meiosis I division of
133 spermatogenesis.
134
135 Differential analyses of RNA-seq results showed that 205 and 2748 genes were significantly
136 deregulated in *Topaz1*$^{-/-}$ testis compared to WT at P16 and P18 respectively (adjust p-value < 0.05 and
137 absolute Log2 Fold Change > 1 (Log2FC>1) (Fig 1A, S1 Table). At P16, out of the 205 DEGs, 97 genes
138 were significantly down-regulated (Log2FC<-1 or FC<0.5) and 108 up-regulated (Log2FC>1 or FC>2).
139 However, at P18 down-regulated DEGs accounted for 91% (2491 genes) and up-regulated genes for
140 only 9% (257genes). Among all these DEGs, 120 were in common between P16 and P18 (Fig 1A). The
141 2748 DEGs of the developmental stage P18 were largely testis-enriched DEGs in mouse testis specific
142 genes according to the mouse gene atlas (Fig 1B).

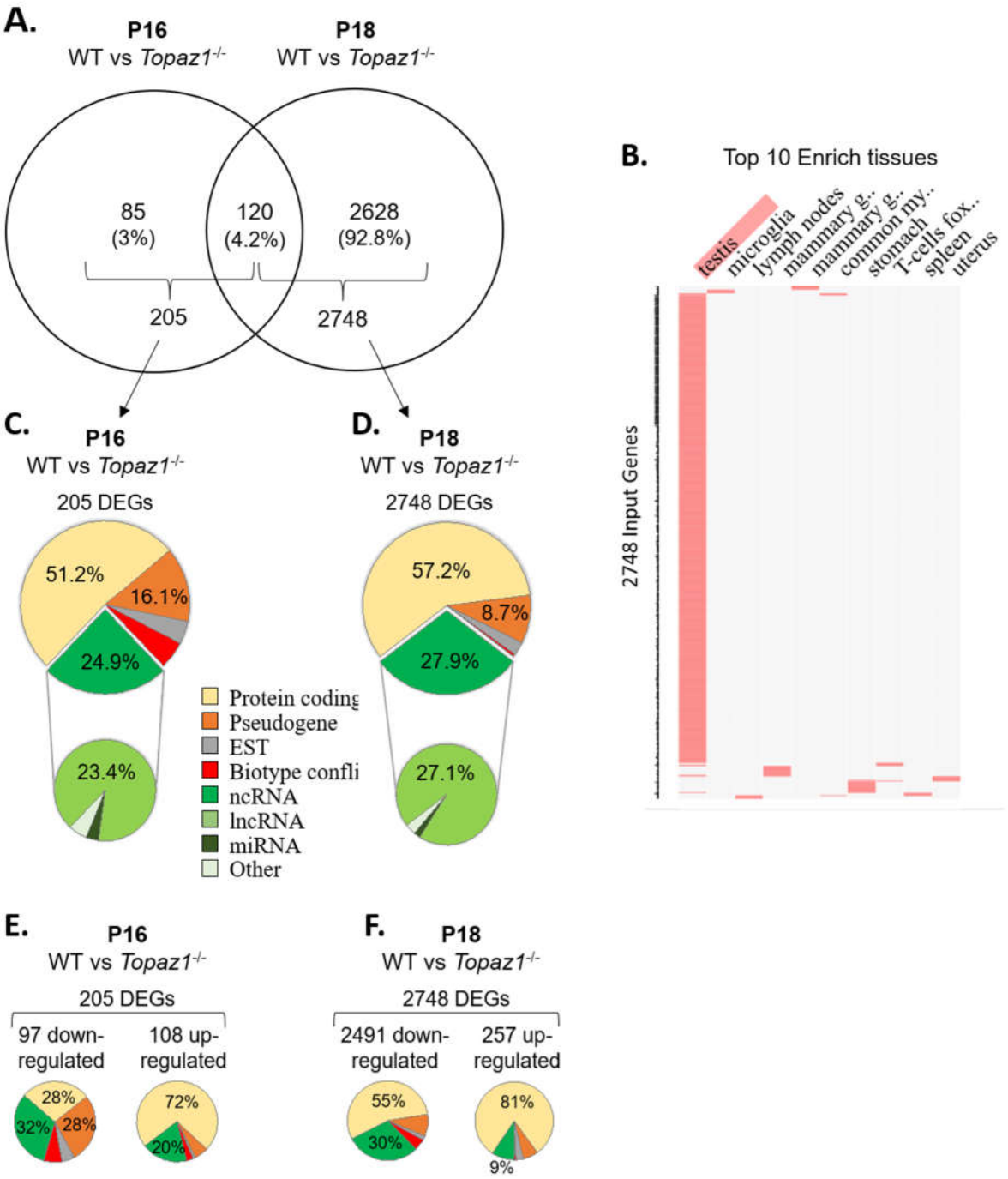


Fig 1: WT vs *Topaz1*$^{-/-}$ deregulated gene analysis from mouse testes.

(A) Venn diagram showing overlap of differentially expressed genes between P16 and P18 *Topaz1*$^{-/-}$ mouse testes (adj p<0.05 and down-regulated FC<0.5 (log2FC<-1) or up-regulated FC>2 (log2FC>1)). (B) Clustergrammer was generated by the Enrichr website. Top 10 enrich tissues are the columns, input genes (2748 DEGs of P18 *Topaz1*$^{-/-}$ compared to normal testes) are the rows, and cells in the matrix indicate if a DEG is associated with a tissue from the Mouse Gene Atlas. (C-D) Biotype of DEGs in *Topaz1*$^{-/-}$ testis of (C) P16 and (D) P18. Around half of them

are protein-coding genes whereas around one quarter is ncRNA at both developmental stages. (E-F) Biotype of DEGs in *Topaz1*$^{-/-}$ testis of (E) P16 and (F) P18 depending on whether they were up- or down-regulated.

Validation of several DEGs was realized by RT-qPCR. Two randomly selected up-regulated genes at P16 (B3galt2 and Hp) and three at P18 (*B3galt2*, *Afm* and *Cx3cr1*), four down-regulated genes at P16 and P18 (*Gstt2*, *4930463O16Rik*, *4921513H07Rik* and *Gm21269*) and two non-differential genes (*Cdc25c* and *Nop10*) were analyzed (S1 Fig). Results confirmed those obtained by RNA-seq.

The biotype of the differential transcripts (protein-coding, non-coding RNAs…) was performed on the annotation of the NCBI, MGI and Ensembl databases. Two major deregulated groups were highlighted at both stages. The protein-coding gene biotype accounted for half of the deregulated genes (51.2% and 57.2% at P16 and P18 respectively) (Fig 1C-D). A quarter of *Topaz1*$^{-/-}$ DEGs, 24.9% and 27.9% at P16 and P18 respectively, was found to belong to the ncRNA second group. Among this latter one, the major biotype was the lncRNAs one at both stages, 23.4% and 27.1% at P16 and P18 respectively. This significant proportion of deregulated lncRNA raises the question of their potential involvement in spermatogenesis.

Pathway and functional analysis of DEGs

To further understand biological functions and pathways, these DEGs were functionally annotated based on GO terms and KEGG pathway or on InterPro databases through the Database for Annotation, Visualization and Integrated Discovery ontology database (DAVID v6.8, https://david.ncifcrf.gov/) with the default criteria [21,22].

At P16, thus before the first meiosis division, out of 205 differentially expressed genes, 32% of down- and 20% of up- regulated genes corresponded to non-coding RNAs with no GO annotation or no pathway affiliation for the vast majority (Fig 1E), leading a less powerful functional annotation clustering (S2 Table). Five clusters with an enrichment score > 1.3 were obtained (enrichment score > 1.3 was used for a cluster to be statistically significant as recommended by Huang et al., [21] but the

number of genes in each cluster was low except for the Annotation cluster number 4. In this one, TOPAZ1 absence seems to affect the extracellular compartment. The others referred to the antioxidant molecular function and the detoxification biological process, suggesting stress conditions.

At P18, corresponding to the first transitions from prophase to metaphase, either considering all DEGs (2748 DEGs; 2404 DAVID IDs) or only the down-regulated genes (2491 DEGs; 2164 DAVID IDs) in the P18 *Topaz1*$^{-/-}$ versus WT testes, resulted in the identification of five identical clusters with an enrichment score higher than 12 (Fig 1F, S3 Table). However, enrichment scores were higher when only the down-regulated genes were considered. These clusters include the following GO terms (*i*) for cellular components: motile cilium, ciliary part, sperm flagellum, axoneme, acrosomal vesicule; (*ii*) for biological processes: microtubule-based process, spermatogenesis, germ cell development, spermatid differentiation (S3 Table).

Finally, using the InterPro database, four clusters with enrichment score > 1.3 were obtained based on down-regulated genes (S3 Table) and with-up-expressed genes, the absence of TOPAZ1 in mouse testes highlighted the biological pathway of the response to external stimulus or of the defense response in the testis, again suggesting, as for P16, stressful conditions in these *Topaz1*$^{-/-}$ testes.

These results indicated that the absence of TOPAZ1 induced alterations of the murine transcriptome of mutant testis transcriptome as early as 16 days after birth. Two days later (P18), these effects were amplified and with a predominantly down-regulation of genes (91% of DEGs). The loss of TOPAZ1 appeared to disrupt the regulation of genes involved in microtubule and/or cilium mobility, spermatogenesis and first meiotic division during the prophase to metaphase transition. This is in agreement with the *Topaz1*$^{-/-}$ phenotype in testes.

Absence of TOPAZ1 leads to drastic cytoplasmic defects before the first meiotic division

According to the preceding GO pathway analyses showing that a majority of deregulated proteins are involved in microtubule cytoskeleton organization, microtubule-based movements and processes, microtubule organizing centers, centrosomes and centrioles in P18 Topaz1$^{-/-}$ testes (S3 Table), we

aimed to better characterize the cytoplasmic components of germ cells in Topaz1$^{-/-}$ testes before the first meiotic division. As the meiotic spindle is a key component of these cells before and during the metaphase stage, we studied it by α- and γ-tubulin immunofluorescence (IF) staining, markers of microtubule spindle and centrosome respectively (Fig 2). We observed a hemispindle centered in the germ cells of the Topaz$^{-/-}$ testis. Moreover, in these cells, centrosome staining was diffuse and weak. This was also observed on entire seminiferous sections (S2 Fig). The chromosomes were not aligned along a metaphase plate but adopted an atypical rosette shape (Fig 2), reflecting a strong perturbation of the microtubule and centrosome pathways in Topaz1-deficient spermatocytes that leads to meiotic arrest.

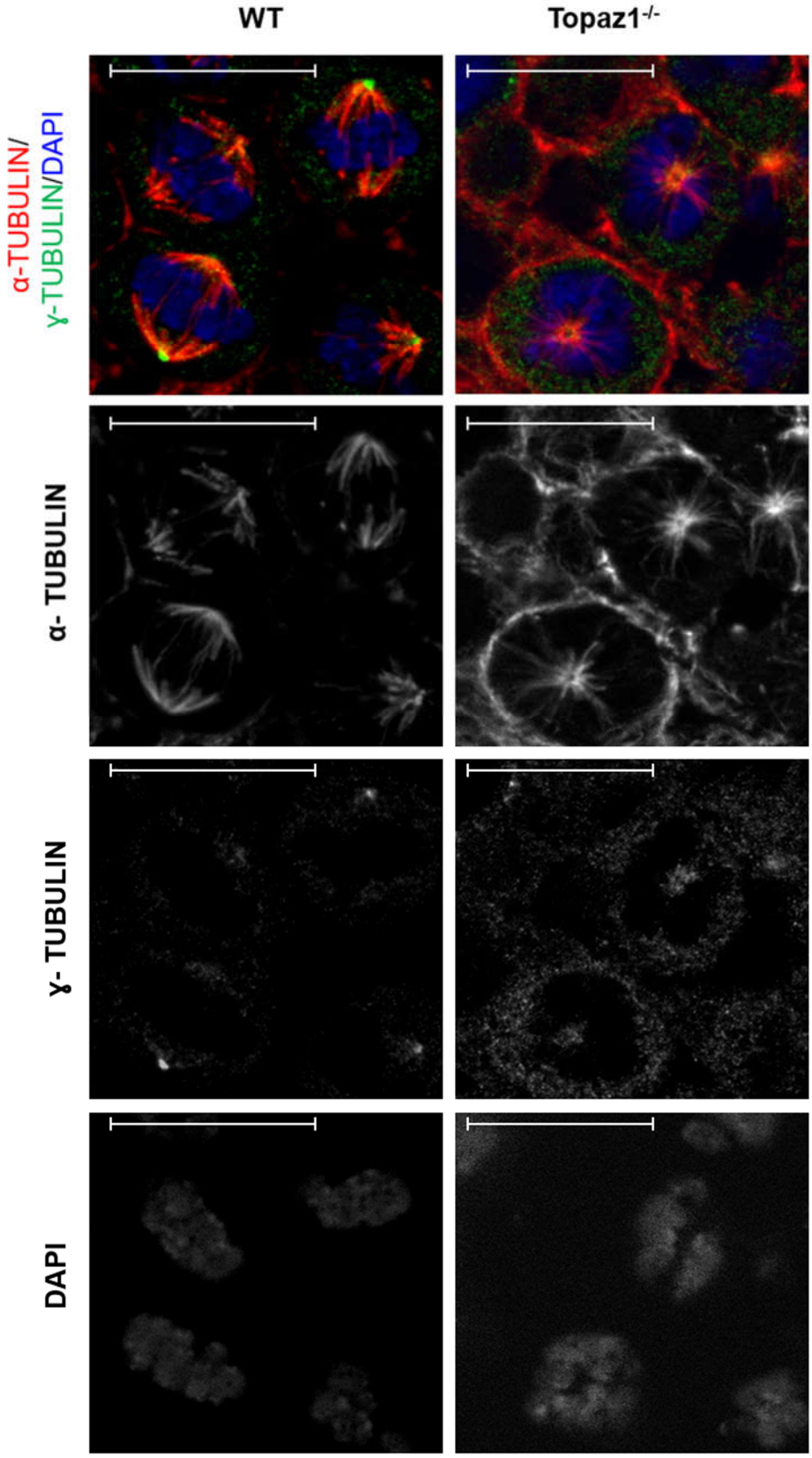


Fig 2: Abnormal metaphase phenotype in Topaz1-deficient gonads.

