SI Appendix

ALKBH5-dependent m6A demethylation controls splicing and stability of long 3'UTR mRNAs in male germ cells

Chong Tang[a,1], Rachel Klukovich[a,1], Hongying Peng[a], Zhuqing Wang[a], Tian Yu[a], Ying Zhang[a], Huili Zheng[a], Arne Klungland[b,c], and Wei Yan[a,d,2]

[a]Department of Physiology and Cell Biology, University of Nevada, Reno School of Medicine, 1664 North Virginia Street, MS575, Reno, NV 89557, USA; [b]Department of Microbiology, Oslo University Hospital, Rikshospitalet, 0027 Oslo, Norway; [c]Department of Molecular Medicine, Institute of Basic Medical Sciences, University of Oslo, 0317 Oslo, Norway; [d]Department of Biology, University of Nevada, Reno, 1664 North Virginia Street, MS575, Reno, NV 89557, USA

[1]C.T. and R.K. contributed equally to this work.

[2]To whom correspondence should be addressed:

Wei Yan M.D., Ph.D.
University of Nevada Reno Foundation Professor
Department of Physiology and Cell Biology
University of Nevada School of Medicine
Center for Molecular Medicine, Room 207B
1664 North Virginia Street, MS/0575
Reno, NV 89557
Tel: 775 784 7765
Fax: 775 784 4362
Email: wyan@med.unr.edu

This file contains SI Materials and Methods, Fig. S1-S10, and Table S1-S3.

Materials and Methods

Animal use and care. All mice used in this study were on the C57BL/6J background, and housed under specific pathogen-free conditions in a temperature- and humidity-controlled animal facility at the University of Nevada, Reno. Animal use protocol was approved by Institutional Animal Care and Use Committee (IACUC) of the University of Nevada, Reno (Protocol number 00494), and are in accordance with the "Guide for the Care and Use of Experimental Animals" established by National Institutes of Health (NIH) (1996, revised 2011). The male *Alkbh5* KO mice used in this study were described previously (1).

Histology. Testes were dissected from WT or KO mouse and were fixed in Bouin's solution for 24 hours, then washed and submerged in 70% ethanol until processed and paraffin-embedded using a Leica tissue processor. Sections were cut at 5μm thick and were attached to positively-charged slides (Genesee Scientific No. 29-107). Testes sections were stained with hematoxylin and eosin and staining was performed in biological duplicates or triplicates. Cauda epididymal sperm from WT or KO were spread onto positively-charged slides (Genesee Scientific No. 29-107), fixed for 5 minutes with 4% PFA, and were stained with hematoxylin and eosin to examine morphology. Images were taken using a Keyence microscope model BZX-710.

PAS staining. Periodic acid-Schiff (PAS) staining was performed on 5 μm thick paraffin sections. After deparaffinization, slides were incubated in 0.5% periodic acid for 10 min, followed by rinsing in distilled water three times. Slides were then submerged in Schiff's reagent for 15 minutes followed by rinsing with distilled water for 10 minutes. Hematoxylin was used to counterstain the slides which were then dipped in 11% total volume HCL in 100% Ethanol until color was not rinsing off. Slides were dehydrated then mounted with Permount. Images were taken using a Keyence microscope model BzX-710. Staining was done in biological triplicates.

TUNEL analyses. Terminal deoxynucleotidyl transferase dUTP nick end-labeling (TUNEL) staining was performed according to the manufacturer's instructions with some modifications (Trevigen *in situ* apoptosis detection kit No. 4810-30-CK). Paraffin-embedded testis sections were 5 μm thick. Proteinase K digestion was incubated at 37°C for 20 minutes and manganese was used as the cation for the TdT reaction. Slides were counterstained with hematoxylin for 15 seconds. Staining and counts were carried out in biological triplicates and the number of TUNEL-positive cells were counted from 100 randomly selected tubules. A student's two-tailed t-test was used to determine significance ($p < 0.05$).

Gross morphology and sperm analysis. Testis and the whole epididymis were dissected from WT and KO mice and rinsed in PBS. After weight measurements and morphology pictures were taken, both cauda epididymides were punctured several times with a needle to release the sperm into HTF medium at 37°C for 30 minutes. Computer-assisted sperm analysis (CASA) was used to analyze motility, performed in

triplicate. A hemocytometer was used to manually count sperm concentration, measured in triplicate.

Immunofluorescence staining and confocal microscopy. After dissection, testes were fixed in 4% paraformaldehyde for 1 hour at room temperature, poked several times with a 20.5 G needle, fixed 2 more hours at room temperature, then transferred to 4°C for 21 more hours. Testes were then washed 3x in 5% sucrose then dehydrated in a sucrose gradient of 5%:20% sucrose at 2:1, 1:1, 1:2 until testes sunk to the bottom of the tube. Testes were then kept in 20% sucrose for 48 hours and were perfused for 1 hour in 1:1 20% sucrose:OCT. Testes were embedded in a Tissue-Tek Cryomold in 1:1 20% sucrose:OCT and rapidly frozen using liquid nitrogen. Blocks were cut into 10µm thick sections on Fisher SuperFrost Plus Slides (No. 22-037-246), were allowed to dry for 30 minutes, and were stored at -80°C until use. Slides were permeablized by boiling in citrate buffer (pH 6.0) 5 minutes, 3 times and were blocked in 1% BSA in TBS supplemented with 5% fetal bovine serum and 5% normal goat serum at room temperature for 1 hour. Primary antibodies were diluted in 1% BSA in TBS: rabbit anti-alkbh5 (Sigma No. HPA007196, 1:150), mouse anti-GM130 (BD Biosciences No. 610822, 1:200), mouse anti-SC35 (No. ab11826, 1:500) and were incubated overnight at 4°C. After washing three times in PBS, the secondary antibody was diluted in 1% BSA in TBS: goat anti-rabbit Alexa Fluor 488 (Thermo Fisher No. A11034, 1:500) and goat anti-mouse Alexa Fluor 594 (Abcam No. ab150116, 1:400) and was incubated for 1 hour at room temperature. Slides were then washed three times in PBS, followed by submersion in 0.1% Sudan Black B in 70% ethanol. After washing three times in PBS, slides were mounted using DAPI (Abcam No. ab104139). An Olympus FV1000 confocal microscope was used to take the images of z-stacks 5µm thick.

Western blots. Total protein lysates were extracted from whole testis in phosphate-buffered saline. The Pierce BCA Protein Assay Kit (Thermo Fisher, Cat#23225) was used to determine concentration. Protein (25 µg) was loaded and run on a 4-15% Mini-PROTEAN TGX Stain-free gel (Bio-Rad No. 456-8084), transferred onto a nitrocellulose membrane (0.45µm, Amersham, Cat#10600003) and blocked with Thermo Fisher Superblock for 1 hour. Primary antibodies: rabbit anti-Alkbh5 (Sigma HPA007196, 1:500) and mouse anti-β-actin (Abcam, Cat#ab8226, 1:5000) were incubated overnight at 4°C. Secondary antibodies: goat anti-rabbit HRP (Abcam, Cat#ab205718) and goat anti-mouse HRP (Southern Biotech, Cat#1030-05) were diluted 1:10,000, the membrane was washed, and StrepTactin-HRP (Bio-Rad, Cat#161-0376) was used to visualize the standards. Advanta Western Bright ECL Kit (Cat#K-12045-D20) was used for detection on the Bio-Rad ChemiDoc imager.

