Supplemental Material

Widespread association of the Argonaute protein AGO2 with meiotic chromatin suggests a distinct nuclear function in mammalian male reproduction

Griffin, Walters, Li, et al.

Other Supplemental Material (See separate .xlsx files)

Supplemental Table S1. AGO2 protein interactors identified by IP-MS.

Supplemental Table S2. AGO2 RNA interactions identified by eCLIP.

Supplemental Table S3. AGO2 ChIP alignment statistics and peaks.

Supplemental Table S4. Shared target genes identified by ChIP and eCLIP (see Fig. 3J).

Supplemental Table S5. Differentially expressed proteins in *Ago2* cKO vs. WT germ cells.

Supplemental Table S6. Differentially expressed proteins in post-meiotic vs. meiotic germ cells.

Supplemental Table S7. Differentially expressed genes in *Ago2* cKO vs. WT germ cells.

Supplemental Table S8. Alternative splicing in *Ago2* vs. WT germ cells.

Supplemental Methods

Antibodies

ChIP-seq and eCLIP-seq were performed using anti-AGO2 (Millipore, 04-642). Immunoblotting was performed using the following primary antibodies and dilutions: Histone H3 (Abcam, ab1791, 1:40000 or Abcam, ab18521, 1:1000), AGO2 (Abcam, ab186733, 1:2000 or Abcam, ab32381, 1:1000), H3K9me3 (Abcam, ab8898, 1:4000), SMARCA5 (Thermo Fisher Scientific, PA-52601, 1:1000), GAPDH (Santa Cruz, sc-32233, 1:1000), CNOT1 (Proteintech, 66507-1-Ig, 1:2000), HMGB2 (Abcam, ab124670, 1:3000), DAZL (Abcam, ab34139, 1: 2000), YBX2 (Abcam, ab154829, 1:4000), AK7 (Novus Biologicals, NBP2-92176, 1: 3000), GTSF1 (NOVUS Biologicals, NBP1-83934,1:1000), PTBP2 (Proteintech, 55186-1-AP, 1:4000), α-TUBULIN (Santa Cruz, sc-8035, 1:1000), SYCP3 (Abcam, ab97672, 1:1000), ACRV1 (guinea pig, gift of Dr. Prabhakara Reddi, University of Illinois (Osuru et al. 2014), 1:5000).

Cell culture

Mouse embryonic stem cells were grown at 37°C and 5% CO_2 on 0.1% gelatin and feeder cells (Gibco CF-1 MEF MITC-Treated) in DMEM (Gibco 11965-092) + 15% fetal bovine serum (Sigma-aldrich F2442), 1:100 pen-strep (Gibco 15140-122), 1:100 GlutaMax (Gibco 35050-061), 1:100 MEM NEAA (Gibco 11140-050), Sodium Pyruvate 1:100 (Gibco 11360-070), HEPES (Gibco 15630-080), β-mercaptoethanol (Sigma-Aldrich M6250), and 1:10000 LIF (Millipore ESG1106). Cells were passaged 1:10 every 2–3 days with 1× trypsin-EDTA 0.25% (Gibco 25200-056).

Testis collection and cell dissociation

Seminiferous tubules were isolated by enzymatic digestion (0.75-1 mg/ml collagenase IV, Gibco 17104-019) for 10 min in a 37°C water bath as described previously (Bryant et al. 2013). Male germ cells were obtained from seminiferous tubules using a second enzymatic digestion (0.25% trypsin (EDTA-free, Sigma), 2 mg/ml DNase I (Stem Cell Technologies, 07900)). The trypsin was neutralized with DMEM containing 10% fetal bovine serum (Life Technologies) and removed by centrifugation at 400×g for 5 mins before resuspending. The cell suspension was filtered in two steps through a 100 μm followed by a 40 μm cell strainer, and then used for either STA-PUT or flow cytometry.

Morphological assessment of spermatozoa

Cauda epididymides were dissected from adult (8-16 week old) male mice and allowed to swim out into PBS in a 12 well culture dish for 15 mins. The sample was then centrifuged for 2 min at 2000×g at 4°C and the resulting pellet was resuspended in 4% paraformaldehyde for 20 mins at room temperature. Fixed spermatozoa were pelleted and washed twice in 100 mM ammonium acetate, then resuspended in a final volume of 100 μl wash buffer. 7 μl of this cell suspension was smeared onto a glass slide and air dried followed by staining for 10 min in Coomassie blue solution or for 30 mins in 10ug/ml fluorescent dye-conjugated lectin-peanut agglutinin (lectin PNA, Invitrogen, L21409) in PBS. Coomassie stained slides were rinsed in water and allowed to air dry before coverslipping with Permount (Fisher Scientific, SP15-100). The proportion of sperm with abnormal heads was visually assessed from >400 sperm for each of 3-4 biological replicates using brightfield microscopy under blinded conditions. Slides stained with lectin PNA were rinsed in PBS, coverslipped with ProLong Gold Antifade Mountant with DAPI (Invitrogen, P36931) and

imaged using confocal microscopy. Statistical significance was tested by an unpaired *t*-test with Welch's correction.

Sperm motility assessment

Cauda epididymal sperm were collected by swim-out in HEPES buffered saline (HS) medium. Aliquots of sperm were placed in a slide chamber (CellVision, 20 mm depth) and motility was examined on a 37°C stage of a Nikon E200 microscope under a 10× phase contrast objective (CFI Plan Achro 10×/0.25 Ph1 BM, Nikon). Images were recorded (40 frames at 50 fps) using a CMOS video camera (Basler acA1300-200um, Basler AG, Ahrensburg, Germany) and analyzed by computer-assisted sperm analysis (CASA, Sperm Class Analyzer version 6.3, Microptic, Barcelona, Spain). Sperm total motility and hyperactivated motility was quantified simultaneously. Over 200 motile sperm were analyzed for each trial. Three and two biological replicates were performed for control and *Ago2* cKO, respectively.