Immunofluorescence staining for α-TUBULIN (red), γ-TUBULIN (green) and DAPI (blue) in WT (left) and Topaz1$^{-/-}$ (right) 28 d*pp* testis sections. Unlike meiotic metaphases seen in normal testes (left), metaphases are abnormal in Topaz1$^{-/-}$ mutants (right) with atypical rosette shape and hemispindle. Scale bar = 20µm

Selection of 3 deregulated lncRNA: all spermatocyte-specific

The vast majority of deregulated lncRNAs in *Topaz1*$^{-/-}$ testes has an unknown function. We decided to study 3 of the 35 down-regulated lncRNAs that are shared by P16 and P18 stages, namely *4930463O16Rik* (ENSMUSG00000020033), *4921513H07Rik* (ENSMUSG00000107042) that is the most down-regulated gene at P16 with a Log2FC of 11.85, both already highlighted in the previous microarray comparative analyses [2] and *Gm21269* (ENSMUSG00000108448), that has the lowest adjust p-value at P18. We quantified these transcripts by qPCR in several somatic tissues (brain, heart, liver, lung, small intestine, muscle, spleen, kidney, epididymis and placenta) and in gonads (testis and ovary). These three lncRNAs were almost exclusively expressed in testis (Fig 3A, C, E). These results were in agreement with RNA-seq data available for *4930463O16Rik* and *Gm21269* on the ReproGenomics viewer (https://rgv.genouest.org/) (S3A and S4A Figs respectively) [23,24]. Our RNA-seq results, summarized using our read density data (bigwig) and the Integrative Genomics Viewer (IGV; http://software.broadinstitute.org/software/igv/), showed a weak or absence of expression of these three genes in *Topaz1*$^{-/-}$ testes (S5A-C Fig).

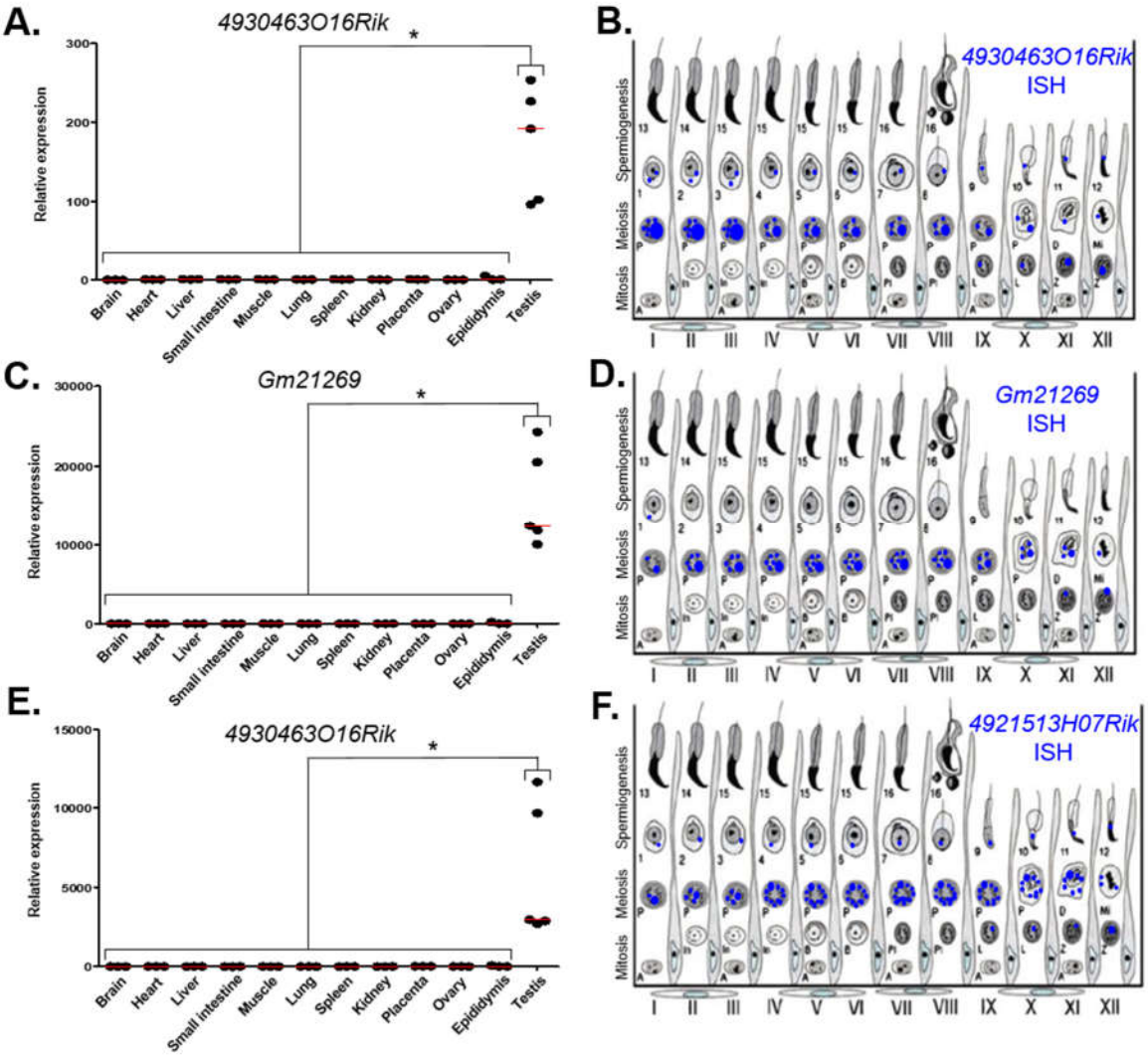


Fig 3: Expression analysis of three lncRNAs.

(A-C-E) RT-qPCR analysis of three different lncRNAs. (A) *4930463O16Rik*; (C) *Gm21269*; (E) *4921513H07Rik* in different two month-old tissues of WT mice. The red lines represent the median for each tissues. n=5 for testes and n=3 for other organs. Statistical analyses were realized with non-parametric Kruskal-Wallis test. * = $p<0.05$ (B-D-F) Schematic representation of the result of (B) *4930463O16Rik*, (D) *Gm21269* and (F) *49215113H07Rik ISH* expression in the meiotic and post-meiotic cells of the WT mouse seminiferous epithelial cycle.

Quantification of these transcripts by qPCR from postnatal to adulthood in WT and *Topaz1*$^{-/-}$ testes was previously reported, as for *4930463O16Rik* and *4921513H07Rik* (Fig 9 in [2]) or realized for *Gm21269* (S6 Fig, also including the postnatal expression of *4930463O16Rik* and *4921513H07Rik*). The

242 difference of expression between normal and *Topaz1*$^{-/-}$ testes was detected significantly as early as
243 P15 (detected insignificant in the previous microarray analysis and *Gm21269* was absent on the used
244 microarray). All showed an absence of expression or at least an important down-regulation in the
245 mutant testis.
246
247 To determine testicular localization of these lncRNA, *in situ* hybridization (*IS*H) on adult WT testis
248 sections was performed (S7 Fig) and results summarized (Fig 3B, D, F). These lncRNAs were expressed
249 in spermatocytes and the most intense probe labeling was observed in the pachytene stage. These
250 results were confirmed by data on the ReproGenomics viewer for *4930463O16Rik* and *Gm21269*
251 (https://rgv.genouest.org/) (S3B and S4B Figs) [23,24].
252 To refine the subcellular localization of these transcripts in the adult mouse testis, we have paired *IS*H
253 experiments and IF staining of the DDX4 protein (or Mvh, Mouse Vasa homolog). DDX4 is a germ cell
254 cytoplasmic marker of germ cells, especially in the testis [25]. Our results showed that the three
255 observed lncRNAs had different expression intensities depending on seminiferous epithelium stages.
256 *4930463O16Rik* was expressed in the nucleus of spermatocytes with a diffuse fluorescence,
257 surrounded by the cytoplasmic DDX4 labelling from the zygotene to the diplotene stages (Fig 4A-B-C).
258 In the same spermatocyte stages (zygotene to diplotene), diffuse labelling of *Gm21269*, similar to that
259 of *4930463O16Rik*, was observed but with the addition of a dot-shaped labelling that co-localized with
260 DDX4 fluorescence (Fig 4D, E, F). *Gm21269* was therefore localized in the cytoplasm and nuclei of
261 spermatocytes during meiosis. *4921513H07Rik* seemed to be cytoplasmic with fluorescent red dots
262 (*IS*H) surrounding the nuclei, and located in close proximity of DDX4 (IF) labelling (Fig 4G, H, I). In other
263 stages, identified by DDX4 staining, *IS*H labelling of these three lncRNA resulted in the observation of
264 single dots in a few spermatogonia and in round spermatids. We performed the same experiment: *IS*H
265 was followed by an IF staining of γH2Ax to highlight the sex body in spermatocytes (S8 Fig). No co-
266 localization between the sex body and the three lncRNA was revealed.

Altogether, these results suggested that these spermatocyte-specific lncRNAs had different subcellular localization into spermatocytes, suggesting functions in these male germ cells.

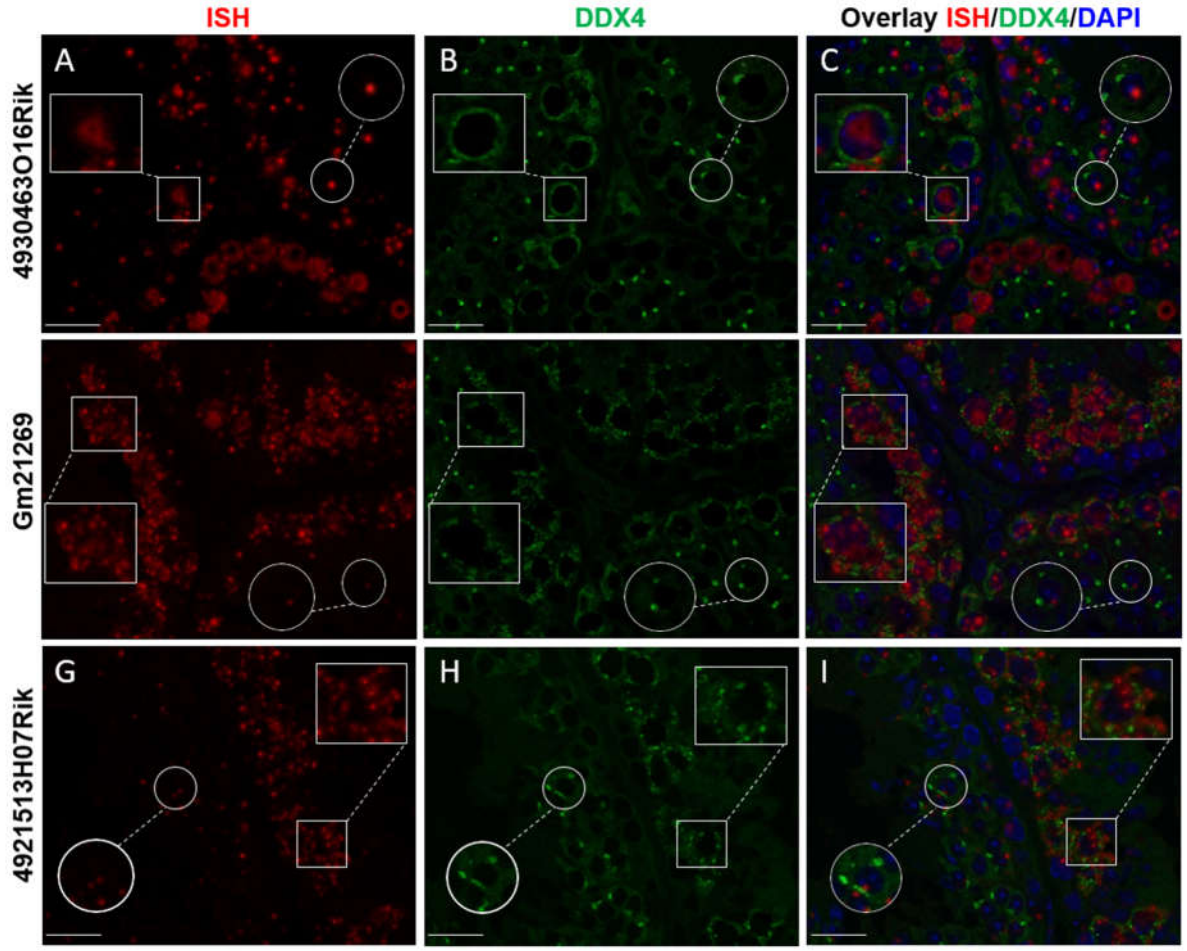


Fig 4: lncRNA cellular localizations on WT two month-old mouse testes.

In situ hybridization using (A) *4930463O16Rik*, (D) *Gm21269* and (G) *4921513H07Rik* probes (red). (B-E-H) Immunofluorescence staining with DDX4 antibody was achieved in the same stage of seminiferous epithelium to identify male germ cells (green). (C-F-I) DAPI (blue), visualizing nuclear chromosomes, was merge with *IS*H (green) and IF (red) signals. Zooms in white squares showed spermatocytes during first meiotic division (zygotene to diplotene stages). Zooms in circles showed spermatid cells with one spot of DDX4 staining per cell. Scale bar = 20 µm.

Generation of *4930463O16Rik*-deleted mice

In order to evaluate a potential role in spermatogenesis of one of these lncRNAs, *4930463O16Rik*, the nuclear expressed gene, was chosen to suppress its expression in a mouse knockout model. *4930463O16Rik* gene (Chr10: 84,488,293-84,497,435 - GRCm38:CM001003.2) was described in public databases as consisting of 4 exons spanning approximately 10 kb in an intergenic locus on the mouse chromosome 10. By PCR and sequencing, we confirmed this arrangement (data not shown). In the goal of understand the role of *4930463O16Rik*, a new mouse line deleted of this lncRNA was created by CRISPR/Cas9 technology (Fig 5A). Briefly, multiple single guide RNAs (sgRNAs) were chosen, 2 sgRNAs in 5' of exon 1 and 2 sgRNAs in 3' of exon 4, to target the entire length of this gene (Fig 5A, C) and increase gene deletion efficiency in mouse [26]. Mice with disruption of the target site were identified after Sanger sequencing of PCR amplification of the genomic region surrounding the deleted locus (Fig 5D). *4930463O16Rik*$^{+/-}$ mice were fertile and growth normally. Males and females *4930463O16Rik*$^{+/-}$ were mated to obtain *4930463O16Rik*$^{-/-}$ mice. Once the mouse line was established, all pups were genotyped with a combination of primers (listed in S4 Table) (Fig 5B).