Purification of spermatogenic cells. Pachytene spermatocytes, round and elongating/elongated spermatids were purified from adult mouse testes using the STA-PUT method (2). The BSA gradients (0.5-4%) were prepared in the EKRB buffer (Cat#K-4002, Sigma), supplemented with sodium bicarbonate (1.26g per 1L), L-glutamine (0.29228g per 1L); Penicillin and Streptomycin mix (Thermo-Fisher, 10,000U per 1L), MEM non-essential amino acids (Thermo-Fisher, 1ml 100X per 1L), MEM amino acids (20ml 50X per 1L) and cycloheximide (100ng/ml), pH7.2-7.3). Eight testes

were pooled each time for cell purification. After being removed and decapsulated, testes were placed into 10ml of the EKRB buffer containing 5mg collagenase (Sigma) for a 12-min digestion at 32°C to disperse the testicular cells. Once dispersed, the testicular cells were washed three times using the EKRB buffer followed by trypsin digestion by incubation in 10ml EKR buffer containing trypsin (Sigma, 0.25mg/ml) and DNase I (Sigma, 20mg/ml) at 37°C for 12min with occasional pipetting to facilitate cell dispersion. Fully dispersed testicular cells were washed three times followed by centrifugation and re-suspension in 10 ml of 0.5% BSA. The cell suspension was passed through 50μm filters, and the filtrate was saved for loading onto the STA-PUT apparatus for sedimentation. After 3h sedimentation at 4°C, fractions were collected from the bottom of the sedimentation chamber. A total of 30 fractions of 15 ml each were collected. After centrifugation, the supernatants were removed, and the cells in each fraction were re-suspended, and the cell purity was determined by microscopy examination based on cell morphology, as described previously (2). Fractions containing the same cell types were pooled followed by centrifugation to collect purified pachytene spermatocytes, round spermatids, and elongating/elongated spermatids. The RNA was extracted from cells by mirVana miRNA Isolation Kit (Life Technology, Cat# AM1560).

RNA extraction. RNA was extracted from HEK293 cells using the mirVana miRNA Isolation Kit (ThermoFisher, Cat#AM1560), according to the manufacturer's instructions. Extracted RNA quantification was done using the Qubit RNA High Sensitivity Assay Kit (Invitrogen No. Q32855) measured on the Qubit 2.0 Fluorometer (Invitrogen).

m6A RNA immunoprecipitation. Protein G Dynabeads (50μL, Invitrogen, Cat#10004D) were washed three times and resuspended with immunoprecipitation buffer (5mL 0.1M Sodium phosphate, 3.5mL 2M sodium chloride, 250μL 10% triton, nuclease-free water to 50mL final volume). Rabbit anti-m^6A antibody (Abcam, Cat#ab151230) or normal rabbit IgG (Invitrogen, Cat#10500C) (6μg each) were added to the beads on a rocker for 1 hour at room temperature, protected from light. The beads were washed and resuspended in immunoprecipitation buffer. 200ng of RNA were heat fragmented into 200nt fragments and were added to the bead suspension on a rocker for 2 hours at room temperature, protected from light. The beads were washed, pelleted, and eluted with sodium lauryl sulfate and tween-20 for 5 minutes at 70°C. The eluate was purified with 2.2x of magnetic beads made from SpeedBeads (Sigma-Aldrich Cat#65152105050250) and run on the Agilent RNA Pico chip (Cat#5067-1513) to verify purity and RNA integrity. All immunoprecipitations were performed in triplicate. The representative results of the procedures are shown in *SI Appendix, Fig. S8*.

RNA library construction. Immunoprecipitated RNA (1ng) and non-immunoprecipitated RNA (300ng) were constructed into next-generation sequencing libraries (Illumina) using the KAPA Stranded RNA-Seq Library Preparation Kit (KK8400) according to the manufacturer's instructions, with some modifications. Briefly, the fragmentation, priming, and elution steps were incubated at 65°C for 3 minutes, and the adaptors and barcodes were substituted with NEBNext Multiplex Oligos for Illumina

(Cat#E7335 and E7500). Library amplification was monitored using the qPCR with the KAPA real-time library amplification kit (Cat#KK2702). All libraries were constructed in triplicate. After final amplification, libraries were size selected using magnetic beads. The eluates were quantified using Qubit DNA High Sensitivity (Invitrogen, Cat#Q32854) and the Agilent DNA High Sensitivity chip (Cat#5067-4626) was used to determine library quality. Immunoprecipitated libraries were pooled according to library quantity and their negative control counterparts were pooled with equal volume to the immunoprecipitated library. Transcriptome libraries were pooled according to quantity. All libraries were sequenced at the Genomics Center of the University of Nevada, Reno on the Illumina Next-Seq 500 using paired-end 75 sequencing.

RNA-Seq data analyses. Quality control of RNA-Seq data are shown in *SI Appendix, Fig. S9*. Trimmomatic was used to remove adaptor sequences, and low-quality reads from the sequencing data (3). To identify all the transcripts, we used Tophat2 and Cufflinks to assemble the sequencing reads based on the UCSC MM9 mouse genome (4). The differential expression analysis was performed by Cuffdiff (4). The global statistics and quality controls are presented in Supplemental Figure S1. The UTR and alternative splicing analyses were performed using the SpliceR pipeline (5). The sequences without identified coding frames were extracted and subjected to coding potential calculating (6). The sequencing depth and alignment ratio were listed in *SI Appendix, Fig. S10*. We performed the data mining using our in-house R script.

Data normalization. FPKMs counts are scaled in Cuffdiff analyses *via* the median of the geometric means of fragment counts across all libraries, as described in (7). The principle was identical to the one used by DESeq (8).

m6a RIP-Seq data analyses. Trimmomatic was used to remove adaptor sequences and low-quality reads from the sequencing data (3). To reduce bias from potential inaccurate gene structure annotation, we aligned the m6A-seq reads to the assembled gene sequences derived from RNA-Seq cufflink results, using Tophat v2.0.14 (4). The longest isoform was used if the gene had multiple isoforms. The peak-calling method was modified from published work (9). To call m6A peaks, the longest isoform of each gene was scanned using a 100nt sliding window with 10nt step. We add one count to all the windows to avoid 0 count in the windows. The peak height threshold was ten counts. For each gene, the read counts in each window were normalized by the median counts of all windows of that gene. A Fisher exact test was used to identify the differential windows between IP and input samples. The window was called positive if the count>=10 and log2(enrichment score)>0.8, which FDR is close to zero. Overlapping positive windows were merged. The following four numbers were calculated to obtain the enrichment score of each window: (a) read counts of the IP samples in the current window, (b) median read counts of the IP sample in all 100nt windows on the current mRNA, (c) read counts of the input sample in the current window, and (d) median read counts of the input sample in all 100nt windows on the current mRNA. The enrichment score of each window was calculated as (a*d)/(b*c). We used HEK293 cell m6A RIP-Seq data as the control to correlate with the 2,000bp genomic range of the published m6a nucleotide resolution site (10). The enrichment

score threshold was set as 0.8 to get the best results. If the high-confident peaks were identified in two of three, we then regarded it as a true m6A peak. We performed the data mining using our in-house R script.

Sequencing data deposition. Both RNA-Seq and m6A RIP-Seq datasets have been deposited into the Sequence Read Achieve (SRA) database in the NCBI with an accession number of PRJNA420607.

References

1. Zheng G*, et al.* (2013) ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol Cell* 49(1):18-29.
2. Zhang Y*, et al.* (2017) MicroRNAs control mRNA fate by compartmentalization based on 3' UTR length in male germ cells. *Genome Biol* 18(1):105.
3. Bolger AM, Lohse M, & Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30(15):2114-2120.
4. Trapnell C*, et al.* (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc* 7(3):562-578.
5. Vitting-Seerup K, Porse BT, Sandelin A, & Waage J (2014) spliceR: an R package for classification of alternative splicing and prediction of coding potential from RNA-seq data. *BMC Bioinformatics* 15:81.
6. Kong L*, et al.* (2007) CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Res* 35(Web Server issue):W345-349.
7. Anders S & Huber W (2010) Differential expression analysis for sequence count data. *Genome Biol* 11(10):R106.
8. Li P, Piao Y, Shon HS, & Ryu KH (2015) Comparing the normalization methods for the differential analysis of Illumina high-throughput RNA-Seq data. *BMC Bioinformatics* 16:347.
9. Zhao BS*, et al.* (2017) m6A-dependent maternal mRNA clearance facilitates zebrafish maternal-to-zygotic transition. *Nature* 542(7642):475-478.
10. Linder B*, et al.* (2015) Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. *Nat Methods* 12(8):767-772.