STA-PUT

STA-PUT was performed as described in (Bellve 1993). For each isolation, 6-8 testes from adult wild type mice were pooled and washed twice in Dulbecco Modified Eagle medium (DMEM). A linear gradient was generated using 350 ml of 2% BSA and 350 ml of 4% BSA solutions in the corresponding chambers. Approximately 1×10^8 male germ cells were resuspended in 20ml of 0.5% BSA solution and loaded to the sedimentation chamber. After 3 hours of sedimentation in the sedimentation chamber, 60 fractions were collected in 15 ml centrifuge tubes and numbered sequentially 1-60. Cells from each fraction were collected by centrifugation at 500×g for 5 min and resuspended in 0.2 ml total volume. An aliquot of each fraction was stained with Hoechst dye (Invitrogen, H3570) and examined thoroughly by eye under phase-contrast and fluorescence microscopes to assess cellular integrity and identify cell types. Fractions containing >80% cells of appropriate size and morphology were pooled as pachytene spermatocytes ('meiotic cells') and round spermatids ('post-meiotic cells'). Purity was estimated at 85% for meiotic and 90% for post-meiotic populations based on visual examination of brightfield and Hoechst-stained cells.

Flow cytometry

For flow sorting, dissociated cells were diluted to 5×10^6 cells/ml. For isolation of round spermatids, the cell suspension was incubated at 37°C with 2 µl/ml Vybrant DyeCycle Green (Life Technologies, V35004) for 20 minutes. Round spermatids were sorted based on 1C DNA content and size using a 2-laser sorter (Bio-Rad S3e with 488nm and 561nm lasers) gated on single cells as previously described (Lesch et al. 2019). Differentiating spermatogonia ('pre-meiotic cells') were isolated as KIT^+ spermatogenic cells using anti-c-Kit antibody directly conjugated to PE (eBioscience 12-1171-82).

Immunoblotting

2 µg of total protein derived from meiotic or post-meiotic germ cell populations was loaded onto a Mini-PROTEAN TGX gel (Bio-Rad, 456-8093) and separated by SDS-PAGE for 1 hour at 150V. Separated proteins were then wet transferred to a 0.45 µm nitrocellulose membrane (GE Healthcare Life Science, 10600003) in Towbin buffer (25mM Tris, 192mM glycine, 20% (v/v) methanol) at a constant current of 250mA for 1 hour. Following transfer, membranes were briefly incubated in Ponceau S (Sigma Aldrich, P7170) to assess transfer quality. Membranes were then blocked in 5% w/v non-fat dry milk-TBST for 30 min with gentle agitation at room temperature.

Primary antibodies were diluted in blocking buffer and incubated with the membrane overnight at 4°C with gentle rocking. Three 3-minute washes were performed with TBST before incubating immunoblotted membranes for 1 hour at RT with peroxidase-conjugated anti-rabbit secondary antibody (Jackson Immuno Research, 111-035-003) diluted 1:20 000 in TBST. Following washing, the membrane was incubated in Pierce ECL substrate (Thermo Fisher Scientific, 32109) for 5 min at RT and the signal was captured with X-ray film (Thermo Fisher Scientific, 34090) for 5-60 sec. Immunoblots were quantified by densitometry using FIJI software (Schindelin et al. 2012). Unpaired *t*-tests were performed for each protein to assess statistical significance.

For Western blots of CLIP samples, protein samples were prepared by eCLIP lysis buffer (50 mM Tris-HCl pH 7.4, 100 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% sodium deoxycholate) containing a protease inhibitor cocktail (Roche, 11697498001). Protein samples were separated by SDS-PAGE and blotted onto the PVDF membranes. Transferred membranes were blocked in 5% w/v non-fat dry milk for 1h at room temperature, followed by incubation with diluted primary antibodies at 4°C overnight. After three washes with 0.05% TBST, the membranes were incubated with secondary antibodies conjugated with horseradish peroxidase at room temperature for 1h. After three washes with 0.05% TBST, immunodetection was performed with chemiluminescence as described above.

Co-immunoprecipitation of AGO2 protein interactors

Whole, cytoplasmic, and nuclear lysates were prepared from seminiferous tubules of adult mice using the enzymatic digestion protocol and subcellular fractionation protocol described above. 200µl of each lysate was incubated with 1µg of normal mouse IgG (Santa Cruz, sc-2025), Ago2 antibody (Millipore, 04-642), or SMARCA5 antibody (Thermo Fisher Scientific, PA-52601) overnight at 4°C with end-over-end mixing. 20 µl (10%) of each lysate was reserved as input. The next day, 20 µl of freshly washed Dynabeads Protein G (Thermo Fisher Scientific, 10004D) were incubated with the lysates for 3 hours at 4°C with-end-over end mixing. The beads were then collected onto a magnetic stand and washed three times for 5 mins using cold lysis buffer. Finally, 60 µl of 1× Laemmli SDS buffer containing β-mercaptoethanol was added to the beads and heated at 95°C for 10 mins. Candidate protein interactors were detected by Western blotting as described above, using 20µl of denatured sample.

RNA isolation

Cell pellets were homogenized in 1 ml of TRIzol reagent (Invitrogen, 15596026) and incubated at room temperature for 5 min. After incubation, 200 µl of chloroform was added, and samples were vortexed, incubated briefly at room temperature, and spun at 12000×g for 15 min at 4°C. The resultant aqueous phase was transferred to a new tube and mixed with an equal volume of 100% ethanol, then transferred to an RNeasy MinElute spin column (Qiagen, 74104). RNA was extracted on the column according to the manufacturer's instructions. The quality of the eluted RNA was determined by measuring 28S/18S ribosomal ratios and RNA integrity number (RIN) using an Agilent Bioanalyzer.