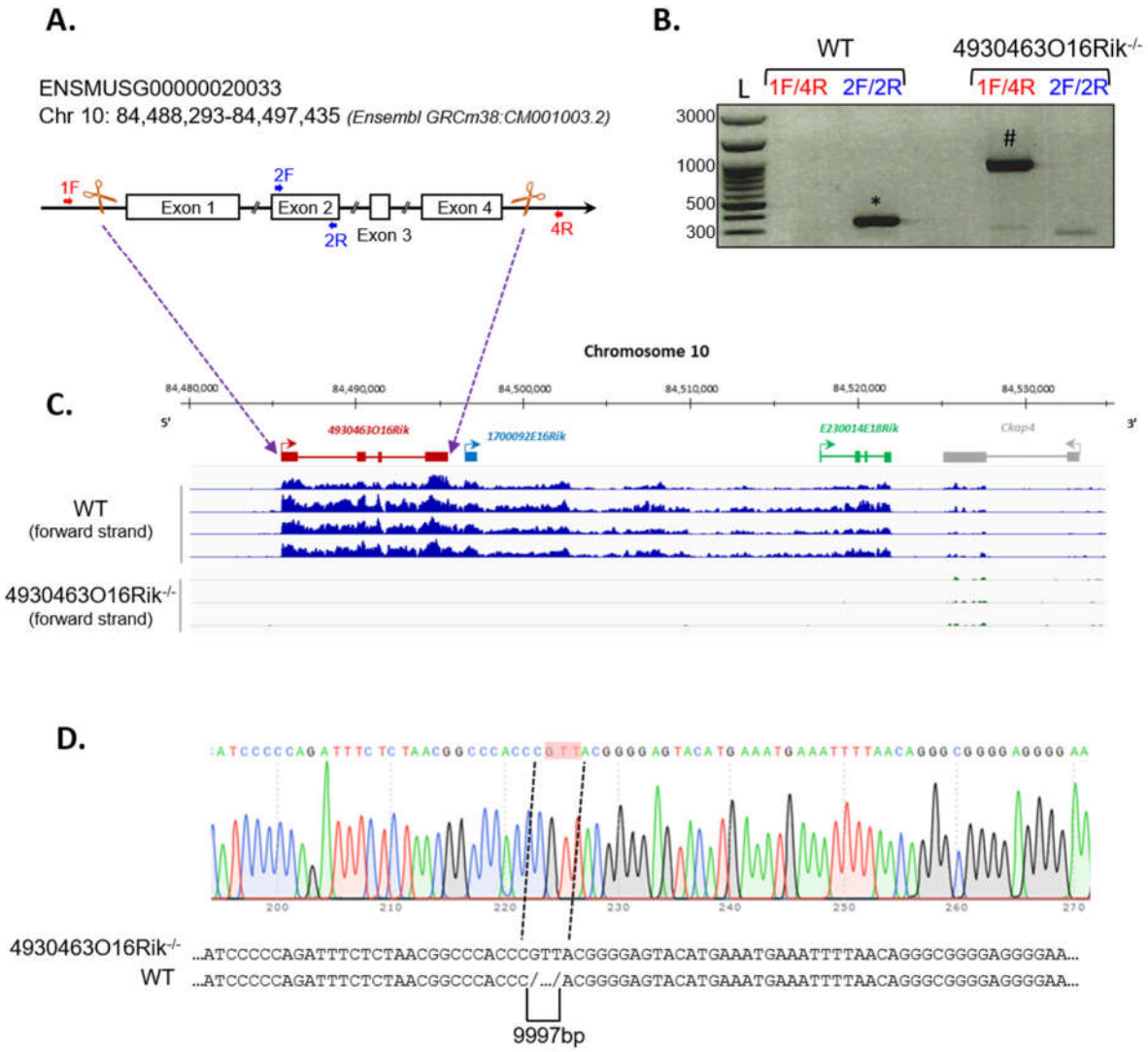
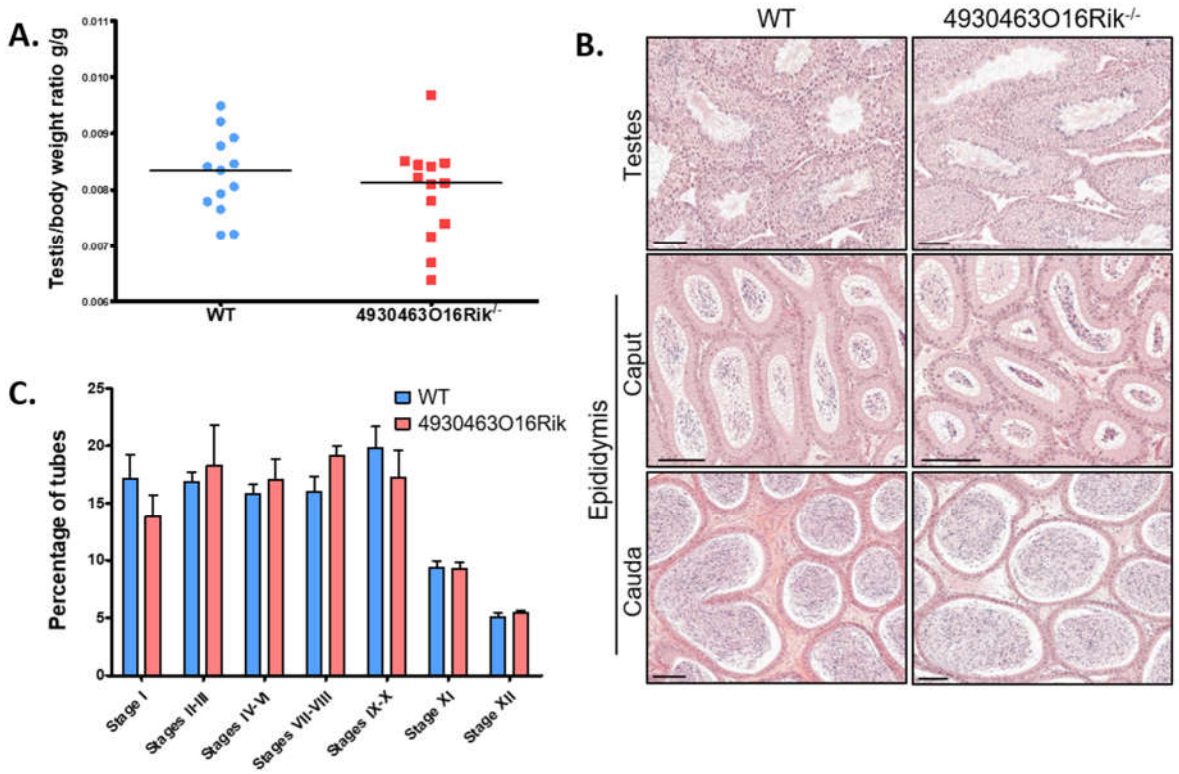


Fig 5: Deletion of the *4930463O16Rik* gene in mouse.

(A) Schematic design of the CRISPR/Cas9 deletion of the *4930463O16Rik* gene with suppression of the 4 exons and 3 introns. The white boxes and lines represent exons and introns, respectively. (B) PCR genotyping on DNA of WT and *4930463O161Rik*$^{-/-}$ mice. The primer pairs used 1F/4R (located in 5' of exon 1 and in 3' of exon 4 of the *4930463O16Rik* gene respectively) or 2F/2R (located in the exon2 of *4930463O16Rik*) were used to determine the genotype of mice. Results showed the following amplicon sizes: (*) 352 bp with the 2F/2R primers in WT (no amplification in mutant mice); (#) 935 bp with the 1F/4R primers on *4930463O161Rik*$^{-/-}$ mice (no amplification in WT mice in the PCR conditions used). (L) DNA ladder. (C) Transcription of the forward strand of chromosome 10 around *4930463O16Rik* gene with RNA-seq coverage (BigWig format representation) in WT (top blue tracks) and *4930463O16Rik*$^{-/-}$ (bottom tracks) mouse P18-testes. A continuous (WT) or very low transcription (*4930463O16Rik*$^{-/-}$) was observed from *4930463O16Rik* until *E230014E18Rik* genes. (D) Electrophoregramm of

304 *4930463O16Rik*$^{-/-}$ mouse genomic DNA showing 9997 bp deletion and an insertion of 3 nucleotides (GTT,
305 highlighted in pink).
306
307 **The absence of *4930463O16Rik* does not affect mouse fertility**
308 Fertility was investigated in *4930463O16Rik* deficient mice. Eight-week-old males and females
309 *4930463O16Rik*$^{-/-}$ mice were matted. *4930463O16Rik*$^{-/-}$ mice were fertile in both sexes. Their litter size
310 (7.5 ± 2.10 pups per litter, n=28) was similar to that of their WT counterparts (6.9 ± 2.12 pups per litter,
311 n=20). There was no significant difference in testicular size, in morphology and histology of testis,
312 cauda and caput epididymis between WT and *4930463O16Rik*$^{-/-}$ adult mice (Fig 6A, B). In addition, the
313 different stages of seminiferous tubules divided into seven groups were quantified between
314 *4930463O16Rik*$^{-/-}$ and WT adult mice. No significant difference was observed between both genotypes
315 (Fig 6C). These results demonstrated that *4930463O16Rik* is not required for mouse fertility.



316
317 **Fig 6: Study of *4930463O16Rik*$^{-/-}$ testicular phenotype.**

(A) Testis/body weight ratio of 8 week-old mice. No significant difference was observed in both mouse lines. Median lines were in black. (B) Haematoxylin and eosin (HE) staining of testis and epididymis sections from WT and *4930463O16Rik*$^{-/-}$ 8 week-old mice. Scale bar=50 µm. Spermatozoa were visible in the lumen of the testis and epididymis of WT and *4930463O16Rik*$^{-/-}$. (C) Quantification of the different seminiferous epithelium stages in WT and *4930463O16Rik*$^{-/-}$ 8 week-old mice. No significant difference was found between WT and *4930463O16Rik*$^{-/-}$ mice.

***4930463O16Rik*$^{-/-}$ mice present modified sperm parameters**

Sperm parameters in 8-week-old testis lacking *4930463O16Rik* were compared to WT testis of the same age. Sperm concentration obtained from epididymis of *4930463O16Rik*$^{-/-}$ mice was significantly reduced by 57.2% compared to WT (Fig 7A) despite an unmodified testis/body weight ratio (Fig 6A). Motility parameters such as the percentage of motile spermatozoa, the motile mean expressed as beat cross frequency (bcf) and the progressive spermatozoa were significantly higher in *4930463O16Rik*$^{-/-}$ mice compared to WT (Fig 7B, C, D). From a morphological point of view, two parameters were significantly modified in the testis of mutant mice: the distal mid-piece reflex (DMR), a defect developing in the epididymis and attesting of the sperm tail abnormality [27] and the percentage of spermatozoa with coiled tail (Fig 7E, F). In addition, two kinetic parameters were also significantly reduced in mutant sperm: the motile mean vsl, related to the progressive velocity in straight line) and the average path velocity, or vap (Fig 7G, H).

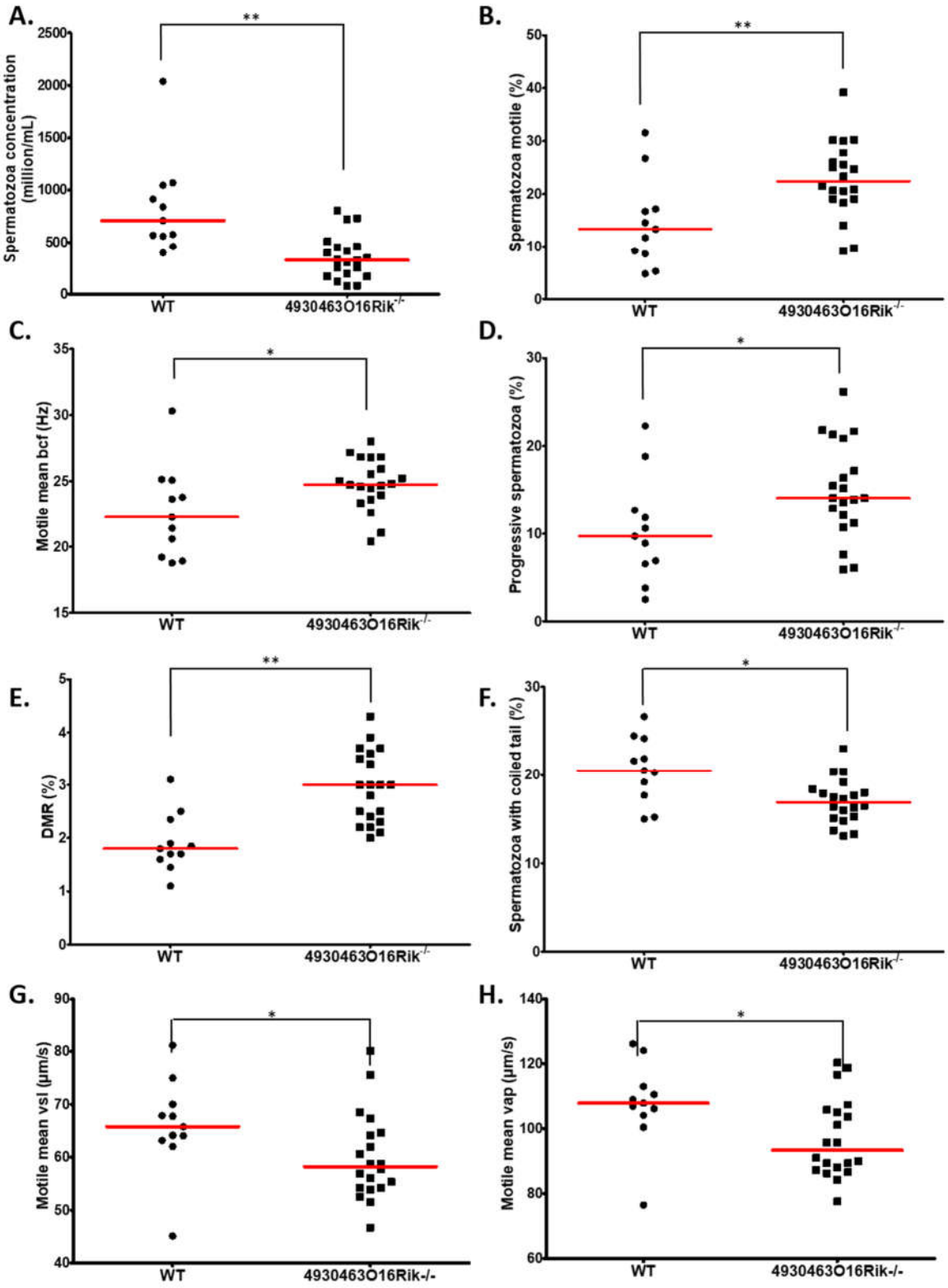


Fig 7: Evaluation of sperm parameters.