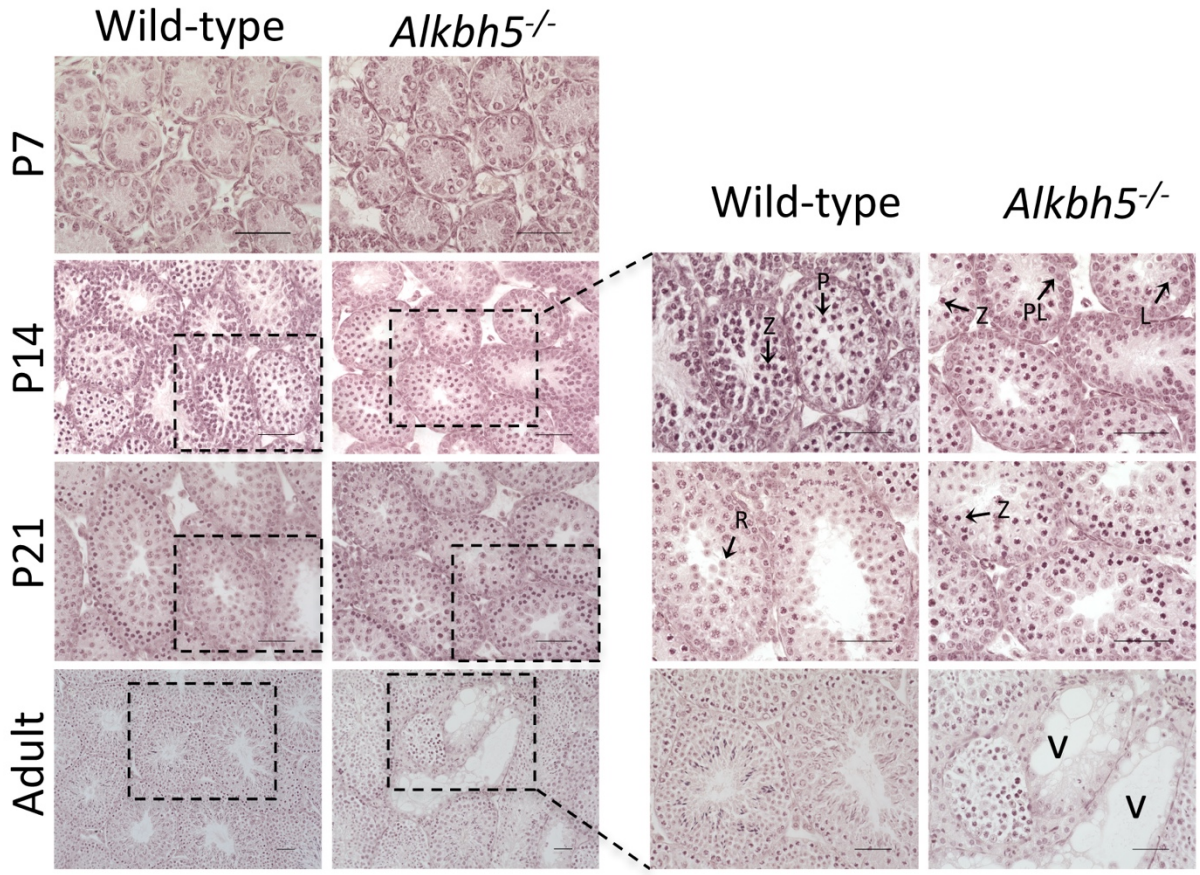

Fig. S1. Disrupted spermatogenesis in *Alkbh5* KO testes. Representative micrographs of hematoxylin and eosin (HE)-stained paraffin sections of wild-type (left column) and *Alkbh5*$^{-/-}$ (right) testis were indistinguishable at postnatal day 7 (P7). A slight delay in meiotic progression began at P14 with much fewer zygotene spermatocytes (Z) in *Alkbh5*$^{-/-}$ testes. At P21, round spermatids (R) already appeared in wild-type testes, but absent in *Alkbh5*$^{-/-}$ testes. P, pachytene spermatocytes; L, leptotene spermatocytes. Adult *Alkbh5*$^{-/-}$ testes lacked elongated spermatids and contained numerous vacuoles (v), indicative of massive germ cell depletion. Scale bars=50μm.

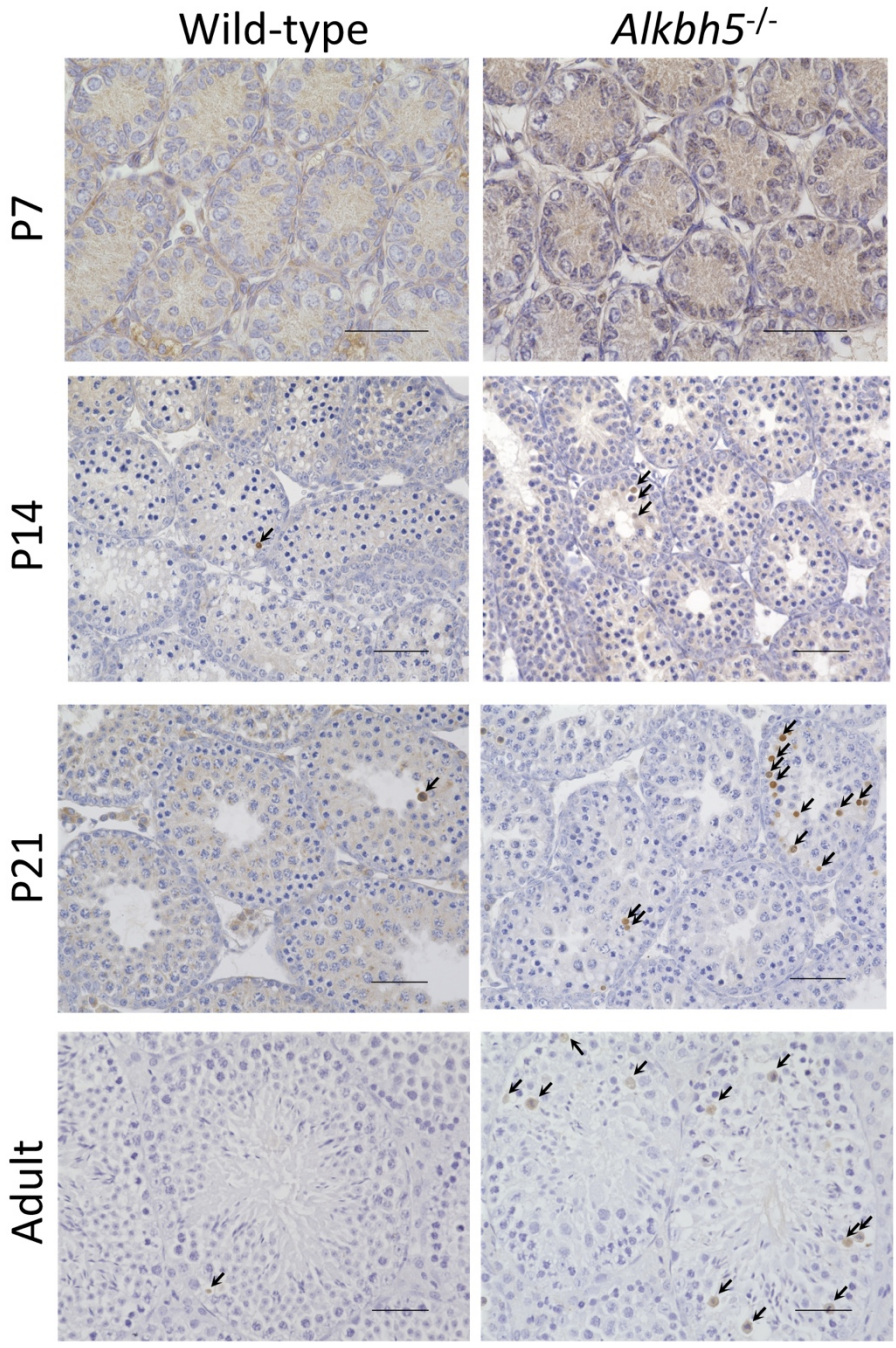

Fig. S2. Enhanced germ cell apoptosis in *Alkbh5*$^{-/-}$ testes. Terminal deoxynucleotidyl transferase dUTP nick end labeling (TUNEL)-staining followed by hematoxylin counterstaining was employed to label apoptotic cells. While no apoptosis was observed in either wild-type or *Alkbh5*$^{-/-}$ testes at postnatal day 7 (P7), the number of apoptotic germ cells was much greater in the *Alkbh5*$^{-/-}$ testes than that in WT controls. Experiments were conducted in triplicates. Scale bars=50μm.