Luciferase assays

GC-2spd(ts) cells (ATCC, CRL-2196) were seeded a 96-well flat bottom plate (Corning, 3903) at 5,000 cells/well in 80µl of complete DMEM. After overnight incubation, the cells were transfected using lipofectamine 3000 (Thermo Fisher Scientific) with pmirGLO Dual-Luciferase miRNA Target Expression Vectors (Promega, E1330) harboring sequences corresponding to targets of

Ago2 binding at *Ak7*, *Spag17*, or *Tdrd9* mRNA. The inserted sequences are centered on the eCLIP peak (12-30bp) and extended in both directions for a final insert length of 400bp (**Supplemental Table S10**). When peaks in the two replicates were not identical, both peaks and surrounded sequences were concatenated for a total of 400bp. Sequences of all peaks were shuffled to generate a control vector. After 24 hours of incubation, transfected cells were incubated with 80µl of Dual-Glo luciferase reagent (Promega, E2920) for 10 mins before measuring firefly luciferase signal on a plate reader. The same volume of Dual-Glo Stop & Glo was then added to each well and incubated for 30 mins before measuring *Renilla* luciferase on the plate reader. Firefly luciferase signals were normalized to *Renilla* luciferase measurements and then a final relative response ratio was calculated for each experimental sample based on normalized ratios from the shuffled control sample, according to the manufacturer's instructions.

Real-time quantitative PCR

Reverse transcription of 1µg of total RNA was performed with random hexamers and SuperScript III reverse transcriptase (Thermo Fisher Scientific #12574026) in a total volume of 20µl according to the manufacturer's instructions. Reaction mixtures were incubated in a thermocycler at 25°C for 10 min then 50°C for 50 min before stopping the reaction at 85°C for 5 min. cDNA was then diluted 1:5 with nuclease-free water. Real-time quantitative PCR (qPCR) assays were performed in a total reaction volume of 20ul consisting of 4µl of diluted cDNA, 0.4µl of 10µM forward/reverse primer mix, 10µl of Power SYBR Green PCR Master Mix (Applied Biosystems 4367659) and 5.6µl nuclease-free water. Primer sequences used for qPCR are listed in **Supplemental Table S9**. Reactions for each target gene were performed in duplicate in a 96 well plate loaded into an Applied Biosystems QuantStudio 3 Real-Time PCR System. Standard cycling conditions were used: Hold stage (×1): 50°C for 2 min, 95°C for 10 min; PCR stage (×40): 95°C for 15 sec, 60 °C for 1 min. Melt curve stage conditions were: 95°C for 15 secs, 60°C for 1 min, 95°C for 15 secs. Relative fold change in transcript abundance was calculated using the delta-delta Ct method by normalizing target gene expression levels to beta-actin. Significant differences in gene expression were evaluated using an unpaired *t*-test.

eCLIP

AGO2 eCLIP in spermatogenic cell fractions and whole testis tissue was performed as described previously (Van Nostrand et al. 2016; Biancon et al. 2022; Brugiolo et al. 2017). Briefly, purified germ cells were isolated from adult mouse testes and crosslinked with 254 nm UV light. Nuclear and cytoplasmic fractions were isolated, and nuclear samples were lysed and sonicated in a Bioruptor sonicator (Diagenode) at 4C, 5 min, 30 sec on/30 sec off. The lysate was then digested with RNase I (AM2295, Thermo Fisher Scientific) and Turbo DNase (AM2239, Thermo Fisher Scientific), and immunoprecipitated using anti-AGO2 antibody (04-642, Millipore). Next, the protein-bound RNAs were ligated with 3' end adapters. Protein-RNA complexes were subjected to gel electrophoresis and transferred to a membrane, and RNA was isolated by cutting the membrane at the expected size range. cDNA was generated by reverse transcription, DNA adapter was ligated to the 3' end of the cDNA, and the library was amplified by PCR. Screening of subclones prior to high-throughput sequencing was performed for quality control.

Chromatin immunoprecipitation

ChIP was performed as described previously (Lesch et al. 2013, 2016). Cells were cross-linked for ChIP-seq with 1% formaldehyde at room temperature for 10 min. The formaldehyde was quenched

with 2.5M glycine at room temperature for 10 minutes. Fixed cells were spun down at 6000×g for 3 minutes at 4°C. Cells were washed with cold PBS twice and resuspended in 100ul ChIP lysis buffer (1% SDS, 10mM EDTA, 50mM Tri-HCl at pH 8.1) before being snap frozen in liquid nitrogen. Antibody-bound Dynabeads (Invitrogen 00821318) were prepared by mixing 10 µl aliquots of beads with 100 µl block solution (0.5% BSA in PBS). Beads were washed twice with 150 µl of block solution and resuspended in 30ul block solution. 1 µl of AGO2 antibody (Millipore, 04-642) was added to each aliquot of beads and incubated for 8 hours rotating at 4°C.