Comparison of sperm-specific parameters from WT (circle, n=11) and *4930463O16Rik*$^{-/-}$ (square, n=20) mice. Significantly affected sperm parameters were (A) spermatozoa concentration (10^6/mL), (B) spermatozoa motility

(%), (C) motile mean bcf (beat cross frequency), (D) progressive spermatozoa (%), (E) DMR (distal midpiece reflex, abnormality of the sperm tail, %), (F) spermatozoa with coiled tail (%), (G) motile mean VSL (µm/s) and (H) VAP (µm/s). Statistical analyses were realized with non-parametric Kruskal-Wallis test. * = p-val<0.05, ** = p-val<0.01.

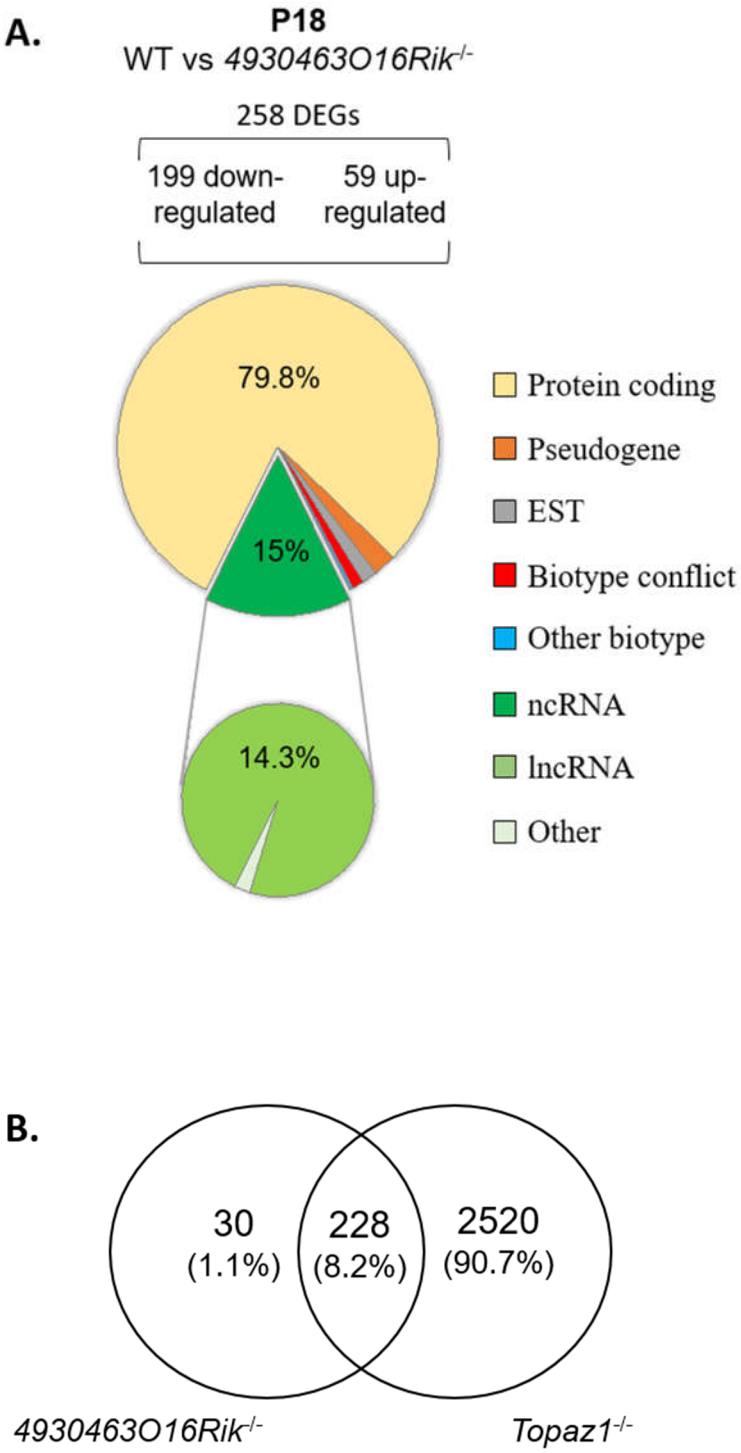
These results showed that several sperm parameters, namely concentration, motility, morphology and kinetics, were impacted by computer-aided sperm analysis (CASA) in the *4930463O16Rik* lncRNA-deficient mice. Some of them would suggest negatively impacting fertility, as the sperm concentration, the DMR, the percentage of coiled tail, the motile mean vsl while others would rather suggest increased fertility, such as the motile mean percentage and bcf, the progressive spermatozoa. These observations might explain the resulting normal fertility of *4930463O16Rik* lncRNA-deficient mice.

Normal male fertility despite *4930463O16Rik*$^{-/-}$ mouse testis transcriptome modified

Transcriptomic RNA-seq analyses were performed in WT and *4930463O16Rik*$^{-/-}$ mouse testes at two developmental stages: P16 and 18, as in the *Topaz1*$^{-/-}$ mouse line.

At P16, seven genes were differentially expressed (adjust p-value<0.05; absolute Log2FC>1) among which, *4930463O16Rik*, but also *1700092E16Rik* and *E230014E18Rik* (S5 Table). These observations reinforced our hypothesis that these last two Riken cDNAs are in fact in the 3' transcribed RNA of *4930463O16Rik* (positioned in Fig 5C) and correspond to a unique locus. Transcriptional activity of this new locus stops towards the 3' end of the *cKap4* gene (cytoskeleton-associated protein 4 or Climp-63). This gene was down-regulated by 1.7 fold in both knock-out lines (*Topaz1* and *4930463O16Rik*). This could suggest a newly discovered positive regulatory role of this lncRNA on the *cKap4* gene.

At P18, 258 genes were differentially expressed (199 down- and 59-up regulated using same statistical parameters, S5 Table). Among them, 206 were protein-coding genes representing 79.8% of DEGs (Fig 8A). Thus, P18 DEGs highlighted a direct or indirect relationship between the loss of the *4930463O16Rik* lncRNA and protein-coding genes. In addition, the loss of this lncRNA also resulted in the deregulation of 37 (14.3%) other lncRNAs.



367

368 **Fig 8: Deregulated genes from *4930463O16Rik*$^{-/-}$ mouse testes.**

369 (A) Biotype of differential expressed gene at P18 in *4930463O16Rik*$^{-/-}$ testis. Most of deregulated gene are coding

370 protein genes (adj p<0.05 and down-regulated FC<2 (log2FC<-1) or up-regulated FC>2 (log2FC>1). (B) Venn

diagram showing overlap of differentially expressed genes between *4930463O16Rik*$^{-/-}$ and *Topaz1*$^{-/-}$ mouse testes at 18 d*pp*.

Validation of several DEGs was performed by RT-qPCR using WT and *4930463O16Rik*$^{-/-}$ testicular RNAs of both developmental stages (P16 and P18) (S9 Fig). The qPCR results for the genes tested were consistent with those of RNA-seq.

The *4930463O16Rik*$^{-/-}$ DEGs were also analyzed with the DAVID database (S6 Table). At P18, six functional clustering have an enrichment score > 1.3 [21]. As for *Topaz1*$^{-/-}$ mouse testis, they included the following GO terms (*i*) for cellular components: cilium movement, ciliary part, axoneme; (*ii*) for biological processes: microtubule-based process, regulation of cilium movement, spermatogenesis, male gamete generation, spermatid development and differentiation (S6 Table). Analyzing by discriminating up- from down-regulated genes only increased the value of the enrichment score for the down-regulated genes. The absence of TOPAZ1 protein or of *4930463O16Rik* lncRNA showed the same enrichment clusters in mutant testes despite different outcomes on the fertility of male mice. The other clusters from the DAVID analysis referred to the GO terms: cell surface, external side of membrane, defense or immune response and response to external stimulus .These clusters were only found in DAVID analysis with up-regulated genes.

Therefore, *4930463O16Rik* gene would appear to regulate genes related to spermatogenesis, microtubule or cil organizations and cytoskeleton in the P18 testis. In the absence of this lncRNA, some genes involved in defense mechanisms or immune response are also deregulated suggesting stressful conditions. It should be noted that the majority (228/258 or 88%) of the DEGs from P18-*4930463O16Rik*$^{-/-}$ testis was in common with those deregulated in *Topaz1*$^{-/-}$ mice (Fig 8B). This led to similar results on ontological analyses of the DEGs of the two mutant lines. These 228 genes could be, in the *Topaz1*$^{-/-}$ testes, a consequence of the down-regulation of the *4930463O16Rik* lncRNA. On the other hand, these 228 DEGs alone do not explain the meiotic arrest in the *Topaz1*$^{-/-}$ testes.

Discussion

Topaz1 was initially reported as a germ-cell specific factor [3], essential for meiotic progression and male fertility in mice [2]. Suppression of *Topaz1* led to an arrest of meiosis progression at the diplotene-metaphase I transition associated with germ cell apoptosis. Moreover, a first transcriptomic approach, based on DNA microarrays, allowed observing a large but not exhaustive repertoire of deregulated transcripts. With this technology, 10% of differentially probes were lncRNAs and presented a deregulated expression in P20 *Topaz1*$^{-/-}$ testes compared to WT.

In this study, we show that the effects of the absence of the *Topaz1* gene are visible on the mouse testicular transcriptome as early as 16 days *post-partum*, i.e. before the first meiotic division and the production of haploid germ cells. These effects are amplified at 18 days post-natal, just before or at the very beginning of the appearance of the first haploid germ cells. The molecular pathways involved in the suppression of Topaz1 belong to spermatogenesis and to the establishment of the cell cytoskeleton. At these two stages, P16 and P18, about a quarter of the deregulated genes in testes are ncRNAs, mainly lncRNAs, some of which show almost no expression in *Topaz1*$^{-/-}$ testis. Suppression of one of them does not prevent the production of haploid spermatids and spermatozoa, but halves the murine sperm concentration. Furthermore, by deleting ~10 kb corresponding to this *4930463O16Rik* lincRNA, we show that the transcriptional extinction is even longer, encompassing ~35 kb in total and two other genes (1700092E16Rik (unknown gene type according to Ensembl) and the lincRNA E230014E18Rik). Our transcriptional data suggest in fact that these three annotated loci belong to a unique gene. Transcription of this lincRNA ended near the 3' region of the *cKap4* gene, a gene known to be associated with the cytoskeleton [28]. Remarkably, cKap4 expression is down-regulated by 1.7 fold in both knockout mice (*Topaz1*$^{-/-}$ and *4930463O16Rik*$^{-/-}$), suggesting a previously unknown positive regulatory role of *4930463O16Rik* on *cKap4*.

TOPAZ1 ablation leads to chromosome misalignments at pro-metaphase I.

Meiosis and its two cell divisions are well-orchestrated sequences of events controlled by different genes. Although these divisions have many similarities between males and females, meiosis is also sex-dimorphic. This particularly concerns timing, synchronization, the number of haploid gametes produced and the periods of meiotic arrest (reviewed by [29]). In females, meiosis is initiated during fetal life and the development of oocytes is arrested at the end of prophase I. Oocytes remain in this arrested state until the onset of ovulatory cycles around puberty. There, the first division of meiosis resumes and leads to the release of a first polar globule with the secondary oocyte. At metaphase II, the oocyte is blocked again. The release of the second polar globule leading to the formation of the female gamete only takes place at fertilization [29]. In males, meiosis is a continuous process taking place in the post-natal period, just before puberty and resulting in the formation of four male gametes from one spermatocyte. Despite these sex-dimorphic differences, the first reductional division of meiosis is highly conserved between species and between sexes in terms of morphology and genetic regulation. It was hypothesized that the mechanisms regulating and controlling prophase I during mammalian meiosis, frequently named "checkpoints", are more stringent in males than in females. This has been demonstrated over the last 25 years by the use of a large number of mutant mouse models, mainly gene knockout mice [30–32].

A major checkpoint in males is the synaptic checkpoint that controls the zygotene-pachytene transition, highlighted in male mice lacking *Sycp3*, *Dmc1*, *Spo11*, *mei1*, *Msh4-5* or *OvoL1* genes [33–40]. In mutant females, this synaptic checkpoint is less stringent. Indeed, female meiotic arrest may occur later, starting from the diplotene, as seen in *Dmc1*, *mei1*, *Msh4-5* knockout mice, during the dictyate-resting phase of the oocyte, evidenced *Spo11*. These female mice can even be fertile, as seen in *OvoL1* knockout mouse [40]. Other gene suppressions have highlighted a second meiotic checkpoint of metaphase I in male, such as those of *Mlh1-3*, or cyclin A1, due to a misalignment of chromosomes on the spindle [41–43]. Mice devoid of *Topaz1* gene could match these latter models. Indeed, *Topaz1*$^{-/-}$

spermatocytes do not progress to metaphase I and the chromosomes are not properly aligned on the metaphase plate.

TOPAZ1 seems to be involved in the cellular shape, structure and movements.

The absence of *Topaz1* gene disturbs the transcriptome of the murine testes as early as 16 days postnatal. Of the 205 DEGs at P16, 85 are specific to this stage of development compared to P18 (Fig 1A), such as *Ptgs1* (*Cox1*), marker of peritubular cells [44] and *Krt18*, marker of Sertoli cell maturation [45]. Moreover, different genes involved in the TGFβ pathway are also P16-DEGs such as *Bmpr1b*, *Amh*, *Fstl3*. This last one, for example, was demonstrated to reduce the Sertoli cells numbers in mouse testis and to limit the organ size testis [46]. *Ptgds* (L-*Pgds*) playing a role in the PGD2 molecular pathway in mammalian testicular organogenesis is also deregulated [47]. All these genes specifically deregulated at P16 due to the absence of *Topaz1* thus seem to participate in the regulation of cell communication. At P18, these genes are no longer differential, but genes belonging to the same gene families can replace them, such as cadherin or keratin families for example.