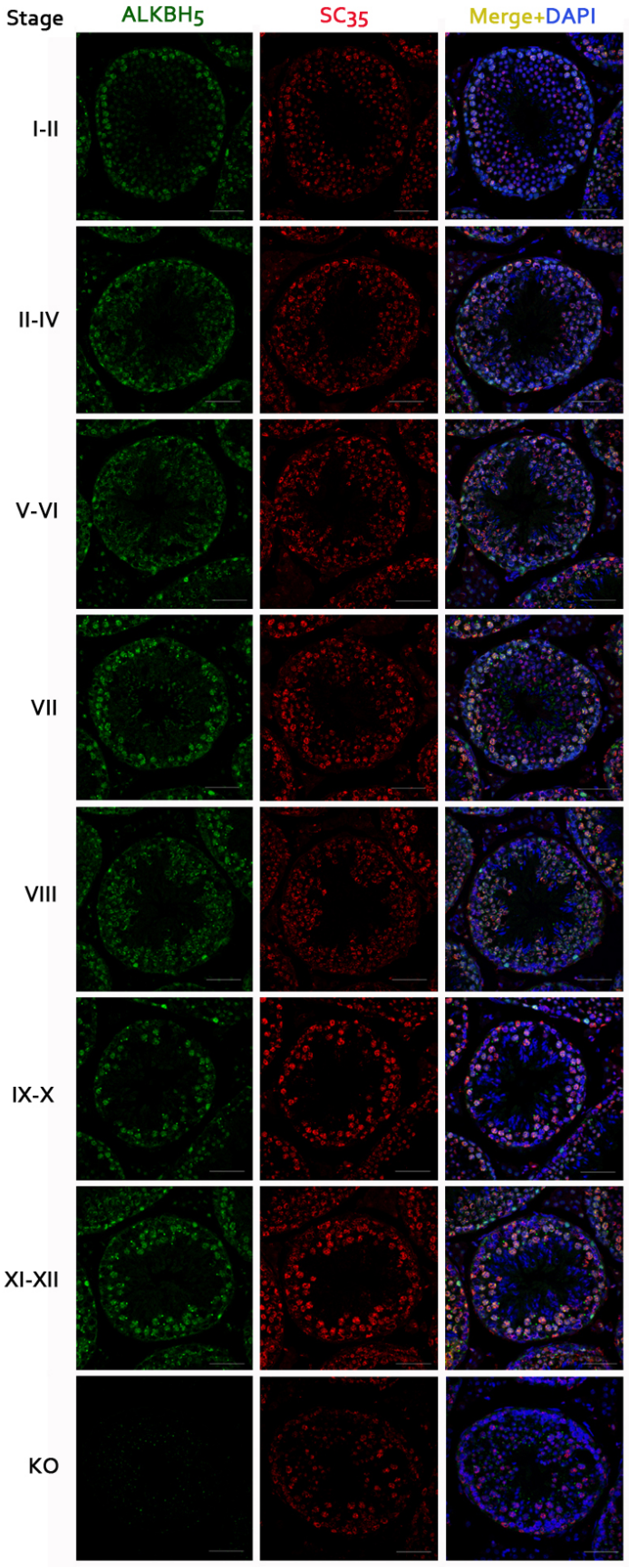

Fig. S3. Immunofluorescent staining of ALKBH5 (green) and SC35 (red) in WT and Alkbh5 KO testes. Representative images of twelve stages of the seminiferous epithelial cycle are shown. Scale bars = 50μm.

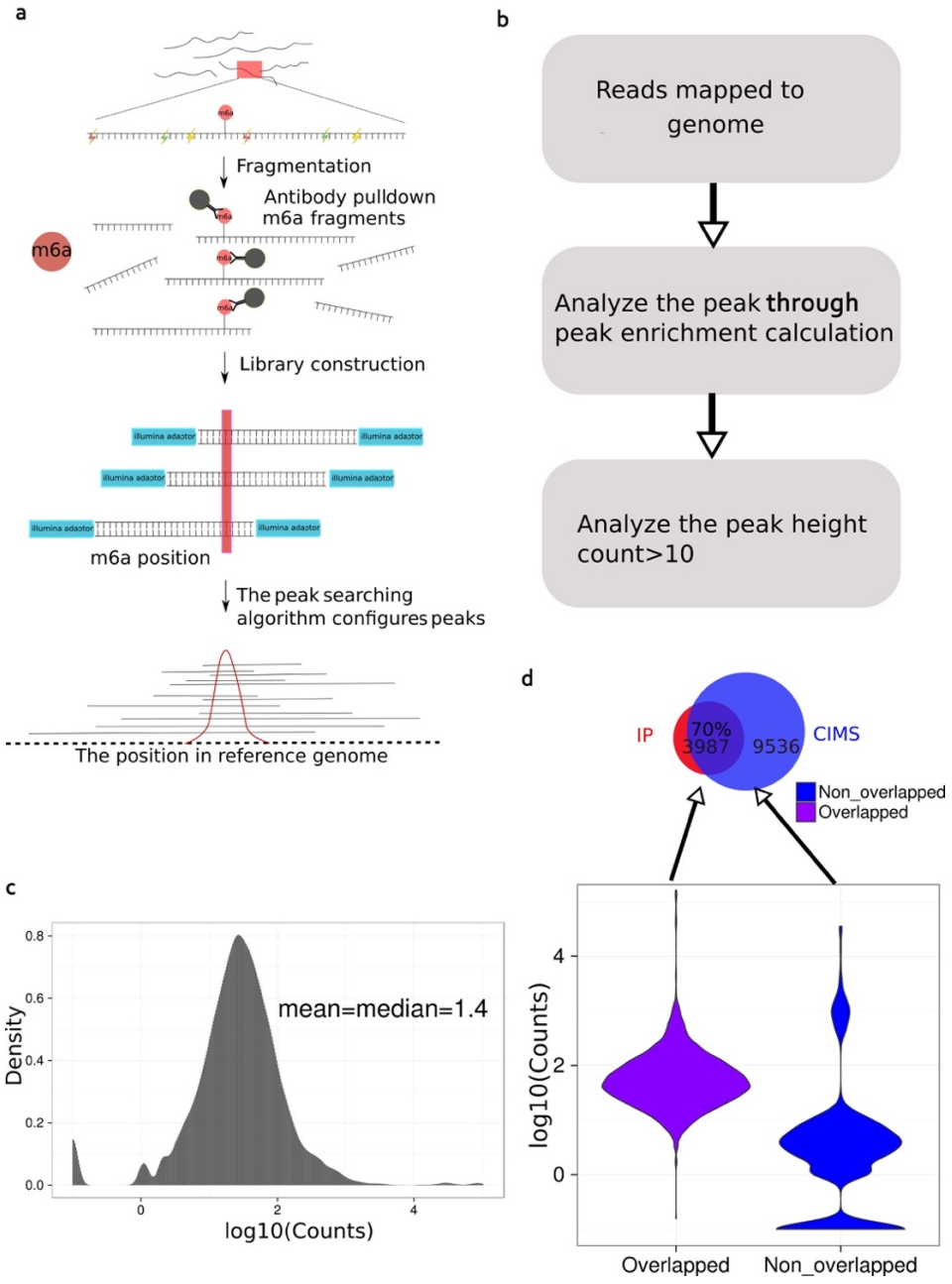

Fig. S4. Validation of the m6a RNA immunoprecipitation followed by deep sequencing (m6A-RIP-Seq) method using HEK293 cells. a. Total RNA was fragmented followed by IP using the m6a antibodies. Sequencing reads should peak around the m6a sites. b. Workflow for calculating the peak enrichment score and height. c. Density plot showing the distribution of m6A reads. d. After data training and proper parameter cutoff, our m6A-RIP-Seq method could detect ~70% of the m6a sites previously identified in HEK293 cells using a similar method [1, 2].