Between 2×10^5 to 5×10^6 cells were used for ChIP-seq. Frozen cells were thawed on ice and ChIP dilution buffer (0.01% SDS, 1.1% Triton X-100, 1.2mM EDTA, 167mM NaCl, 16.7mM Tris-HCl at pH 8.1) was added to reach a total volume of 300 µl. Cells were split into 150 µl aliquots in 0.5 ml Eppendorf tubes and sonicated at 4°C for 30 cycles (30 seconds on/off) using a Bioruptor (Diagenode). Aliquots of the same samples were pooled and spun down at 12,000×g for 5 min. The supernatant was moved to a new 1.5 ml Eppendorf tube and 600 µl of dilution buffer and 100 µl protease inhibitor cocktail (Roche #11836153001) were added. 50 µl of each sample was set aside as input before an aliquot of anti-AGO2 antibody bound Dynabeads was added and left overnight rotating at 4°C. After overnight incubation, beads were washed twice with low-salt immune complex wash buffer (0.1% SDS, 1% Triton X-100, 2mM EDTA, 150mM NaCl, and 20mM Tris-HCL at pH 8.1), twice with LiCl wash buffer (0.25M LiCl, 1% NP40, 1% deoxycholate, 1mM EDTA, and 10mM Tris-HCl at pH 8.1), and twice with TE (1 mM EDTA and 10mM Tris-HCl at pH 8.0). Bound DNA was eluted twice with 125 µl elution buffer (0.2% SDS, 0.1M NaHCO3, and 5mM DTT in TE) at 65°C and crosslinking was reversed with an 8 hour incubation at 65°C. ChIP and input samples were incubated for 2 hours at 37°C with 0.2 mg/ml RNase A (Millipore 70856-3), and 2 hours at 55°C with 0.1 mg/ml Proteinase K (NEB P8107S).

Samples were prepared for sequencing using a Zymo ChIP DNA Clean & Concentrator kit (Zymo Research #D5201) according to the manufacturer's instructions. Columns were washed twice with 200 µl wash buffer and DNA was eluted twice into fresh Eppendorf tubes: first with 7 µl of elution buffer and then with 6ul elution buffer.

RNase treatment

mESC nuclei were treated with RNase according to a protocol adapted from (Sigova et al. 2015). Approximately 1×10^6 mouse ESCs were grown in culture as described above. Cells were dissociated with 1× trypsin-EDTA 0.25%, spun down at 1500 rpm for 3 minutes, and resuspended in PBS. Cells were then washed twice with cold PBS, then resuspended in 5ml cold hypotonic solution (20mM HEPES, 20% glycerol, 10mM NaCl, 1.5mM MgCl2, 0.2 mM EDTA, 10% Triton X-100, 0.5 mM DTT, 1mM protease inhibitor, in H_2O). Cells were incubated on ice for 10 min before nuclei were spun down and supernatant removed. The nuclear pellet was resuspended in 250 µl of cold PBS and incubated at room temperature with 15 µl RNase A (Millipore Sigma 70856-3) and 10 µl RNase H (New England Biolabs M0297S). The nuclei were prepared for ChIP as described above.

Mass spectrometry analysis for AGO2 protein interactors

Nuclear and cytoplasmic fractions from wild type round spermatids purified by STA-PUT were isolated as described above and incubated with anti-AGO2 antibody (Millipore, 04-642). A parallel sample was incubated with normal mouse IgG as a negative control. Immunoprecipitated samples were then separated by SDS-PAGE and the gel was stained with Coomassie blue. All bands except immunoglobulin heavy chain and light chain were excised from the gel and subjected

to mass spectrometry protein identification at the Keck MS & Proteomics Resource at Yale University. Briefly, individual excised gel bands in 1.5 ml Eppendorf tubes were washed 4 times; first with 500 µL 60% acetonitrile containing 0.1% TFA and then with 5% acetic acid, then with 250 µL 50% H2O/50% acetonitrile followed by a 250 µL 50% CH3CN/ 50 mM NH4HCO3, and a final wash with 250 µL 50% CH3CN/10 mM NH4HCO3 prior to removal of wash and complete drying of gel pieces in a Speed Vac. 10 µL of a 0.1 mg/mL stock solution of trypsin (Promega Trypsin Gold MS grade) in 5 mM acetic acid was freshly diluted into a 140 µL solution of 10 mM NH4HCO3 to make the working digestion solution. 124 µL of the working digestion solution was added to the dried gel pieces (additional 10 mM NH4HCO3 may be added to ensure gel pieces are completely submerged in the digestion solution) and incubated at 37 ºC for 18 hours. Sample was then stored at -20 ºC until analysis. Digested sample was injected onto a Q-Exactive Plus (Thermo Fisher Scientific) LC MS/MS system equipped with a Waters nanoACQUITY™ UPLC system, and used a Waters Symmetry® C18 180 µm × 20 mm trap column and a 1.7 µm, 75 µm × 250 mm nanoACQUITY™ UPLC™ column (37ºC) for peptide separation. Trapping was done at 5 µl/min, 99% Buffer A (100% water, 0.1% formic acid) for 3 min. Peptide separation was performed with a linear gradient over 140 minutes at a flow rate of 300 nL/min. The data were processed and protein identification was searched using Proteome Discoverer (v. 2.2, Thermo Fisher Scientific, Waltham, MA) and Mascot search algorithm (v. 2.6, Matrix Science LLC, London, UK). Mascot search parameters included: fragment ion mass tolerance of 0.020 Da, parent ion tolerance of 10.0 ppm, strict trypsin fragments (enzyme cleavage after the C-terminus of Lysine or Arginine, but not if it is followed by Proline), variable modification of Phospho Ser, Thr, and Tyr, Oxidation of Met, and Propioamidation of Cys, deamidation of Asn and Gln, and Carbamidomethlyation of Cys. Scaffold (version 4.8.6, Proteome Software Inc., Portland, OR). A protein was considered identified when Mascot listed it as significant and more than 2 unique peptides matched the same protein. The Mascot significance score match is based on a MOWSE score and relies on multiple matches to more than one peptide from the same protein. The Mascot search results were exported to an .xml file using a False Discovery Rate of 1% or less for the protein ID.