Many of the 205 DEGs at P16 are involved in the defense response pathways. For example, *Ifit3* and *Gbp3* are immune response genes in spermatocyte-derived GC-2spd(ts) cell [48].

Two days later, at P18, just before the prophase I-metaphase 1 transition, 10 times more genes are deregulated. Among the 120 DEGs common to P16 and P18, there is at least one gene that may be involved in meiosis such as *Aym1*, activator of yeast meiotic promoters 1. The absence of *Topaz1* leads to the absence of testicular expression of *Aym1*. This gene is germ cell-specific [49]. In male mice, *Aym1* is expressed from 10 d*pp* in early meiotic spermatocytes. The small murine AYM1 protein (44 amino acids) is immunolocalized in the nucleus of primary spermatocytes, mainly late pachytene and diplotene, suggesting a nuclear role of AYM1 in germ cells in first meiotic division[49].

At P18, the testicular transcriptome of *Topaz1*$^{-/-}$ mice is largely disturbed compared to WT and most of DEGs are down-regulated (Fig 1F), suggesting that TOPAZ1 promotes gene expression in normal mice. As TOPAZ1 is predicted to be an RNA-binding protein, it is tempting to speculate that its absence

473 disorganized ribonucleic-protein complexes, including their instabilities and degradation. This could
474 partly explain why 90% of DEGs are down-regulated at P18 with a large proportion of lincRNAs. These
475 down-regulated genes at P18 referred to microtubule-based movement and microtubule-based
476 process and cellular components refer to motile cilium, ciliary part, sperm flagellum, axoneme. In
477 addition, DAVID analysis revealed GO terms such as centriole, microtubule and spermatogenesis. All
478 these terms relate to elements of the cytoskeleton, which are indispensable for mitotic and/or meiotic
479 divisions, motility, differentiation and are also widely involved in spermiogenesis, as would be
480 expected with this latter GO terms since most DEGS are testis-specific. Centriole is a widely conserved
481 organelle in most organisms. A pair of centrioles is located into the heart of the centrosome, and the
482 whole is grouped together as the main microtubule-organizing center (MTOC). Here, the staining of
483 the meiotic spindle and centrosomes shows a disturbance of these pathways (Fig 2 and S2 Fig). Such
484 abnormal metaphase-like chromosomes arranged in rosette instead of being neatly aligned at the cell
485 equator and such hemispindles centered in the spermatocytes have already been observed. For
486 example, aberrant prometaphase-like cells were observed in *Mlh1*- or *Meioc*-deficient testes (*Meioc* is
487 down-regulated by 1.51 fold in P18 in *Topaz1*$^{-/-}$ testis) [41,50]. These mutant mice have been described
488 with an arrest of male meiosis, testes devoid of haploid germ cells leading to male sterility like mice
489 lacking the *Topaz1* gene. In *Topaz1* Mlh1 is not a DEG.
490
491 During spermatogenesis, dysregulation of centrosomal proteins may affect meiotic division and
492 genome stability. The centriole proteins CEP126, CEP128, CEP63 are down-regulated (FC from 2.1 to
493 2.7 compared to WT) at P18 in *Topaz1*$^{-/-}$ testis. CEP126 is localised with γ-tubulin on centriole during
494 mitosis of hTERT-RPE-1, human telomerase-immortalized retinal pigmented epithelial [51] but has
495 never been studied in germ cells during meiosis. CEP128 was localized to the mother centriole and
496 required for regulating ciliary signalling in zebrafish [52]. *Cep128* deletion decreased the stability of
497 centriolar microtubules in F9 cells (epithelial cells from testicular teratoma of mouse embryo) [53].
498 Normally, centriole separation occurs at the end of the prophase I or in early metaphase I and CEP63

499 is associated with the mother centrioles. The mouse model devoid of *Cep63* lead to male infertility
500 [54]. In spermatocytes of these mice, the centriole duplication was impaired. Finally, the ontology
501 analysis of *Topaz1*$^{-/-}$ P18-DEGs revealed significant enrichment score for the several clusters referred
502 to the spermatozoa final structure, like tetratricopeptide repeat (TPR) and dynein heavy chain
503 (DNAH1). Dynein chains are macromolecular complexes connecting the central pair or doublet pairs of
504 microtubules together to form the flagellar axoneme, the motility apparatus of spermatozoa (ref in
505 [55]). Dynein proteins have also been identified and involved in microtubule-based, intracellular
506 transport of vesicles, and in both mitosis and meiosis [56].
507 The TPR or PPR (pentatricopeptide repeat) domains consist of several 34 or 36 amino acid repeats
508 respectively constituting αα-hairpin repeat units [57]. The functions of TPR or PPR proteins were firstly
509 documented in plants and are involved in RNA editing [57,58]. In mouse, *Cfap70*, a tetratricopeptide
510 repeat-containing gene was shown to be expressed in the testis [59], or as *Spag1* rather in late-
511 pachytene spermatocytes or round spermatids [60]. Moreover, knockout mouse for *Ttc21a* shown
512 sperm structural defects of the flagella and the connecting piece. In human, *Ttc21a* was associated
513 with asthenoteratospermia in the Chinese population [61]. Many components of the intraflagellar
514 transport (IFT) complex contain TPR. Several genes coding for such tetratricopeptide repeat-containing
515 proteins are down-regulated in the P18 testis devoid of *Topaz1* like *Cfap70*, *Spag1*, *Tct21a*, *Ift140*.
516 Based on TPRpred [62], that predicted TPR- or PPR-containing proteins, the TOPAZ1 protein was
517 predicted to contain such domains, 7 in mice (p-val= 7.5E-08, probability for being PPR= 46.80%) and
518 10 in humans (p-val= 3.4E-09, probability for being PPR = 88.76%).
519 A recent study of single cell-RNA-seq from all types of homogeneous spermatogenetic cells identified
520 clusters of cells at similar developmental stages [63]. This study shown that most of genes involved in
521 spermiogenesis begin to be expressed from the early pachytene stage. This is consistent with our RNA-
522 seq results. Together, these data indicated that the absence of *Topaz1* down-regulated a significant
523 number of cytoskeleton-related genes which leads to a defect in the formation of the meiotic spindle
524 and to a deficient duplication and/or migration of centrosomes as early as 18 days post-natal. *Topaz1*

525 could lead to impaired chromosome dynamics via activation of cytoskeleton-genes and thus shows an
526 essential role for centrosome in promoting division and then fertility. The action of TOPAZ1 could be
527 carried out via its TPR domains.
528
529 **Topaz1 ablation deregulates a high proportion of lncRNAs**
530 DEGs between *Topaz1*$^{-/-}$ and WT mouse testes shown also a high proportion of deregulated lncRNAs.
531 We showed that three lincRNAs, whose expression is almost abolished as early as P16 in mouse
532 testicles lacking *Topaz1*, are testis- and germ cell-specific. We have shown that these genes are
533 expressed in spermatocytes and round spermatids suggesting a role in spermatogenesis. Their
534 functions are unknown.
535 Several investigations revealed that the testis is an organ allowing the expression of many lncRNAs
536 [64]. In mammals, the testis is the organ with the highest transcription rate [65]. However, during the
537 long stage of prophase I, the levels of transcription are not consistent. Indeed, transcription is strongly
538 reduced or even abolished in the entire nucleus of the spermatocytes during the early beginning of
539 prophase I. This is accompanied in particular by the nuclear processes of DNA division, pairing
540 homologous chromosome, telomeric rearrangement [66–68] and also by the onset of the MSCI,
541 meiotic sex chromosome inactivation, markers [69]. These processes are supported by epigenetic
542 changes such as histone modifications, recruitment of specific histone variants (references in Page et
543 al., 2012). Then, transcription takes up an important part in the late-pachytene until diplotene
544 spermatocytes [70]. The above-mentioned scRNA-seq study from individual spermatogenic cells
545 shown that almost 80% of annotated autosomal lncRNAs were expressed in spermatogenetic cells,
546 mainly from the mid-pachytene- to the metaphase I-spermatocytes but also in round spermatids [63].
547 The three lncRNAs investigated in our study (*4930463O16Rik, Gm21269* and *4921513H07Rik*) are also
548 expressed at these stages of development in the mouse testes [71,72]. In this last mentioned study
549 [72], the authors identified certain male germline-associated lncRNAs as potentially important for
550 spermatogenesis *in vivo*, based on several computational and experimental data sets. Among them

551 are *Gm21269* and *4921513H07Rik*. The localization of the lncRNAs in cells may be an indicator of their
552 potential function [73]. *4930463O16Rik* is expressed in the nucleus of spermatocytes. As mentioned
553 above, *4930463O16Rik* could have a positive role on the expression of *cKap4* at the neighboring locus.
554 Some nuclear lncRNA show transcription regulation with *cis*-regulatory role, such as *Malat1* or *Air*
555 [74,75] on their nearby gene. Other nuclear lncRNAs act in *trans* and regulate gene transcription at
556 another locus, such as *HOTAIR* [76]. In addition, some cytoplasmic lncRNA have been shown to play a
557 role in miRNA competition as miRNA sponges or decoys (such as *linc-MD1* in human myoblasts [77]).
558 *Gm21269* is localized in the cytoplasm and nuclei of spermatocytes during meiosis. Both cytoplasmic
559 and nuclear lncRNAs may act as molecular scaffold to assemble functional protein complexes, as
560 *HOTAIR* or *Dali* [78,79], to regulate protein localization and/or direct protein degradation, or as miRNA
561 precursor [80]. Finally, multiple other roles can be observed for lncRNAs. For example, the *Dali* lincRNA
562 locally regulates his neighboring *Pou3f3* gene, acts as a molecular scaffold for POU3F3 protein and
563 interacts with DNMT1 for regulation of DNA methylation status of CpG island-associated promoters in
564 *trans* during neural differentiation [79].
565

566 **Deletion of one lncRNA alters sperm parameters without affection fertility**

567 To decipher the biological function of a lncRNA affected by *Topaz1* invalidation, mouse model devoid
568 of *4930463O16Rik* were produced, under the same genetic background as for *Topaz1*$^{-/-}$ mice. This
569 knockout, did not disturb the meiosis and fertility of the mutant mice under standard laboratory
570 conditions. On a similar approach, *Sox30* is a testis-specific factor indispensable to obtain haploid germ
571 cells during spermatogenesis [81]. SOX30 regulates *Dnajb8* expression. However, deletion of *Dnajb8* is
572 not essential for spermatogenesis and male fertility [82].
573 Several mutant mouse deprived of testis-specific genes proved to be fertile, without any role being
574 stated during spermatogenesis. This was noted in particular for the *Flacc1*, *Trim69*, *Tex55*,
575 *4930524B15Rik* genes [83–86] or for highly testis-enriched genes like *Kdm4d*, *Tex37*, *Ccdc73* or *Prss55*
576 [87,88]. Some of them are down-regulated genes in *Topaz1*$^{-/-}$ or *4930463O16Rik*$^{-/-}$ testis (*Trim69* in

577 *Topaz1*$^{-/-}$ FC = 3.99 and in *4930463O16Rik*$^{-/-}$ FC = 2.27; *Kdm4d* in *Topaz1*$^{-/-}$ FC = 2.70 and in
578 *4930463O16Rik*$^{-/-}$ FC = 1.83; *Ccdc73* in *Topaz1*$^{-/-}$ FC = 1.46). Recently, some laboratories have also
579 generated several dozen of testis-enriched knockout mouse lines using the CRISPR/Cas9 system and
580 shown that all these genes are individually dispensable for male fertility in mice [89,90].
581 The abundant expression of lncRNAs during spermatogenesis has also prompted other laboratories to
582 perform knockout mouse model of testis-specific lncRNAs. This is the case for the *1700121C10Rik* or
583 *lncRNA5512* lncRNAs where mutant mice were also fertile without variation in sperm parameters
584 [91,92]. One working hypothesis could be that some lncRNAs might be regulators of subset of
585 functional spermatogenetic-gene expression, in agreement with their nuclear localization, by binding
586 to their regulatory genomic region.
587 Nevertheless, in our *4930463O16Rik*-knockout mouse model, several sperm parameters were altered,
588 including a reduction of epididymal sperm concentration by more than half and sperm motility. In the
589 *Tslrn1* knockout mice, testis-specific long non-coding RNA 1, males are fertile, showed a significant
590 20% sperm reduction, but no reduction in litter size or major defects in testis histology or variation of
591 sperm motility [93]. In *Kif9*-mutant male mice, no testis abnormalities were found [94]. They were
592 subfertile due to impaired sperm motility: the velocity parameters VSL and VAP were reduced, like in
593 the *4930463O16Rik* knockout mice. The authors concluded that *Kif9* mutant mice were still fertile
594 probably due to variations in the motility of individual spermatozoa. These spermatozoa with good
595 motility could still fertilize oocytes. The same conclusion could apply to mice *4930463O16Rik*$^{-/-}$.
596 Suppression of a gene, in this case the *4930463O16Rik* lincRNA, whose expression is highly down-
597 regulated in the testes of sterile *Topaz1*$^{-/-}$ mice (FC = 40), has no effect on spermatogenesis. Our data
598 suggest that the expression of *4930463O16Rik* is dispensable for meiotic divisions, but add to the
599 terminal differentiation of male germ cells.
600
601 Various genes, testis-specific or highly expressed in the testis, show no effect on reproduction when
602 deleted independently [90,91]. One explanation, given the large number of lncRNAs expressed in the

603 meiotic testis, could be that the function of *4930463O16Rik* is partly redundant with that of other
604 testicular lncRNAs.
605 However, outside the laboratory, in wild reproductive life, we could imagine that biological functions
606 in more natural conditions could be different due to stress and reproductive competition. This has
607 been shown in particular for *Pkdrej*-deficent male mice which are fertile, whereas the *Pkdrej* gene,
608 polycystin family receptor for egg jelly, is important in postcopulatory reproductive selection [90,95].
609
610 The absence of a specific anti-TOPAZ1 antibody does not allow us to futher advance in the
611 understanding of its function during murine spermatogenesis. The creation of a Flag-tagged *Topaz1*
612 knockin mouse model will allow us to get deeper inside. Rip-seq experiments will allow us to determine
613 RNA-TOPAZ1 complexes during spermatogenesis.
614
615

616 ## Materials and methods

617 ### Ethics statement

618 All animal experiments were performed in strict accordance with the recommendations in the
619 guidelines of the Code for Methods and Welfare Considerations in Behavioral Research with Animals
620 (Directive 2016/63/UE). Experiments were approved by the INRAE local animal experiment ethics
621 committee of Jouy-en-Josas (Comethea, number 18-12) and authorization issued from the French
622 Ministry of Higher Education, Research and Innovation (Number 815-2015073014516635).
623

624 ### Mice

625 Generation and preliminary analysis of the *Topaz1*-null transgenic mouse line was previously described
626 (Luangpraseuth-Prosper *et al*. 2015).