References

1. Dominissini D*, et al.* Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* **485**, 201-206 (2012).
2. Meyer KD, Saletore Y, Zumbo P, Elemento O, Mason CE, Jaffrey SR. Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell* **149**, 1635-1646 (2012).

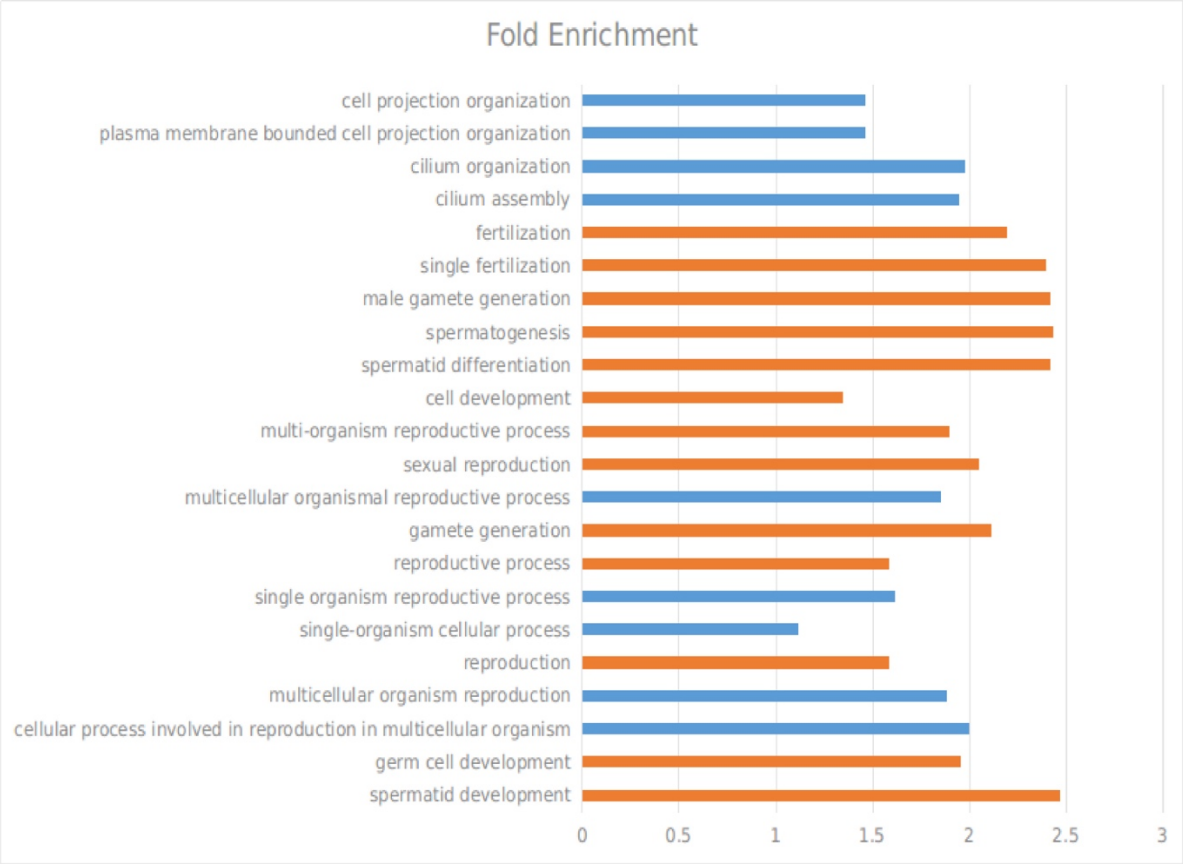

Fig. S5. Gene ontology analyses of 30 longer 3'UTR (>3,000nt) transcripts that are normally expressed in WT round spermatids, but downregulated in *Alkbh5* KO round spermatids. These mRNAs are mostly involved in the control of spermatogenesis and male fertility.

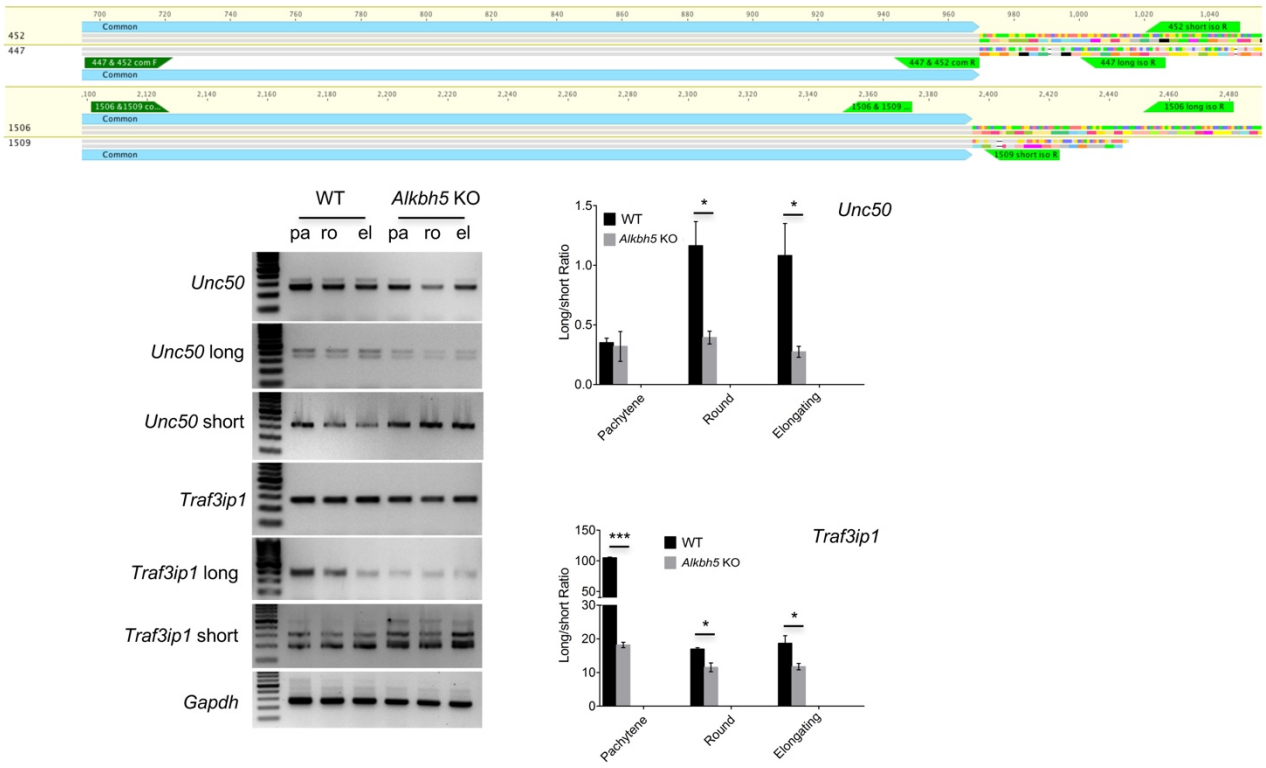

Fig. S6. qPCR analyses of expression levels of *Unc50* and *Traf3ip1* in WT and *Alkbh5* KO pachytene spermatocytes (pa), round (ro) and elongating (el) spermatids. Primer locations are shown in upper panels, and the primer sequences are listed in the table below. *Gapdh* was used as an internal control.