Quantitative mass spectrometry analysis of protein expression

Pachytene spermatocytes or round spermatids were isolated by STA-PUT as described above and pellets containing 1-3×10^6 pachytene spermatocytes or 0.4-1×10^7 round spermatids were washed in PBS with protease inhibitors before sonicating in 400µl of RIPA buffer and then processed for Label Free Quantitative mass spectrometry at the Keck MS & Proteomics Resource. Briefly, the protein suspension was centrifuged at 14K rpm for 10 min, and 150 uL of the supernatant was removed. Chloroform:methanol:water protein precipitation was performed, and dried protein pellet was resuspended in 100 uL 8M urea containing 400mM ABC, reduced with DTT alkylated with iodoacetamide, and dual enzymatic digestion with LysC and trypsin (carried out at 37 C for overnight and 6 hrs., respectively). Digestion was quenched with 0.1% formic acid during the de-salting step with C18 MacroSpin columns (The Nest Group). The effluents from the de-salting step were dried and re-dissolved in 5 µl 70% FA and 35 µl 0.1% TFA. An aliquot was taken and concentration measured via Nanodrop, and diluted to 0.05 µg/µl with 0.1% TFA. 1:10 dilution of 10× Pierce Retention Time Calibration Mixture (Cat#88321) was added to each sample prior to injecting on the UPLC Q-Exactive Plus mass spectrometer to check for retention time variability and instrument QC during LFQ data collection.

LFQ data dependent acquisition (DDA) was performed on a Thermo Fisher Scientific Q-Exactive plus, Q-Exactive HFX, or Orbitrap Fusion Mass spectrometers connected to a Waters nanoACQUITY UPLC system equipped with a Waters Symmetry® C18 180 µm × 20 mm trap column and a 1.7-µm, 75 µm × 250 mm nanoACQUITY UPLC column (35°C). 5 µl of each digests (in triplicates) at 0.05 µg/µl concentration was injected in block randomized order. To ensure a high level of identification and quantitation integrity, a resolution of 60,000 was utilized for MS and 15 MS/MS spectra was acquired per MS scan using HCD. All MS (Profile) and MS/MS (centroid) peaks were detected in the Orbitrap. Trapping was carried out for 3 min at 5 µl/min in 99% Buffer A (0.1% FA in water) and 1% Buffer B [(0.075% FA in acetonitrile (ACN)] prior to eluting with linear gradients that reached 30% B at 140 min, 40% B at 155 min, and 85% B at 160 min. Two blanks (1st 100% ACN, 2nd Buffer A) followed each injection to ensure against sample carry over.

The LC-MS/MS data was processed with Progenesis QI software (Nonlinear Dynamics, version 2.4) with protein identification carried out using in-house Mascot search engine (v2.6). The Progenesis QI software performs chromatographic/spectral alignment (one run is chosen as a reference for alignment of all other data files to), mass spectral peak picking and filtering (ion signal must satisfy the 3 times standard deviation of the noise), and quantitation of proteins and peptides. A normalization factor for each run was calculated to account for differences in sample load between injections as well as differences in ionization. The normalization factor was determined by calculating a quantitative abundance ratio between the reference run and the run being normalized, with the assumption being that most proteins/peptides are not changing in the experiment so the quantitative value should equal 1. The experimental design was set up to group multiple injections (technical and biological replicates) from each run into each comparison set. The algorithm then calculates the tabulated raw and normalized abundances, and ANOVA *p* values for each feature in the data set. The MS/MS spectra was exported as .mgf (Mascot generic files) for database searching. Mascot search algorithm was used for searching against the Swiss Protein database with taxonomy restricted to *Mus musculus*; and carbamidomethyl (Cys), oxidation of Met, Phospho (Ser, Thr, Tyr), and Acetylation (N-term and Lys) , and deamidation (Asp, Arg) were entered as variable modifications. Two missed tryptic cleavages were allowed, precursor mass tolerance was set to 10ppm, and fragment mass tolerance was set to 0.02 Da. The significance threshold was set based on a False Discovery Rate (FDR) of 1%, and MASCOT peptide score of >95% confidence. The Mascot search results was exported as .xml files and then imported into the processed dataset in Progenesis QI software where peptides ID'ed were synched with the corresponding quantified features and their corresponding abundances. Protein abundances (requiring at least 2 unique peptides) were then calculated from the sum of all unique normalized peptide ions for a specific protein on each run.

High-throughput DNA sequencing

RNA sample integrity was assessed by Bioanalyzer (Agilent Technologies), and samples with RIN >7 were used for paired-end sequence library construction using the KAPA mRNA HyperPrep Kit (Roche # 08098123702). For ChIP samples, DNA integrity and fragment size were confirmed on a Bioanalyzer. Approximately 5-10 ng of DNA was end-repaired, A-tailed, adapter-ligated, and PCR enriched (8-10 cycles) using the KAPA Hyper Library Preparation kit (KAPA Biosystems, #KK8504). eCLIP libraries were prepared as described above according to (Van Nostrand et al. 2016). All indexed libraries were quantitated by qPCR using the KAPA Library Quantification Kit (KAPA Biosystems #KK4854). Samples were sequenced on an Illumina NovaSeq using 100bp

paired-end sequencing or an Illumina HiSeq 2500 instrument using 75 bp single-end reads. Sample de-multiplexing was performed using CASAVA 1.8.2 (Illumina).

ChIP-seq data analysis

ChIP datasets were quality filtered with the fastq_quality_filter tool from FASTX-Toolkit using parameters –q 20 –p 80 (Hannon 2010) and quality was assessed using FastQC (Andrews 2010). ChIP-seq data was aligned to the mouse genome build mm10 using Bowtie2 with --end-to-end --fast parameters (Langmead and Salzberg 2012, 2). Peaks were called with MACS2 at q=0.05 for all datasets (Zhang et al. 2008). Pachytene spermatocyte replicates were overlapped using the intersect tool from BEDTools (Quinlan and Hall 2010). For other cell types, the replicate with highest quality was used for analysis. Shuffled controls were generated by randomly re-assigning peaks in the dataset to new genome coordinates and leaving peak size unchanged, using the shuffle function in BEDTools (Quinlan and Hall 2010).