The generation of 4630493O16Rik-null transgenic mouse line was obtained using the CrispR-Cas9 genome editing technology. RNA mix comprising an mRNA encoding for the SpCas9-HF1 nuclease and the four sgRNA (S4 Table) targeting the *4930463o16Rik* gene (NC_000076: 84324157-84333540) were prepared. These sgRNAs were chosen following the CRISPOR software (http://crispor.tefor.net/) in order to remove the 4 exons and introns of the *4930463o16Rik* gene. Cas9-encoding mRNA and the 4 sgRNAs were injected at 100 ng/µL each into 1 cell-fertilized C57Bl/6N mouse eggs [96].

Surviving injected eggs were transferred into pseudo-pregnant recipient mice. Tail-DNA analysis of the resulting alive pups was performed by PCR using genotyping oligonucleotides (S4 Table) and Takara Ex Taq® DNA Polymerase kit. PCR conditions were 94 °C 30s, 60 °C 30s and 72 °C 30s with 35 amplification cycles.

Two transgenic founder mice were crossed with wild-type C57Bl/6N mice to establish transgenic lines. F1 heterozygote mice were crossed together in each line to obtained F2 homozygote mice, thus establishing the 4630493O16Rik$^{-/-}$ mouse lines. Both mouse lines were fertile, the number of pups was equivalent, so we worked with one mouse line.

All mice were fed *ad libitum* and were housed at a temperature of 25°C under a 12h/12h light/dark cycle at the UE0907 unit (INRAE, Jouy-en-Josas, France). Mice were placed in an enriched environment to improve their reception while respecting the 3R. All mice were sacrificed by cervical dislocation. Tissues at different developmental stages were dissected and fixed as indicated later or flash frozen immediately in liquid nitrogen before storage at -80°C. Frozen tissues were used for molecular biology experiments described later.

Histological and immunohistochemical analyses

For histological studies, fresh tissues from 8-week-old mice were fixed in 4% paraformaldehyde (Electron Microscopy Sciences reference 50-980-495) in phosphate buffer saline (PBS) at 4°C. After rinsing tissues in PBS, they were stored in 70% ethanol at 4°C. Then, paraffin inclusions were performed with a Citadel automat (Thermo Scientific Shandon Citadel 1000) with standard protocol. Tissues

653 including in paraffin blocks were sectioned at 4µm and organized on Superfrost Plus Slides (reference
654 J1800AMNZ). Once dry, slides were stored at 4°C. On the day of the experiment, these slides with the
655 sectioned tissues were deparaffinized and rehydrating in successive baths of xylene and ethanol at
656 room temperature. For histology, testis sections were stained with hematoxylin and eosin (HE) by the
657 @Bridge platform (INRAE, Jouy-en-Josas) using an automatic Varistain Slide Stainer (Thermo Fisher
658 Scientific). To determine the seminiferous epithelium stages, we used periodic acid-Schiff staining
659 (PAS).
660 *In situ* hybridization experiments were performed using the RNAscope® system (ACB, Bio-Techne SAS,
661 Rennes, France). Briefly, probes (around 1000 nt long) for *Topaz1* (NM_001199736.1), *4930463o16Rik*
662 (NR_108059.1), *Gm21269* (NR_102375.1) and *4921513H07Rik* (NR_153846.1) were designed by the
663 ACB company and referenced with catalog numbers 402321, 431411, 549421 and 549441 respectively.
664 Negative and positive control were ordered from ACD with *Bacillus subtilis* dihydrodipicolinate
665 reductase (dapB) and *Homo sapiens* ubiquitin C (Hs-UBC) respectively. Hybridization was performed
666 according to the manufacturer's instructions using the labelling kit (RNAscope® 2.5HD assay-brown).
667 Brown labelling slides were counterstained PAS staining protocol and then, were observed as visible
668 signal. Hybridization was considered as positive when at least one dot was observed in one cell.
669 Colored sections were scanned using a 3DHISTECH panoramic scanner at the @Bridge platform (INRAE,
670 Jouy-en-Josas) and were analyses with Case Viewer software (3DHISTECH). We also used the
671 RNAscope® 2.5HD assay-red kit in combination with immunofluorescence in order to achieve
672 simultaneous visualization of RNA and protein on the same slide. In this way, the *IS*H protocol was
673 stopped with water immersion before hematoxylin counterstain. Instead, slides were washed in PBS
674 at room temperature. Mouse on mouse (M.O.M.) kit (BMK-2202, Vector laboratories) was used and
675 slides were incubated one hour in Blocking Reagent, 5 minutes in Working solution and 2 hours with a
676 primary antibody: DDX4 (ab13840, Abcam) or γH2AX(Ser139) (Merck), diluted at 1:200 in Blocking
677 Reagent. Detection was performed by using secondary antibody conjugated to DyLight 488 (green,
678 KPL). Then, diluted DAPI (1:1000 in PBS) was applied during eight minutes on slides. Slides were

679 mounted with Vectashield Hard Set Mounting Medium for fluorescence H-1400 and images were
680 captured at the MIMA2 platform (https://www6.jouy.inrae.fr/mima2/,
681 https://doi.org/10.15454/1.5572348210007727E12) with an inverted ZEISS AxioObserver Z1
682 microscope equipped with an ApoTome slider, a Colibri light source, Axiocam MRm camera. Images
683 were analyzed with the Axiovision software 4.8.2 (Carl Zeiss, Germany).
684
685 **Total RNA extraction and Quantitative RT-PCR (RT-qPCR)**
686 Total RNAs from post-natal mouse testis or other organs were isolated using Trizol reagent. RNAs were
687 purified with RNeasy Mini kit (Qiagen) following manufactures instruction and DNAse-treated
688 (Qiagen). Quantification of total RNAs was realized with a Qbit® Fluorometric Quantitation. Maxima
689 First-Strand cDNA Synthesis Kit (Thermo Scientific) was used to reverse transcript RNA in cDNA. Step
690 One system with Fast SYBR™ Green Master Mix (Applied Biosystems, ThermoFisher France) was used
691 for qPCR. qPCR was performed in duplicates for all tested genes and results were normalized with
692 qBase+ software (Biogazelle) [97]. Gapdh, Ywahz and Mapk1 were used as reference genes. For each
693 experiment, median values were plotted with GraphPad Prism, and statistical analyses were
694 performed with KrusKall-Wallis test in R software (Rcmdr package (p-value<0.05)). Primer sequences
695 used for RT-qPCR were provided in S4 Table.
696
697 **RNA-sequencing**
698 Total RNA quality was verified on Agilent 2100 Bioanalyser (Matriks, Norway) and samples with a RIN>9
699 were made available for RNA-sequencing. This work has benefited from the platform and expertise of
700 the High-throughput Sequencing Platform of I2BC (Gif-sur-Yvette, Université Paris-Saclay, France) for
701 oriented library preparation (Illumina Truseq RNA Sample Preparation Kit) and sequencing (Paired-end
702 75 bp; NextSeq). More than 38 million of 75 bp paired-end reads per sample were generated.
703
704 **Transcriptomic analysis**

Sequence libraries were aligned with the Ensembl 95 genome, with TopHat [98], and gene table counts were obtained by applying featureCounts to these alignments [99]. Data normalization and single-gene level analysis of the differential expression were performed using the DESeq2 [100]. Some samples were sequenced several months apart. A batch effect was observed after the computation of hierarchical clustering of samples. In order to take this effect into account, we introduced the batch number into the DESeq2 model, as well as the studied condition. Differences were considered significant for Benjamini-Hochberg adjusted p-values < 0.05, and absolute fold change >2 (absolute Log2FC>1) [101]. The RNA-seq raw data have been deposited through SRA Submission portal (https://submit.ncbi.nlm.nih.gov/subs/sra/) BioProject ID PRJNA698440.

Biotype determination of DEGs

Data available on the websites of NCBI, MGI (http://www.informatics.jax.org) and Ensembl (https://www.ensembl.org/) were used simultaneously to determine the DEG biotypes. For this purpose, the mouse genome information were obtained by ftp from NCBI (ftp://ftp.ncbi.nih.gov/gene/DATA/GENE_INFO/Mammalia/Mus_musculus.gene_info.gz); the annotation BioMart file from Ensembl (http://www.ensembl.org/biomart/martview; Ensembl genes 99, Mouse genes GRCm28.p6) and the feature types from MGI (http://www.informatics.jax.org/marker/; with protein coding gene, non-coding RNA gene, unclassified gene and pseudogenic region). Only data corresponding to DEGs were conserved. The files from these three databases were therefore cross-referenced to determine the biotype of DEGs. When the biotype of a gene was different between the databases, the annotation was then listed as genes with "biotype conflict".

Gene ontology enrichment

729 Identified mouse DEGS were analyzed with Gene Ontology (GO) and Kyoto Encyclopedia of Genes and
730 Genomes (KEGG) pathway membership with Database performed using the DAVID Bioinformatic
731 Database 6.8 (https://david.ncifcrf.gov/). These analyses and pathways were considered significant for
732 a Benjamini corrected enrichment p-value of less than 0.05. Mouse Atlas Genome of differentially
733 expressed genes extracted from this study was performed via the Enrichr website
734 (https://maayanlab.cloud/Enrichr/).
735
736 **Sperm analysis**
737 Evaluation of the semen concentration and motility of WT and *4930463O16Rik*$^{-/-}$ 8-week-old mice were
738 realized by using the IVOS II Computer Assisted Sperm Analysis (CASA) system (Hamilton Thorne,
739 Beverly, MA, USA). The both fresh cauda epididymis per individual were removed and plunged into
740 200 µL of TCF buffer (Tris, citrate and fructose buffer) where they were chopped with small scissors.
741 For sperm release, the samples were incubated 10 minutes at 37°C. A 4 µl aliquot was placed in a
742 standardized four-chamber Leja counting slide (Leja Products B.V., Nieuw-Vennep, the Netherlands).
743 Ten microscope fields were analyzed using the predetermined starting position within each chamber
744 with an automated stage. Statistical analyses were performed using the mean of the 10 analysed fields
745 with at least 300 cells. The IVOS settings chosen was those defined for mice sperm-cell analysis (by
746 Hamilton Thorne). The principal parameters were fixed as follows: 45 frames were captured at 60 Hz.
747 Concerning the cell detection, the camera considered a signal as a spermatozoon when the elongation
748 percentage was between 70 (maximum) and 2 (minimum); the minimal brightness of the head at 186,
749 and the minimum and maximum size of the head at 7 and 100 µm² respectively. Used kinematic
750 thresholds were: cell travel max at 10µm, progressive STR at 45%, progressive VAP at 45µm/s, slow
751 VAP at 20µm/s, slow VSL at 30µm/s, static VAP at 4µm/S and static VSL at 1µm/s. The full settings used
752 was listed in S7 Table. The recorded CASA parameters included the average path velocity (VAP in
753 µm/s), straight line velocity (VSL in µm/s), curvilinear velocity (VCL in µm/s), amplitude of lateral head
754 displacement (ALH in µm), motility (percentage), and sperm concentration (.10^6/mL). The slow cells

were recorded as static. Median and interquartile range were plotted with GraphPad. To compare the sperm parameters between WT and *4930463O16Rik*$^{-/-}$ mice, statistical analyses were performed with the Kruskal-Wallis non-parametric test.

Acknowledgments

We warmly thank Jean-Luc Vilotte for allowing lncRNA deletion to take place in his laboratory, also for the proofreading of this manuscript and for English corrections. We would like to thank the TACGENE facility within the U1154-UMR7196, MNHN, Paris, especially Anne De Cian and Jean-Paul Concordet, for the synthesis of gRNAs and Cas9 mRNA. We are grateful to all the members of our mouse experimental unit (IERP, INRA, Jouy en Josas, France). We thank the @Bridge plateform for Agilent Bioanalyzer using and for the histology facility (UMR 1313 GABI, Jouy-en-Josas, France), particularly, Marthe Vilotte, for HE staining. We also thank the MIMA2 platform for providing access to the virtual slide scanner (Panoramic SCAN, 3DHISTECH). We are grateful to all the members of the "Gonad Differentiation and its Disturbances" team for our scientific and technical exchanges on Monday mornings either in person or by videoconference.