Primer	**Sequence (5'→3')**
Gapdh F	ACAGTCCATGCCATCACTGCC
Gapdh R	GCCTGCTTCACCACCTTCTTG
Unc50 (447 & 452 com) F	GGTGTTGGCCTTCTGATCTCAACGTTA
Unc50 (447 & 452 com) R	CACTGTATCCCAGGAAGGTCACATAG
Unc50 (447) short iso R	CATGAGAGGTGCGAATGGATAGAGCA
Unc50 (452) long iso R	CATATGAGTGAACCATAACACCTCAGGCA
Traf3ip1 (1506 &1509 com) F	CAGGAGCAGAGTATCACAGACAGTGC
Traf3ip1 (1506 & 1509 com) R	AGCAGGCCATTTCTCAGCAGACT
Traf3ip1 (1506) long iso R	CTGGAGGCTCTAAGTAGTAAGCACTGTAGC
Traf3ip1 (1509) short iso R	ACTGTCCTCTTCCCAACCGAGAC

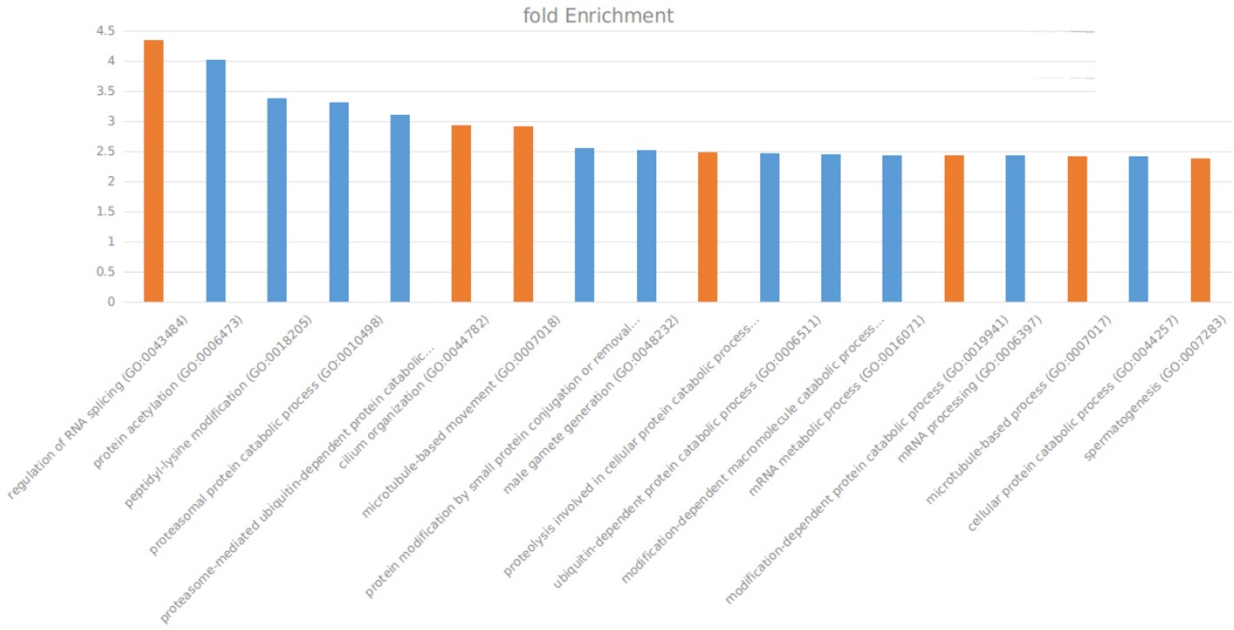

Fig. S7. Gene ontology analyses of the shorter transcripts that quickly degrade in *Alkbh5* KO elongating spermatids. Note that many of these short mRNAs are involved in the regulation of RNA splicing, protein and mRNA metabolism.

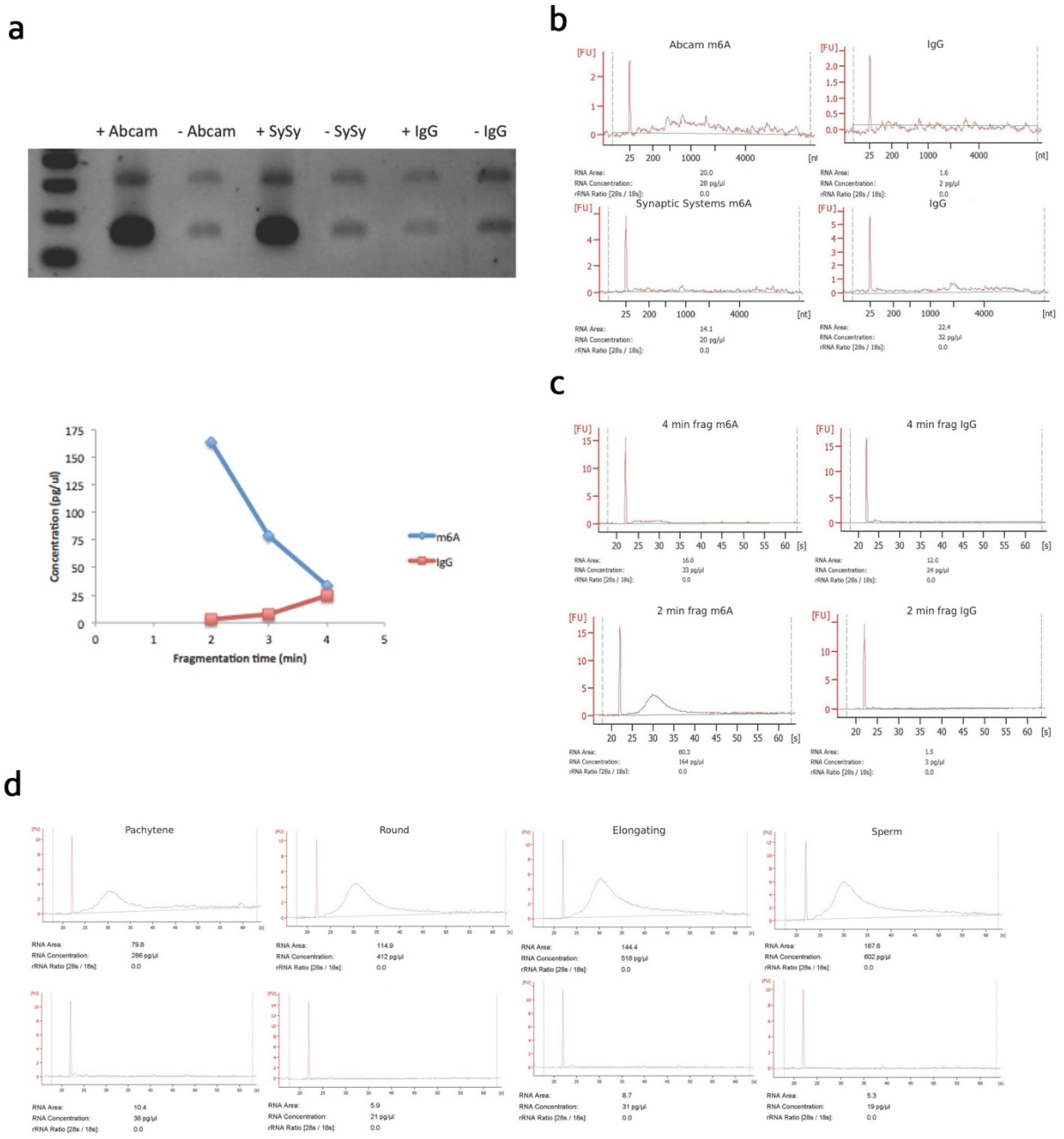

Fig. S8. Optimization of the m6A-RIP-Seq method. a. Validation of the anti-m^{6}A antibodies used. m^{6}A-positve and negative RNAs were synthesized (350nt), combined in a 1:1 ratio, immunoprecipitated using the anti-m^{6}A antibodies from either Abcam or Synaptic Systems, then digested with BamHI. m^{6}A-containing RNAs had a BamHI digestion site incorporated. IgG immunoprecipitation was used as a negative control. b. Agilent profiles of m^{6}A-immunoprecipitated total RNA using Abcam or Synaptic Systems antibodies, fragmented for 10 min. Normal rabbit IgG immunoprecipitation was used as a negative control. c, d. Optimization of RNA fragmentation times (2-4 min) followed by m^{6}A-immunoprecipitation shown as a graph (s) and Agilent profiles (d). e. Immunoprecipitation results from large RNA from spermatids fragmented for 2 minutes. Top row: m^{6}A. Bottom row: IgG.