RNA-seq data analysis

Raw sequence files were filtered and quality was assessed using FASTX-toolkit and FastQC as described above. Filtered paired-end reads were matched to yield a file of common reads which were then pseudoaligned to the mouse (mm10) transcriptome from Ensembl release 102 (Howe et al. 2021) and quantified using kallisto v0.45.0 (Bray et al. 2016). Resulting transcript counts were summed to gene level and DEGs were called using DESeq2 with batch correction (Love *et al*. 2014). Lowly expressed genes with < 10 TPM were filtered from the analysis. Genes with fold changes ≥1.5 and adjusted *P* values ≤0.05 were considered differentially expressed.

RNA-seq libraries were sequenced to a depth of 100 million reads to allow splice variant analysis. Splicing was assessed using rMATS with default parameters (Shen et al. 2014), using FASTQ files were used as input. Differential splicing calls and downstream analysis were conducted using the maser package in R (Veiga 2018).

eCLIP analysis

AGO2 IP and size-matched input control (SMInput) samples were demultiplexed as previously described using the script demux_paired_end.py (Van Nostrand et al. 2016). 3’ and 5’ adapters were trimmed using cutadapt v1.18 (Martin 2011) with default parameters. To eliminate artifacts from rRNA and other repetitive sequences, trimmed reads were then mapped to Repbase (mouse mm10 RepeatMasker, last modified time: 05-May-2017 06:32) using STAR software v.2.6.1d (Dobin et al. 2013). Unmapped reads recovered after the RepeatMasker step were then mapped to the mouse mm10 genome assembly with options --outFilterScoreMin 0 --outFilterScoreMinOverLread 0.33 --outFilterMatchNminOverLread 0.33 --alignEndsType Local. The script barcode_collapse_pe.py (Van Nostrand et al. 2016) was used to remove PCR duplicates based on unique molecular identifiers (UMIs). Peaks were called using the tag2peak.pl script from the CLIP Tool Kit (CTK) software (Shah et al. 2017) with an adjusted p-value threshold of 0.01 and normalized to the mm10 background model. Each AGO2 eCLIP replicate was also normalized to the corresponding SMInput data using the Peak_input_normalization_wrapper.pl script from (Van Nostrand et al. 2016). eCLIP targets were confirmed by an additional independent peak calling pipeline (Biancon et al. 2022).

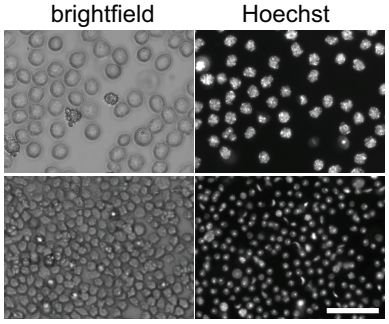
Supplemental Table S9. Primers.

RT-qPCR Primers		
Gene	**Forward (5' to 3')**	**Reverse (5' to 3')**
Actb	CCACTGTCGAGTCGCGTCC	GCCCACGATGGAGGGGAATA
Ago1	CGGGAAGACGCCAGTGTATG	TGAGGCAGAGGTTGGACAGAGT
Ago2	CACCATGTACTCGGGAGCC	GGCATATCCTGGGATGGGTGA
Ago3	CCTTTATACAGCCAATCCACTTC	GTCGGAGCACCACGTCAA
Ago4	GTGGACAGATGAAACGAAAA	TTGAAGGCAGGGAAGGTG
L15UTR	GGCGAAAGGCAAACGTAAGA	GGAGTGCTGCGTTCTGATGA
L1ORF2	GGAGGGACATTTCATTCTCATC	GCTGCTCTTGTATTTGGAGCATAGA
LINE1	GAGAACATCGGCACAACAATC	TTTATTGGCGAGTTGAGACCA
SINE	TTTGATCCCAGCACTCAGGAGG	CGAGACAGGGTTTCTCTGTGTATC
Genotyping Primers		
Gene	**Forward (5' to 3')**	**Reverse (5' to 3')**
Ago2 f1/r1 (floxed vs wt allele)	TGATCATGGTTGAGGTCTGA	GTGAGCCACTCACTGTGCAC
Ago2 f4/r4 (del. allele)	CTCAGTTAAGGTTAATGCCTCTTG	TGCTGCTTCTTCTACCGTCTC
Cre	TGGGCGGCATGGTGCAAGTT	CGGTGCTAACCAGCGTTTTC

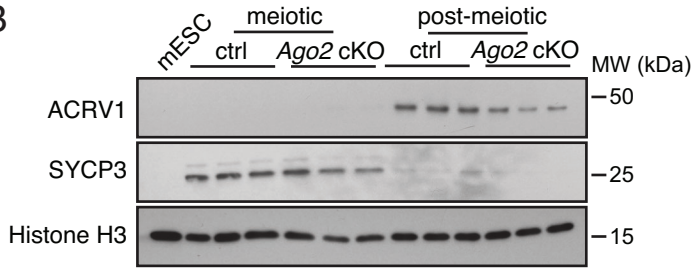
Supplemental Table S10. Target sequences for luciferase assay.