References

1. Kleckner N. Meiosis: how could it work? Proc Natl Acad Sci U S A. 1996;93: 8167–8174. doi:10.1073/pnas.93.16.8167

2. Luangpraseuth-Prosper A, Lesueur E, Jouneau L, Pailhoux E, Cotinot C, Mandon-Pépin B. TOPAZ1, a germ cell specific factor, is essential for male meiotic progression. Dev Biol. 2015;406: 158–171. doi:10.1016/j.ydbio.2015.09.002

3. Baillet A, Le Bouffant R, Volff JN, Luangpraseuth A, Poumerol E, Thépot D, et al. TOPAZ1, a novel germ cell-specific expressed gene conserved during evolution across vertebrates. PLoS One. 2011;6: e26950. doi:10.1371/journal.pone.0026950

4. Brannan CI, Dees EC, Ingram RS, Tilghman SM. The product of the H19 gene may function as an RNA. Mol Cell Biol. 1990;10: 28–36. doi:10.1128/mcb.10.1.28

5. Brockdorff N, Ashworth A, Kay GF, McCabe VM, Norris DP, Cooper PJ, et al. The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. Cell. 1992;71: 515–526. doi:10.1016/0092-8674(92)90519-i

6. Gil R, Latorre A. Factors behind junk DNA in bacteria. Genes (Basel). 2012;3: 634–650. doi:10.3390/genes3040634

7. ENCODE Project Consortium, Birney E, Stamatoyannopoulos JA, Dutta A, Guigó R, Gingeras TR, et al. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature. 2007;447: 799–816. doi:10.1038/nature05874

8. Carninci P, Kasukawa T, Katayama S, Gough J, Frith MC, Maeda N, et al. The transcriptional landscape of the mammalian genome. Science. 2005;309: 1559–1563. doi:10.1126/science.1112014

9. Morris KV, Mattick JS. The rise of regulatory RNA. Nat Rev Genet. 2014;15: 423–437. doi:10.1038/nrg3722

10. 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. doi:10.1002/mrd.23041

11. Quan G, Li J. Circular RNAs: biogenesis, expression and their potential roles in reproduction. J Ovarian Res. 2018;11: 9. doi:10.1186/s13048-018-0381-4

12. Yadav RP, Kotaja N. Small RNAs in spermatogenesis. Mol Cell Endocrinol. 2014;382: 498–508. doi:10.1016/j.mce.2013.04.015

13. Jarroux J, Morillon A, Pinskaya M. History, Discovery, and Classification of lncRNAs. Adv Exp Med Biol. 2017;1008: 1–46. doi:10.1007/978-981-10-5203-3_1

14. Necsulea A, Kaessmann H. Evolutionary dynamics of coding and non-coding transcriptomes. Nat Rev Genet. 2014;15: 734–748. doi:10.1038/nrg3802

15. Sarropoulos I, Marin R, Cardoso-Moreira M, Kaessmann H. Developmental dynamics of lncRNAs across mammalian organs and species. Nature. 2019;571: 510–514. doi:10.1038/s41586-019-1341-x

16. Wen K, Yang L, Xiong T, Di C, Ma D, Wu M, et al. Critical roles of long noncoding RNAs in Drosophila spermatogenesis. Genome Res. 2016;26: 1233–1244. doi:10.1101/gr.199547.115

17. Arun G, Akhade VS, Donakonda S, Rao MRS. mrhl RNA, a long noncoding RNA, negatively regulates Wnt signaling through its protein partner Ddx5/p68 in mouse spermatogonial cells. Mol Cell Biol. 2012;32: 3140–3152. doi:10.1128/MCB.00006-12

18. Anguera MC, Ma W, Clift D, Namekawa S, Kelleher RJ, Lee JT. Tsx produces a long noncoding RNA and has general functions in the germline, stem cells, and brain. PLoS Genet. 2011;7: e1002248. doi:10.1371/journal.pgen.1002248

19. Zhang L, Lu H, Xin D, Cheng H, Zhou R. A novel ncRNA gene from mouse chromosome 5 trans-splices with Dmrt1 on chromosome 19. Biochem Biophys Res Commun. 2010;400: 696–700. doi:10.1016/j.bbrc.2010.08.130

20. Drumond AL, Meistrich ML, Chiarini-Garcia H. Spermatogonial morphology and kinetics during testis development in mice: a high-resolution light microscopy approach. Reproduction. 2011;142: 145–155. doi:10.1530/REP-10-0431

21. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009;4: 44–57. doi:10.1038/nprot.2008.211

22. Huang DW, Sherman BT, Zheng X, Yang J, Imamichi T, Stephens R, et al. Extracting biological meaning from large gene lists with DAVID. Curr Protoc Bioinformatics. 2009;Chapter 13: Unit 13.11. doi:10.1002/0471250953.bi1311s27

23. Darde TA, Sallou O, Becker E, Evrard B, Monjeaud C, Le Bras Y, et al. The ReproGenomics Viewer: an integrative cross-species toolbox for the reproductive science community. Nucleic Acids Res. 2015;43: W109-116. doi:10.1093/nar/gkv345

24. Darde TA, Lecluze E, Lardenois A, Stévant I, Alary N, Tüttelmann F, et al. The ReproGenomics Viewer: a multi-omics and cross-species resource compatible with single-cell studies for the reproductive science community. Bioinformatics. 2019;35: 3133–3139. doi:10.1093/bioinformatics/btz047

25. Toyooka Y, Tsunekawa N, Takahashi Y, Matsui Y, Satoh M, Noce T. Expression and intracellular localization of mouse Vasa-homologue protein during germ cell development. Mech Dev. 2000;93: 139–149. doi:10.1016/s0925-4773(00)00283-5

26. Han J, Zhang J, Chen L, Shen B, Zhou J, Hu B, et al. Efficient in vivo deletion of a large imprinted lncRNA by CRISPR/Cas9. RNA Biol. 2014;11: 829–835. doi:10.4161/rna.29624

27. Johnson WH. The significance to bull fertility of morphologically abnormal sperm. Vet Clin North Am Food Anim Pract. 1997;13: 255–270. doi:10.1016/s0749-0720(15)30339-x

28. Vedrenne C, Hauri H-P. Morphogenesis of the endoplasmic reticulum: beyond active membrane expansion. Traffic. 2006;7: 639–646. doi:10.1111/j.1600-0854.2006.00419.x

29. Handel MA, Eppig JJ. Sexual dimorphism in the regulation of mammalian meiosis. Curr Top Dev Biol. 1998;37: 333–358. doi:10.1016/s0070-2153(08)60179-9

30. Morelli MA, Cohen PE. Not all germ cells are created equal: aspects of sexual dimorphism in mammalian meiosis. Reproduction. 2005;130: 761–781. doi:10.1530/rep.1.00865

31. Su Y, Li Y, Ye P. Mammalian meiosis is more conserved by sex than by species: conserved co-expression networks of meiotic prophase. Reproduction. 2011;142: 675–687. doi:10.1530/REP-11-0260

32. Handel MA, Schimenti JC. Genetics of mammalian meiosis: regulation, dynamics and impact on fertility. Nat Rev Genet. 2010;11: 124–136. doi:10.1038/nrg2723

33. Yuan L, Liu JG, Zhao J, Brundell E, Daneholt B, Höög C. The murine SCP3 gene is required for synaptonemal complex assembly, chromosome synapsis, and male fertility. Mol Cell. 2000;5: 73–83. doi:10.1016/s1097-2765(00)80404-9

34. Pittman DL, Cobb J, Schimenti KJ, Wilson LA, Cooper DM, Brignull E, et al. Meiotic prophase arrest with failure of chromosome synapsis in mice deficient for Dmc1, a germline-specific RecA homolog. Mol Cell. 1998;1: 697–705. doi:10.1016/s1097-2765(00)80069-6

35. Baudat F, Manova K, Yuen JP, Jasin M, Keeney S. Chromosome synapsis defects and sexually dimorphic meiotic progression in mice lacking Spo11. Mol Cell. 2000;6: 989–998. doi:10.1016/s1097-2765(00)00098-8

36. Romanienko PJ, Camerini-Otero RD. The mouse Spo11 gene is required for meiotic chromosome synapsis. Mol Cell. 2000;6: 975–987. doi:10.1016/s1097-2765(00)00097-6

37. Reinholdt LG, Schimenti JC. Mei1 is epistatic to Dmc1 during mouse meiosis. Chromosoma. 2005;114: 127–134. doi:10.1007/s00412-005-0346-4

38. Kneitz B, Cohen PE, Avdievich E, Zhu L, Kane MF, Hou H, et al. MutS homolog 4 localization to meiotic chromosomes is required for chromosome pairing during meiosis in male and female mice. Genes Dev. 2000;14: 1085–1097.

39. Edelmann W, Cohen PE, Kneitz B, Winand N, Lia M, Heyer J, et al. Mammalian MutS homologue 5 is required for chromosome pairing in meiosis. Nat Genet. 1999;21: 123–127. doi:10.1038/5075

40. Li B, Nair M, Mackay DR, Bilanchone V, Hu M, Fallahi M, et al. Ovol1 regulates meiotic pachytene progression during spermatogenesis by repressing Id2 expression. Development. 2005;132: 1463–1473. doi:10.1242/dev.01658

41. Eaker S, Cobb J, Pyle A, Handel MA. Meiotic prophase abnormalities and metaphase cell death in MLH1-deficient mouse spermatocytes: insights into regulation of spermatogenic progress. Dev Biol. 2002;249: 85–95. doi:10.1006/dbio.2002.0708

42. Lipkin SM, Moens PB, Wang V, Lenzi M, Shanmugarajah D, Gilgeous A, et al. Meiotic arrest and aneuploidy in MLH3-deficient mice. Nat Genet. 2002;31: 385–390. doi:10.1038/ng931

43. Liu D, Matzuk MM, Sung WK, Guo Q, Wang P, Wolgemuth DJ. Cyclin A1 is required for meiosis in the male mouse. Nat Genet. 1998;20: 377–380. doi:10.1038/3855

44. Rey-Ares V, Rossi SP, Dietrich K-G, Köhn F-M, Schwarzer JU, Welter H, et al. Prostaglandin E2 (PGE2) is a testicular peritubular cell-derived factor involved in human testicular homeostasis. Mol Cell Endocrinol. 2018;473: 217–224. doi:10.1016/j.mce.2018.01.022

45. Tarulli GA, Stanton PG, Lerchl A, Meachem SJ. Adult sertoli cells are not terminally differentiated in the Djungarian hamster: effect of FSH on proliferation and junction protein organization. Biol Reprod. 2006;74: 798–806. doi:10.1095/biolreprod.105.050450

46. Oldknow KJ, Seebacher J, Goswami T, Villen J, Pitsillides AA, O'Shaughnessy PJ, et al. Follistatin-like 3 (FSTL3) mediated silencing of transforming growth factor β (TGFβ) signaling is essential for testicular aging and regulating testis size. Endocrinology. 2013;154: 1310–1320. doi:10.1210/en.2012-1886

47. Moniot B, Declosmenil F, Barrionuevo F, Scherer G, Aritake K, Malki S, et al. The PGD2 pathway, independently of FGF9, amplifies SOX9 activity in Sertoli cells during male sexual differentiation. Development. 2009;136: 1813–1821. doi:10.1242/dev.032631

48. Kurihara M, Otsuka K, Matsubara S, Shiraishi A, Satake H, Kimura AP. A Testis-Specific Long Non-Coding RNA, lncRNA-Tcam1, Regulates Immune-Related Genes in Mouse Male Germ Cells. Front Endocrinol (Lausanne). 2017;8: 299. doi:10.3389/fendo.2017.00299

49. Malcov M, Cesarkas K, Stelzer G, Shalom S, Dicken Y, Naor Y, et al. Aym1, a mouse meiotic gene identified by virtue of its ability to activate early meiotic genes in the yeast Saccharomyces cerevisiae. Dev Biol. 2004;276: 111–123. doi:10.1016/j.ydbio.2004.08.026

50. Abby E, Tourpin S, Ribeiro J, Daniel K, Messiaen S, Moison D, et al. Implementation of meiosis prophase I programme requires a conserved retinoid-independent stabilizer of meiotic transcripts. Nat Commun. 2016;7: 10324. doi:10.1038/ncomms10324

51. Bonavita R, Walas D, Brown AK, Luini A, Stephens DJ, Colanzi A. Cep126 is required for pericentriolar satellite localisation to the centrosome and for primary cilium formation. Biol Cell. 2014;106: 254–267. doi:10.1111/boc.201300087

52. Mönnich M, Borgeskov L, Breslin L, Jakobsen L, Rogowski M, Doganli C, et al. CEP128 Localizes to the Subdistal Appendages of the Mother Centriole and Regulates TGF-β/BMP Signaling at the Primary Cilium. Cell Rep. 2018;22: 2584–2592. doi:10.1016/j.celrep.2018.02.043

53. Kashihara H, Chiba S, Kanno S-I, Suzuki K, Yano T, Tsukita S. Cep128 associates with Odf2 to form the subdistal appendage of the centriole. Genes Cells. 2019;24: 231–243. doi:10.1111/gtc.12668

54. Marjanović M, Sánchez-Huertas C, Terré B, Gómez R, Scheel JF, Pacheco S, et al. CEP63 deficiency promotes p53-dependent microcephaly and reveals a role for the centrosome in meiotic recombination. Nat Commun. 2015;6: 7676. doi:10.1038/ncomms8676

55. Miyata H, Morohoshi A, Ikawa M. Analysis of the sperm flagellar axoneme using gene-modified mice. Exp Anim. 2020;69: 374–381. doi:10.1538/expanim.20-0064

56. Mountain V, Compton DA. Dissecting the role of molecular motors in the mitotic spindle. Anat Rec. 2000;261: 14–24. doi:10.1002/(SICI)1097-0185(20000215)261:1<14::AID-AR5>3.0.CO;2-E

57. D'Andrea LD, Regan L. TPR proteins: the versatile helix. Trends Biochem Sci. 2003;28: 655–662. doi:10.1016/j.tibs.2003.10.007

58. Schmitz-Linneweber C, Small I. Pentatricopeptide repeat proteins: a socket set for organelle gene expression. Trends Plant Sci. 2008;13: 663–670. doi:10.1016/j.tplants.2008.10.001

59. Shamoto N, Narita K, Kubo T, Oda T, Takeda S. CFAP70 Is a Novel Axoneme-Binding Protein That Localizes at the Base of the Outer Dynein Arm and Regulates Ciliary Motility. Cells. 2018;7. doi:10.3390/cells7090124

60. Takaishi M, Huh N h. A tetratricopeptide repeat-containing protein gene, tpis, whose expression is induced with differentiation of spermatogenic cells. Biochem Biophys Res Commun. 1999;264: 81–85. doi:10.1006/bbrc.1999.1477

61. Liu W, He X, Yang S, Zouari R, Wang J, Wu H, et al. Bi-allelic Mutations in TTC21A Induce Asthenoteratospermia in Humans and Mice. Am J Hum Genet. 2019;104: 738–748. doi:10.1016/j.ajhg.2019.02.020

62. Karpenahalli MR, Lupas AN, Söding J. TPRpred: a tool for prediction of TPR-, PPR- and SEL1-like repeats from protein sequences. BMC Bioinformatics. 2007;8: 2. doi:10.1186/1471-2105-8-2

63. Chen Y, Zheng Y, Gao Y, Lin Z, Yang S, Wang T, et al. Single-cell RNA-seq uncovers dynamic processes and critical regulators in mouse spermatogenesis. Cell Res. 2018;28: 879–896. doi:10.1038/s41422-018-0074-y

64. Necsulea A, Soumillon M, Warnefors M, Liechti A, Daish T, Zeller U, et al. The evolution of lncRNA repertoires and expression patterns in tetrapods. Nature. 2014;505: 635–640. doi:10.1038/nature12943

65. Soumillon M, Necsulea A, Weier M, Brawand D, Zhang X, Gu H, et al. Cellular source and mechanisms of high transcriptome complexity in the mammalian testis. Cell Rep. 2013;3: 2179–2190. doi:10.1016/j.celrep.2013.05.031

66. Bolcun-Filas E, Schimenti JC. Genetics of meiosis and recombination in mice. Int Rev Cell Mol Biol. 2012;298: 179–227. doi:10.1016/B978-0-12-394309-5.00005-5