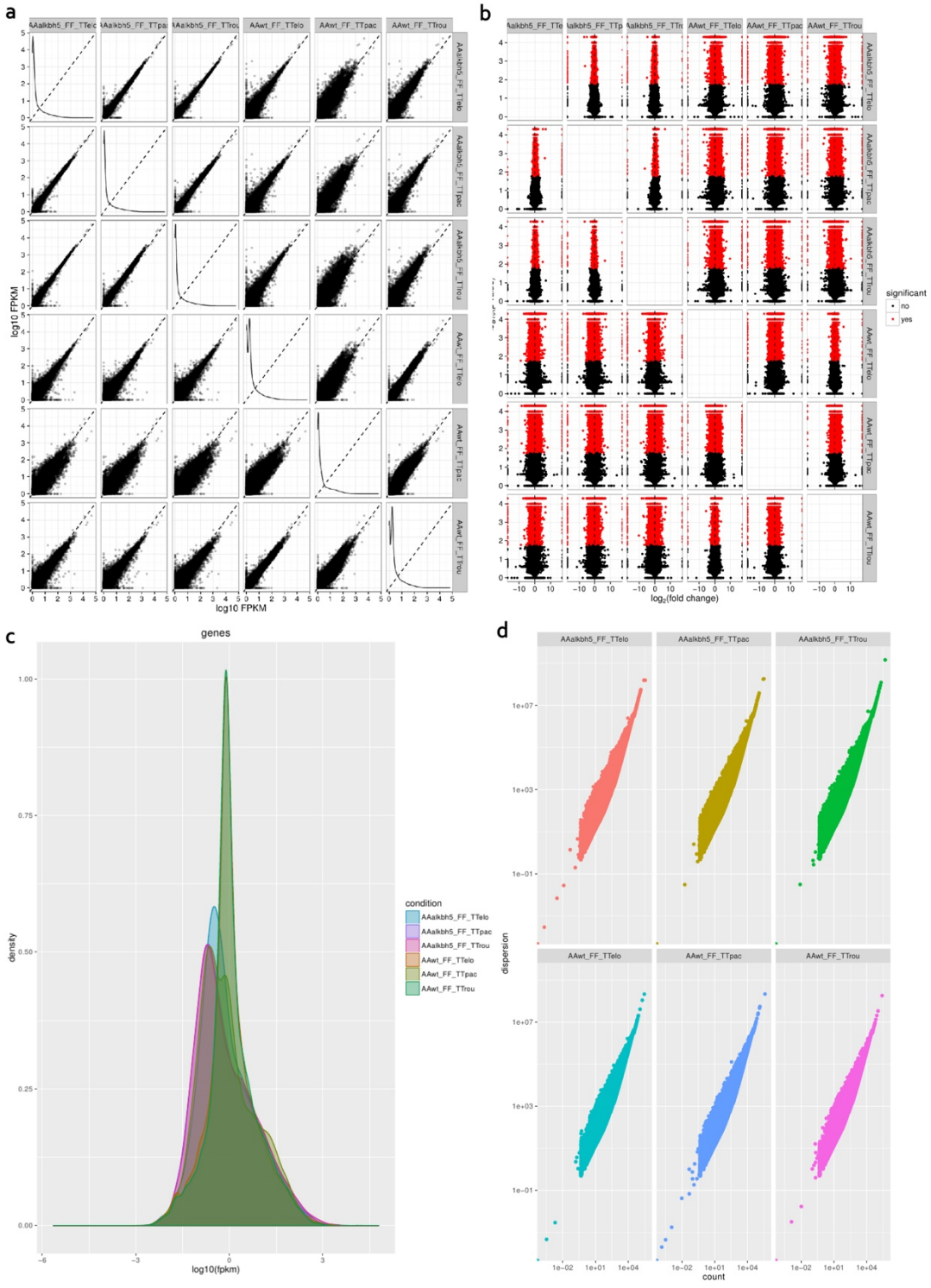

Fig. S9. Evaluation of the quality of the RNA-Seq data generated in this study. a. Pairwise scatterplots of the log10 normalized FKMP scores across all six RNA samples. b. Volcano plots showing the log10(p-value) in y-axes and log2(fold change) in x-axes by pairwise comparison across all six RNA samples. c. Density plots showing the distribution of log10 normalized FPKM scores across biological replicates of all six RNA samples. d. Overdispersion plots demonstrating the estimated overdispersion for each sample as a quality control measure.

Filename	Sequencing_depth	Mapping_ratio
AAwt-FF-TTrou-SS1	34735112	33442268 (96.3% of input)
AAalkbh5-FF-TTrou-SS1	38024748	36154779 (95.1% of input)
AAwt-FF-TTelo-SS2	32836366	31676981 (96.5% of input)
AAwt-FF-TTelo-SS1	36664543	35392210 (96.5% of input)
AAalkbh5-FF-TTpac-SS2	29271271	27864562 (95.2% of input)
AAalkbh5-FF-TTpac-SS1	34920033	33202198 (95.1% of input)
AAalkbh5-FF-TTelo-SS1	28998202	27622759 (95.3% of input)
AAwt-FF-TTpac-SS1	37619674	35967874 (95.6% of input)
AAwt-FF-TTrou-SS2	31757348	30593614 (96.3% of input)
AAwt-FF-TTpac-SS2	36056407	34486666 (95.6% of input)
AAalkbh5-FF-TTrou-SS2	31337681	29924225 (95.5% of input)
AAalkbh5-FF-TTelo-SS2	37342745	35523803 (95.1% of input)

Fig. S10. Sequencing depth and mapping rate of RNA-Seq analyses.

Table S1. Dysregulated splicing factor genes in Alkbh5 KO spermatogenic cells.

Gene symbol	Gene name	Function
Sfswap	Splicing factor, suppressor of white-apricot homolog;Sfswap;ortholog	mRNA splicing factor(PC00148)
U2af2	Splicing factor U2AF 65 kDa subunit;U2af2;ortholog	
Srsf1	Serine/arginine-rich splicing factor 1;Srsf1;ortholog	
Khdrbs3	KH domain-containing, RNA-binding, signal transduction-associated protein 3;Khdrbs3;ortholog	mRNA splicing factor(PC00148);transcription cofactor(PC00217)
Cdk12	Cyclin-dependent kinase 12;Cdk12;ortholog	non-receptor serine/threonine protein kinase(PC00167);non-receptor tyrosine protein kinase(PC00168)
Son	Protein SON;Son;ortholog	
Snrnp70	U1 small nuclear ribonucleoprotein 70 kDa;Snrnp70;ortholog	mRNA splicing factor(PC00148)
Brdt	Bromodomain testis-specific protein;Brdt;ortholog	
Hnrnpl	Heterogeneous nuclear ribonucleoprotein L;Hnrnpl;ortholog	
Celf5	CUGBP, Elav-like family member 5;Celf5;ortholog	
Tra2b	Transformer-2 protein homolog beta;Tra2b;ortholog	mRNA splicing factor(PC00148)
Srek1	Splicing regulatory glutamine/lysine-rich protein 1;Srek1;ortholog	RNA binding protein(PC00031)
Zfp326	DBIRD complex subunit ZNF326;Znf326;ortholog	
Sf3a1	Splicing factor 3A subunit 1;Sf3a1;ortholog	mRNA splicing factor(PC00148)
Hnrnpk	Heterogeneous nuclear ribonucleoprotein K;Hnrnpk;ortholog	enzyme modulator(PC00095);mRNA splicing factor(PC00148);ribonucleoprotein(PC00201);serine protease(PC00203)
Ddx17	Probable ATP-dependent RNA helicase DDX17;Ddx17;ortholog	
Tmbim6	Bax inhibitor 1;Tmbim6;ortholog	
Rbm25	RNA-binding protein 25;Rbm25;ortholog	

Table S2. Dysregulated cilliogenic genes in Alkbh5 KO spermatogenic cells.