Target gene name	Genome coordinates (mm10)	Cloned DNA sequence (5' to 3')
Ak7	chr12:105747175-105747575	GTGGCATACAGGAGGCGATTGTCGCACCCAAGGACTCAGTGGAAGGAG AAGAGGAAGGCGAAGAGGAGGAGGAGGAAGAGAACGTGGATGATGCC CAGGAGCTCCTTGATGGCATCAAGGAAAGCATGGAGCAGAACGCAGGT GACGCCCAGAGAGCCGGTGGGGTGAGAGAGCCGAGCCAGGCGGCCCA GAGCTCTTCCGATCTGGTTCTAAAATGCCCGGCTGCAGGCTGGAAAGAT AGCTCAGTGGTCAAGACCTTCGGCATGCTTTCCCAGAGTACCCCAGTGT GGTTCCCAGCACCCACACACCCGGGTCACAACCTTGAACTCCAGGGGAT CCAAAAACTTCTGGACTGCACGGGGACTGCCCTGGTGCTTGCAAATGTA CATGGGCGCTCACAC
Spag17	chr3:99952583-99952783, chr3:99937301-99937501	CACCCCCTGTCCACTTTCCTTCCCAGGTCTCTGACCCATAATTGTTCCTG TCTGAAAGAATTACATGGATGGGAATGAAGAGAGGAGCCTGAGAAAAA GAAGGTCCAACGACAGGCCCAATATGGGATCCAGCTCAAGGGGAGGTC CCAAGGCCTGACACTATTACTGAGGCTATGGAGTTCTCACAATAAGGAA CCTATCGGCAAAGAAACGTGCAGGGAGAAGGGGGCAGTGGGAGCAGTT TTTATTGCTGAAACTTTAACACACCGCCTACCATGAGCAATGAATACTT TTCCATGAGAACGGAAGAGACTTTCCTCCTTCAGCTCTCTCTATAGATG CAGCTATAATGGGAGGGGTAAAAGCGGTACCCAGTGCAAAGAGGGTCA GTTTGTGGCAATA
Tdrd9	chr12:112010438-112010638, chr12:111971948-111972148	AATCGCCTGTGAGCTCACACTCACCCCAACAGTCACCGGGATCTTGATA AGAGGCCAAAACTACACATGGGAGAGAAGATAGCATCAGCAGCAAGT GGGGCTGGTCAGACTGCTACATAGCAGGTAAAGAACAGAACTAGATCT TCCCTCACACCCTGTCCACACCAGTTAGTGCCAAGTGGATCAGGTACCA CAGCGTAGAAGCCGAGCAACCTAAGCATTTCGTAATTTGTTTATTATAA TTTGCAGCCTGGGTTTTTTTGTTTTTGTTTTTGTTTGTTTTGAAAGAGTCC AGTGTTGCCCAGTCTGGCCTCAACTTACACCTCTGTGGATAACTTTGAA TTCTTCAGCCTCTGGCTTCTACTTCCTCAAGTCCTGAGATTACAGGGTTA CATCTGCAAC
Shuffled control	n/a	GGAGAAGTAAAGAGAAAGGCAGTCCTCAGTTAGCTTGATTGGAATACT ATTCGCAGCGATGTGCAAGAATGAATGGACAAAGGAAATATTTGCAGA GTCCACATGATCGTCTTATCGCGCGGTACGGGTCCTGGGTCGAGAAAAT CCCATAGTTGATAGATTGCTACCAACCTCGATGTCGCTCCATGAGTCTG GACGTAAATGTCCAGAGGCATTGTACGTTCCGCGTATCGCTTCCAGATG CGTCGCAGCTTGCGTACCTGAGGAGAAAGAGATCAATTGCCTGAGGTG CGAGGCAGATTGCCCAAGGACACTACAGGATTATATGATACTTGACTAC AGCAACCAGTTCCGACAGTAGGAGCCACGCAACCGAACAGTCACTCGT TAGAGCTTTTAAAGGGGT