67. Baudat F, Imai Y, de Massy B. Meiotic recombination in mammals: localization and regulation. Nat Rev Genet. 2013;14: 794–806. doi:10.1038/nrg3573

68. Shibuya H, Watanabe Y. The meiosis-specific modification of mammalian telomeres. Cell Cycle. 2014;13: 2024–2028. doi:10.4161/cc.29350

69. Page J, de la Fuente R, Manterola M, Parra MT, Viera A, Berríos S, et al. Inactivation or non-reactivation: what accounts better for the silence of sex chromosomes during mammalian male meiosis? Chromosoma. 2012;121: 307–326. doi:10.1007/s00412-012-0364-y

70. Monesi V. RIBONUCLEIC ACID SYNTHESIS DURING MITOSIS AND MEIOSIS IN THE MOUSE TESTIS. J Cell Biol. 1964;22: 521–532. doi:10.1083/jcb.22.3.521

71. Chen Y, Zheng Y, Gao Y, Lin Z, Yang S, Wang T, et al. Single-cell RNA-seq uncovers dynamic processes and critical regulators in mouse spermatogenesis. Cell Res. 2018;28: 879–896. doi:10.1038/s41422-018-0074-y

72. Li K, Xu J, Luo Y, Zou D, Han R, Zhong S, et al. Panoramic transcriptome analysis and functional screening of long noncoding RNAs in mouse spermatogenesis. Genome Res. 2021;31: 13–26. doi:10.1101/gr.264333.120

73. Chen L-L. Linking Long Noncoding RNA Localization and Function. Trends Biochem Sci. 2016;41: 761–772. doi:10.1016/j.tibs.2016.07.003

74. Zhang B, Arun G, Mao YS, Lazar Z, Hung G, Bhattacharjee G, et al. The lncRNA Malat1 is dispensable for mouse development but its transcription plays a cis-regulatory role in the adult. Cell Rep. 2012;2: 111–123. doi:10.1016/j.celrep.2012.06.003

75. Sleutels F, Zwart R, Barlow DP. The non-coding Air RNA is required for silencing autosomal imprinted genes. Nature. 2002;415: 810–813. doi:10.1038/415810a

76. Chu C, Qu K, Zhong FL, Artandi SE, Chang HY. Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. Mol Cell. 2011;44: 667–678. doi:10.1016/j.molcel.2011.08.027

77. Cesana M, Cacchiarelli D, Legnini I, Santini T, Sthandier O, Chinappi M, et al. A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. Cell. 2011;147: 358–369. doi:10.1016/j.cell.2011.09.028

78. Tsai M-C, Manor O, Wan Y, Mosammaparast N, Wang JK, Lan F, et al. Long noncoding RNA as modular scaffold of histone modification complexes. Science. 2010;329: 689–693. doi:10.1126/science.1192002

79. Chalei V, Sansom SN, Kong L, Lee S, Montiel JF, Vance KW, et al. The long non-coding RNA Dali is an epigenetic regulator of neural differentiation. Elife. 2014;3: e04530. doi:10.7554/eLife.04530

80. Cai X, Cullen BR. The imprinted H19 noncoding RNA is a primary microRNA precursor. RNA. 2007;13: 313–316. doi:10.1261/rna.351707

81. Bai S, Fu K, Yin H, Cui Y, Yue Q, Li W, et al. Sox30 initiates transcription of haploid genes during late meiosis and spermiogenesis in mouse testes. Development. 2018;145. doi:10.1242/dev.164855

82. Wang F, Kong S, Hu X, Li X, Xu B, Yue Q, et al. Dnajb8, a target gene of SOX30, is dispensable for male fertility in mice. PeerJ. 2020;8: e10582. doi:10.7717/peerj.10582

83. Chotiner JY, Leu NA, Wang PJ. FLACC1 is testis-specific but dispensable for fertility in mice. Mol Reprod Dev. 2020;87: 1199–1201. doi:10.1002/mrd.23435

84. He X, Xie W, Li H, Cui Y, Wang Y, Guo X, et al. The testis-specifically expressed gene Trim69 is not essential for fertility in mice. J Biomed Res. 2020;35: 47–60. doi:10.7555/JBR.34.20200069

85. Jamin SP, Petit FG, Demini L, Primig M. Tex55 encodes a conserved putative A-kinase anchoring protein dispensable for male fertility in the mouse. Biol Reprod. 2021. doi:10.1093/biolre/ioab007

86. Khan R, Ye J, Yousaf A, Shah W, Aftab A, Shah B, et al. Evolutionarily conserved and testis-specific gene, 4930524B15Rik, is not essential for mouse spermatogenesis and fertility. Mol Biol Rep. 2020;47: 5207–5213. doi:10.1007/s11033-020-05595-0

87. Iwamori N, Zhao M, Meistrich ML, Matzuk MM. The testis-enriched histone demethylase, KDM4D, regulates methylation of histone H3 lysine 9 during spermatogenesis in the mouse but is dispensable for fertility. Biol Reprod. 2011;84: 1225–1234. doi:10.1095/biolreprod.110.088955

88. Khan M, Jabeen N, Khan T, Hussain HMJ, Ali A, Khan R, et al. The evolutionarily conserved genes: Tex37, Ccdc73, Prss55 and Nxt2 are dispensable for fertility in mice. Sci Rep. 2018;8: 4975. doi:10.1038/s41598-018-23176-x

89. Lu Y, Oura S, Matsumura T, Oji A, Sakurai N, Fujihara Y, et al. CRISPR/Cas9-mediated genome editing reveals 30 testis-enriched genes dispensable for male fertility in mice†. Biol Reprod. 2019;101: 501–511. doi:10.1093/biolre/ioz103

90. Miyata H, Castaneda JM, Fujihara Y, Yu Z, Archambeault DR, Isotani A, et al. Genome engineering uncovers 54 evolutionarily conserved and testis-enriched genes that are not

required for male fertility in mice. Proc Natl Acad Sci U S A. 2016;113: 7704–7710. doi:10.1073/pnas.1608458113

91. Li C, Shen C, Shang X, Tang L, Xiong W, Ge H, et al. Two novel testis-specific long noncoding RNAs produced by *1700121C10Rik* are dispensable for male fertility in mice. J Reprod Dev. 2020;66: 57–65. doi:10.1262/jrd.2019-104

92. Zhu Y, Lin Y, He Y, Wang H, Chen S, Li Z, et al. Deletion of lncRNA5512 has no effect on spermatogenesis and reproduction in mice. Reprod Fertil Dev. 2020;32: 706–713. doi:10.1071/RD19246

93. Wichman L, Somasundaram S, Breindel C, Valerio DM, McCarrey JR, Hodges CA, et al. Dynamic expression of long noncoding RNAs reveals their potential roles in spermatogenesis and fertility. Biol Reprod. 2017;97: 313–323. doi:10.1093/biolre/iox084

94. Miyata H, Shimada K, Morohoshi A, Oura S, Matsumura T, Xu Z, et al. Testis-enriched kinesin KIF9 is important for progressive motility in mouse spermatozoa. FASEB J. 2020;34: 5389–5400. doi:10.1096/fj.201902755R

95. Sutton KA, Jungnickel MK, Florman HM. A polycystin-1 controls postcopulatory reproductive selection in mice. Proc Natl Acad Sci U S A. 2008;105: 8661–8666. doi:10.1073/pnas.0800603105

96. Henao-Mejia J, Williams A, Rongvaux A, Stein J, Hughes C, Flavell RA. Generation of Genetically Modified Mice Using the CRISPR-Cas9 Genome-Editing System. Cold Spring Harb Protoc. 2016;2016: pdb.prot090704. doi:10.1101/pdb.prot090704

97. Hellemans J, Mortier G, De Paepe A, Speleman F, Vandesompele J. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. Genome Biol. 2007;8: R19. doi:10.1186/gb-2007-8-2-r19

98. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics. 2009;25: 1105–1111. doi:10.1093/bioinformatics/btp120

99. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014;30: 923–930. doi:10.1093/bioinformatics/btt656

100. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15: 550. doi:10.1186/s13059-014-0550-8

101. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B (Methodological). 1995;57: 289–300.

Supporting information

S1 Fig. Validation of several DEGs by RT-qPCR (RNA-seq *Topaz1*$^{-/-}$ *vs* WT testes).

Validation of several differentially expressed up- or down-regulated genes and of non-DEGs of RNA-seq analysis by qRT-PCR from P16 (A) or P18 (B) mouse testis RNAs. The lines represent the median of each genotype (blue: WT; red: *Topaz1*$^{-/-}$). A statistical test of Kruskal-Wallis was carried out (*p<0.05).

S2 Fig. Abnormal centrosome labelling in Topaz1-deficient gonads.

Immunofluorescence staining for γ-TUBULIN (red) and DAPI (blue) in WT (left) and *Topaz1*$^{-/-}$ (right) 30 d*pp* testis sections. Unlike the two red dots locating centrosomes in the meiotic metaphases seen in normal testes (left), centrosomes are abnormal in *Topaz1*$^{-/-}$ mutants (right) with one diffuse labelling. Zooms in white squares showed spermatocytes in metaphase I (WT) or in metaphase I-like (*Topaz1*$^{-/-}$). Scale bar = 50µm

S3 Fig. Reprogenomics data on dynamic expression of *4930463O16Rik*.

Dynamic expression of *4930463O16Rik* in five different tissues in male and female adult mice (A), in embryo primordial germ cells and adult male germ cells (B). *4930463O16Rik* is expressed in testis in germ cell during post-natal life. The strongest dynamic expression is found in pachytene spermatocytes.

S4 Fig. Reprogenomics data on dynamic expression of *Gm21269*.

Dynamic expression of *Gm21269* in five different tissues in male and female adult mice (A), in embryo primordial germ cells and adult male germ cells (B). *Gm21269* is expressed in testis in germ cell during post-natal life. The strongest dynamic expression is found in pachytene spermatocytes.

S5 Fig. IGV representation of P18-testis RNA-seq.

Expression of *4930463O16Rik* (A), *Gm21269* (B) and *4921513H07Rik* (C) from BigWig files of strand-specific RNA-seq data. The first four tracks represent transcripts of WT testes at P18; the next three tracks represent transcripts of *Topaz1*$^{-/-}$ testis at the same developmental stage.Representations of the

genes (from mm10 or GRCm38) are at the bottom of each graph. At the top of each one, there is the representation of the size of *4930463O16Rik* (A), *Gm21269* (B) and *492151H07Rik* (C) transcripts (red) from Ensembl data (GRCm38). *4930463O16Rik* and *4921513H07Rik* gene transcriptions overlap in 3' or 5' respectively.

S6 Fig. Expression of *Gm21269*, *4930463O16Rik* and *492151H07Rik* mRNAs in testes from 5 days to adulthood.

Quantitative RT-PCR analysis of *Gm21269*, *4930463O16Rik* and *492151H07Rik* gene expressions at different developmental stages in WT (blue) and *Topaz1*$^{-/-}$ (red) testes. The lines represent the median of each genotype. A statistical test of Kruskal-Wallis was carried out (*p<0.05; **p<0.01).

S7 Fig. *IS*H with PAS counterstained in WT mouse testis.

Visualisation of 493O16Rik (A), *Gm21269* (B) and *4921513H07Rik* (C) mRNAs respectively by *IS*H at different seminiferous epithelium stages highlighted by PAS staining. Scale bar = 20µm

S8 Fig. LncRNA cellular localizations on WT two month-old mouse testes.

*IS*H using (A) *4930463O16Rik*, (D) *Gm21269* and (G) *4921513H07Rik* probes (red). (B-E-H) Immunofluorescence staining with γH2Ax antibody was achieved in the same stage of seminiferous epithelium to identify male germ cells (green). (C-F-I) DAPI (blue), visualizing nuclear chromosomes, was merge with *IS*H (green) and IF (red) signals. Zooms in white squares showed spermatocytes during prophase I. No colocation between the sex body (γH2Ax) and lncRNAs (red) was evident. Scale bar = 20 µm.

S9 Fig. Validation of several DEGs by RT-qPCR (RNA-seq *4930463O16Rik*$^{-/-}$ *vs* WT testes).

Validation of several differentially expressed up- or down-regulated genes and of non-DEGs of RNA-seq analysis by RT-qPCR from P16 (A) or P18 (B) mouse testis RNAs. The lines represent the median of

each genotype (blue: WT; red: 4930463O16Rik$^{-/-}$). A statistical test of Kruskal-Wallis was carried out (*p<0.05).

S1 Table. List of DEGs in *Topaz1*$^{-/-}$ testis compared to WT.

List of deregulated genes in *Topaz1* KO testes at P16 (sheet 1) and P18 (sheet 2) (adjusted p-value < 0.05 and absolute Log2FC>1).

S2 Table. Functional annotation of P16 DEGs (RNA-seq *Topaz1*$^{-/-}$ *vs* WT testes).

DAVID functional Annotation Clustering (DAVID 6.8) analysis (based on GO terms and KEGG pathway) of all P16-differentially expressed genes (sheet 1) or only up-regulated (sheet 2) or down-regulated DEGs (sheet 3) in *Topaz1*$^{-/-}$ testis.

S3 Table. Functional annotation of P18 DEGs (RNA-seq *Topaz1*$^{-/-}$ *vs* WT testes).

DAVID functional Annotation Clustering (DAVID 6.8) analysis (based on GO terms and KEGG pathway) of P18-differentially expressed genes (sheet 1) or only up-regulated (sheet 2) or down-regulated DEGs (sheet 3) in *Topaz1*$^{-/-}$ testis. Annotation clusters based on InterPro database of P18-down-regulated DEGs have been mentioned in sheet 4.

S4 Table. List of primers.

List of different primers used in this study for genotyping, RT-qPCR and gRNAs).

S5 Table. List of DEGs in *4930463O16Rik*$^{-/-}$ testis compared to WT.

List of deregulated genes in *4930463O16Rik* KO testes at P16 (sheet 1) and P18 (sheet 2) (adjusted p-value < 0.05 and absolute Log2FC>1).

S6 Table. Functional annotation of P16 DEGs (RNA-seq *4930463O16Rik*$^{-/-}$ *vs* WT testes).

1124 DAVID functional Annotation Clustering (DAVID 6.8) analysis of P18-differentially expressed genes

1125 (sheet 1) or only up-regulated (sheet 2) or down-regulated DEGs (sheet 3) in *4930463O16Rik*$^{-/-}$ testis.

1126

1127 **S7 Table. Casa system settings.**

1128