Gene symbol	Gene name	Function
Lca5	Lebercilin;Lca5;ortholog	
C2cd3	C2 domain-containing protein 3;C2cd3;ortholog	
Foxj1	Forkhead box protein J1;Foxj1;ortholog	
Nek1	Serine/threonine-protein kinase Nek1;Nek1;ortholog	
Hydin	Hydrocephalus-inducing protein;Hydin;ortholog	
Dnaaf3	Dynein assembly factor 3, axonemal;Dnaaf3;ortholog	
Cfap126	Protein Flattop;Cfap126;ortholog	
Ttll8	Protein monoglycylase TTLL8;Ttll8;ortholog	
Ube2b	Ubiquitin-conjugating enzyme E2 B;Ube2b;ortholog	
Ttc21a	Tetratricopeptide repeat protein 21A;Ttc21a;ortholog	
Cep89	Centrosomal protein of 89 kDa;Cep89;ortholog	
Ttll3	Tubulin monoglycylase TTLL3;Ttll3;ortholog	
Dnah7a	Dynein, axonemal, heavy chain 7A;Dnah7a;ortholog	hydrolase(PC00121);microtubule binding motor protein(PC00156)
Traf3ip1	TRAF3-interacting protein 1;Traf3ip1;ortholog	
Rfx4	Transcription factor RFX4;Rfx4;ortholog	winged helix/forkhead transcription factor(PC00246)
Ttc26	Intraflagellar transport protein 56;Ttc26;ortholog	
Pcnt	Pericentrin;Pcnt;ortholog	
Ttc21b	Tetratricopeptide repeat protein 21B;Ttc21b;ortholog	
Wdr60	WD repeat-containing protein 60;Wdr60;ortholog	microtubule family cytoskeletal protein(PC00157)
Ift80	Intraflagellar transport protein 80 homolog;Ift80;ortholog	
Actr2	Actin-related protein 2;Actr2;ortholog	
Ccdc114	Coiled-coil domain-containing protein 114;Ccdc114;ortholog	
Ccdc63	Coiled-coil domain-containing protein 63;Ccdc63;ortholog	
Sclt1	Sodium channel and clathrin linker 1;Sclt1;ortholog	
Rabl2	Rab-like protein 2A;Rabl2;ortholog	
Dnaic1	Dynein intermediate chain 1, axonemal;Dnai1;ortholog	microtubule family cytoskeletal protein(PC00157)
Dnah7b	Dynein, axonemal, heavy chain 7B;Dnah7b;ortholog	hydrolase(PC00121);microtubule binding motor protein(PC00156)

Table S3. Dysregulated spermatogensis genes in *Alkbh5* KO spermatogenic cells.

Gene symbol	Gene name	Function
Catsperb	Cation channel sperm-associated protein subunit beta;Catsperb;ortholog	
Adcy10	Adenylate cyclase type 10;Adcy10;ortholog	
Sycp2	Synaptonemal complex protein 2;Sycp2;ortholog	
Ccr6	C-C chemokine receptor type 6;Ccr6;ortholog	
Tdrd12	Putative ATP-dependent RNA helicase TDRD12;Tdrd12;ortholog	nuclease(PC00170)
Stra8	Stimulated by retinoic acid gene 8 protein;Stra8;ortholog	
Strbp	Spermatid perinuclear RNA-binding protein;Strbp;ortholog	mRNA processing factor(PC00147)
Ccnyl1	Cyclin Y-like 1;Ccnyl1;ortholog	
Nek1	Serine/threonine-protein kinase Nek1;Nek1;ortholog	
Sycp1	Synaptonemal complex protein 1;Sycp1;ortholog	
Fanca	Fanconi anemia group A protein homolog;Fanca;ortholog	
Capza3	F-actin-capping protein subunit alpha-3;Capza3;ortholog	non-motor actin binding protein(PC00165)
Zfp37	Zinc finger protein 37;Zfp37;ortholog	KRAB box transcription factor(PC00029)
Grip1,Ncoa2	Nuclear receptor coactivator 2;Ncoa2;ortholog	acetyltransferase(PC00038);transcription factor(PC00218)
Fancl	E3 ubiquitin-protein ligase FANCL;Fancl;ortholog	ubiquitin-protein ligase(PC00234)
Crem	cAMP-responsive element modulator;Crem;ortholog	
Cabyr	Calcium-binding tyrosine phosphorylation-regulated protein;Cabyr;ortholog	
Pum1	Pumilio homolog 1;Pum1;ortholog	mRNA processing factor(PC00147);translation factor(PC00223)
Adad1	Adenosine deaminase domain-containing protein 1;Adad1;ortholog	DNA binding protein(PC00009);RNA binding protein(PC00031);deaminase(PC00088);defense/immunity protein(PC00090);kinase activator(PC00138)
Pax5	Paired box protein Pax-5;Pax5;ortholog	
Ube2b	Ubiquitin-conjugating enzyme E2 B;Ube2b;ortholog	
Odf3	Outer dense fiber protein 3;Odf3;ortholog	structural protein(PC00211)
Sfmbt1	Scm-like with four MBT domains protein 1;Sfmbt1;ortholog	chromatin/chromatin-binding protein(PC00077);transcription factor(PC00218)
Micalcl	MICAL C-terminal-like protein;Micalcl;ortholog	
Tex14	Inactive serine/threonine-protein kinase TEX14;Tex14;ortholog	
Chd5	Chromodomain-helicase-DNA-binding protein 5;Chd5;ortholog	
Rara	Retinoic acid receptor alpha;Rara;ortholog	C4 zinc finger nuclear receptor(PC00169);nucleic acid binding(PC00171);receptor(PC00197)
Ctnnb1	Catenin beta-1;Ctnnb1;ortholog	
Tle3	Transducin-like enhancer protein 3;Tle3;ortholog	transcription cofactor(PC00217)
Atat1	Alpha-tubulin N-acetyltransferase 1;Atat1;ortholog	
Eif4g3	Eukaryotic translation initiation factor 4 gamma 3;Eif4g3;ortholog	translation initiation factor(PC00224)
Prm2	Protamine-2;Prm2;ortholog	
Spag1	Sperm-associated antigen 1;Spag1;ortholog	
Tgfbr1	TGF-beta receptor type-1;Tgfbr1;ortholog	TGF-beta receptor(PC00035);serine/threonine protein kinase receptor(PC00205)
Pacrg	Parkin coregulated gene protein homolog;Pacrg;ortholog	
Brca2	Breast cancer type 2 susceptibility protein homolog;Brca2;ortholog	damaged DNA-binding protein(PC00086)
Brdt	Bromodomain testis-specific protein;Brdt;ortholog	
Odf4	Outer dense fiber protein 4;Odf4;ortholog	
Tbc1d21	TBC1 domain family member 21;Tbc1d21;ortholog	G-protein modulator(PC00022);cysteine protease(PC00081)
Abhd2	Monoacylglycerol lipase ABHD2;Abhd2;ortholog	serine protease(PC00203)
Ednra	Endothelin-1 receptor;Ednra;ortholog	
Ttc26	Intraflagellar transport protein 56;Ttc26;ortholog	
Hmga2	High mobility group protein HMGI-C;Hmga2;ortholog	DNA binding protein(PC00009)
Stx2	Syntaxin-2;Stx2;ortholog	SNARE protein(PC00034)
Immp2l	Mitochondrial inner membrane protease subunit 2;Immp2l;ortholog	
Msh4	MutS protein homolog 4;Msh4;ortholog	DNA binding protein(PC00009)
Stag3	Cohesin subunit SA-3;Stag3;ortholog	chromatin/chromatin-binding protein(PC00077)
Rnf17	RING finger protein 17;Rnf17;ortholog	
Ccdc63	Coiled-coil domain-containing protein 63;Ccdc63;ortholog	
Rabl2	Rab-like protein 2A;Rabl2;ortholog	
Marf1	Meiosis arrest female protein 1;Marf1;ortholog	RNA binding protein(PC00031)
Arid4b	AT-rich interactive domain-containing protein 4B;Arid4b;ortholog	transcription cofactor(PC00217)
Fignl1	Fidgetin-like protein 1;Fignl1;ortholog	non-motor microtubule binding protein(PC00166)