A



B

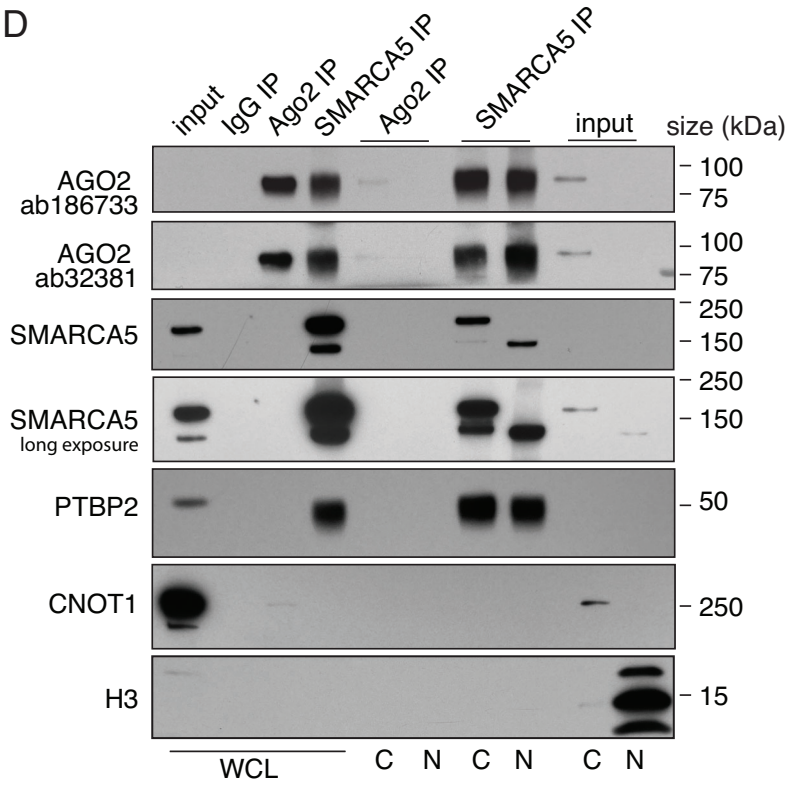


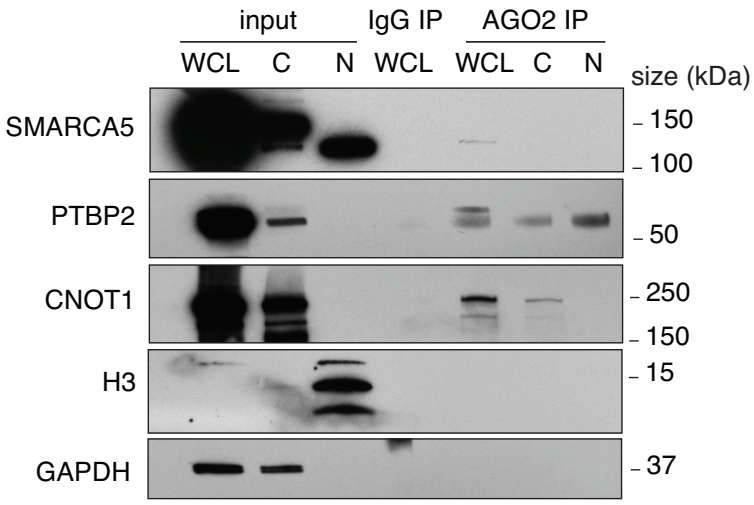
C

Selected nuclear AGO2-interacting proteins

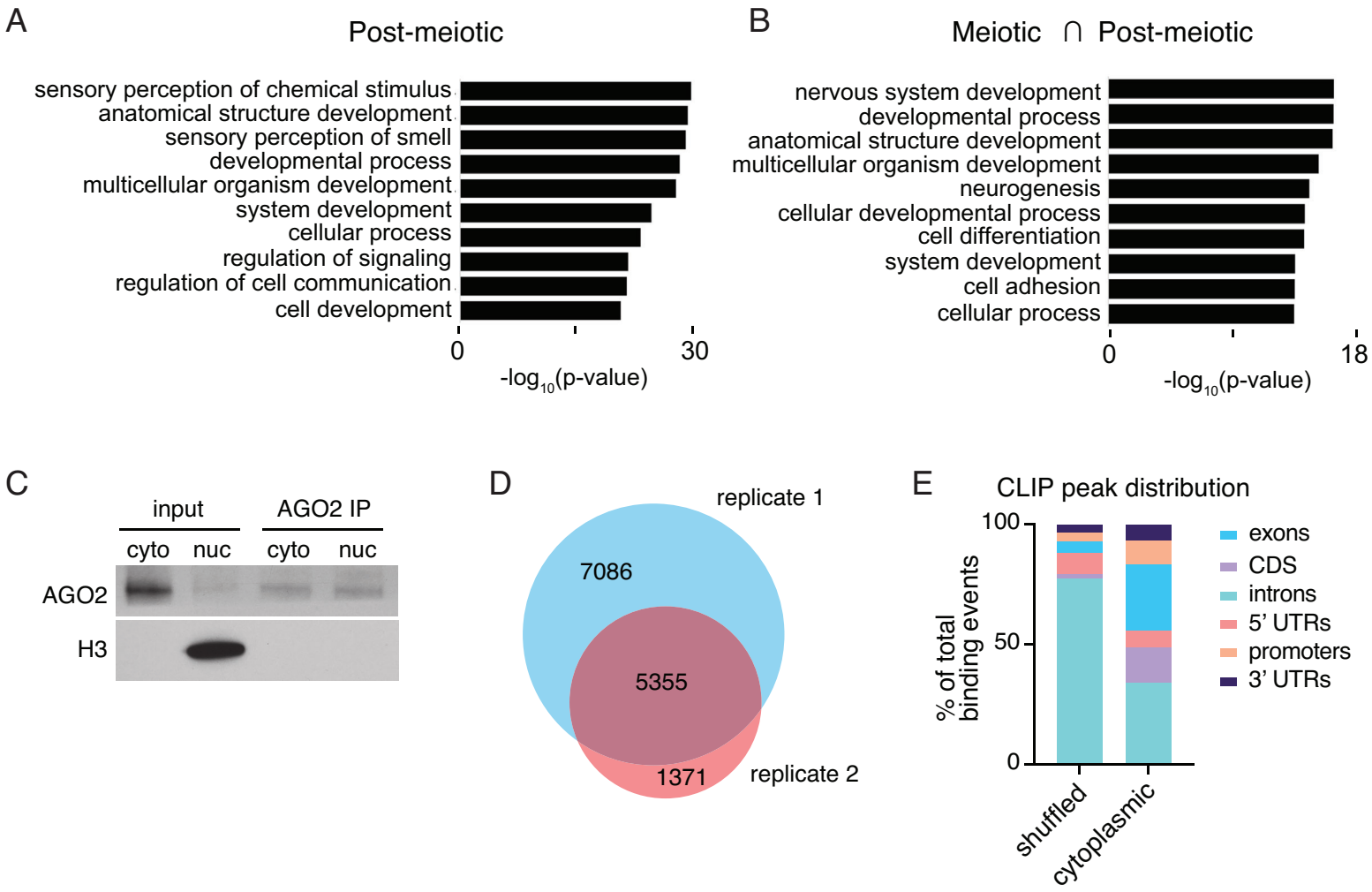
Protein ID	Cytoplasm IP # peptides	Nucleus IP # peptides
Splicing		
PTBP2	0	5
PSPC1	67	43
RBM39	3	3
SUGP2	3	22
PRPF8	4	10
HNRNPC	2	11
Ribonucleoprotein nuclear export		
NUP210L	16	25
NUP160	3	14
NUP98	7	15
NUP133	3	9
NXF1	5	23
Nucleosome positioning		
H1-2	3	7
H1-3	2	4
SMARCA5	0	3

D

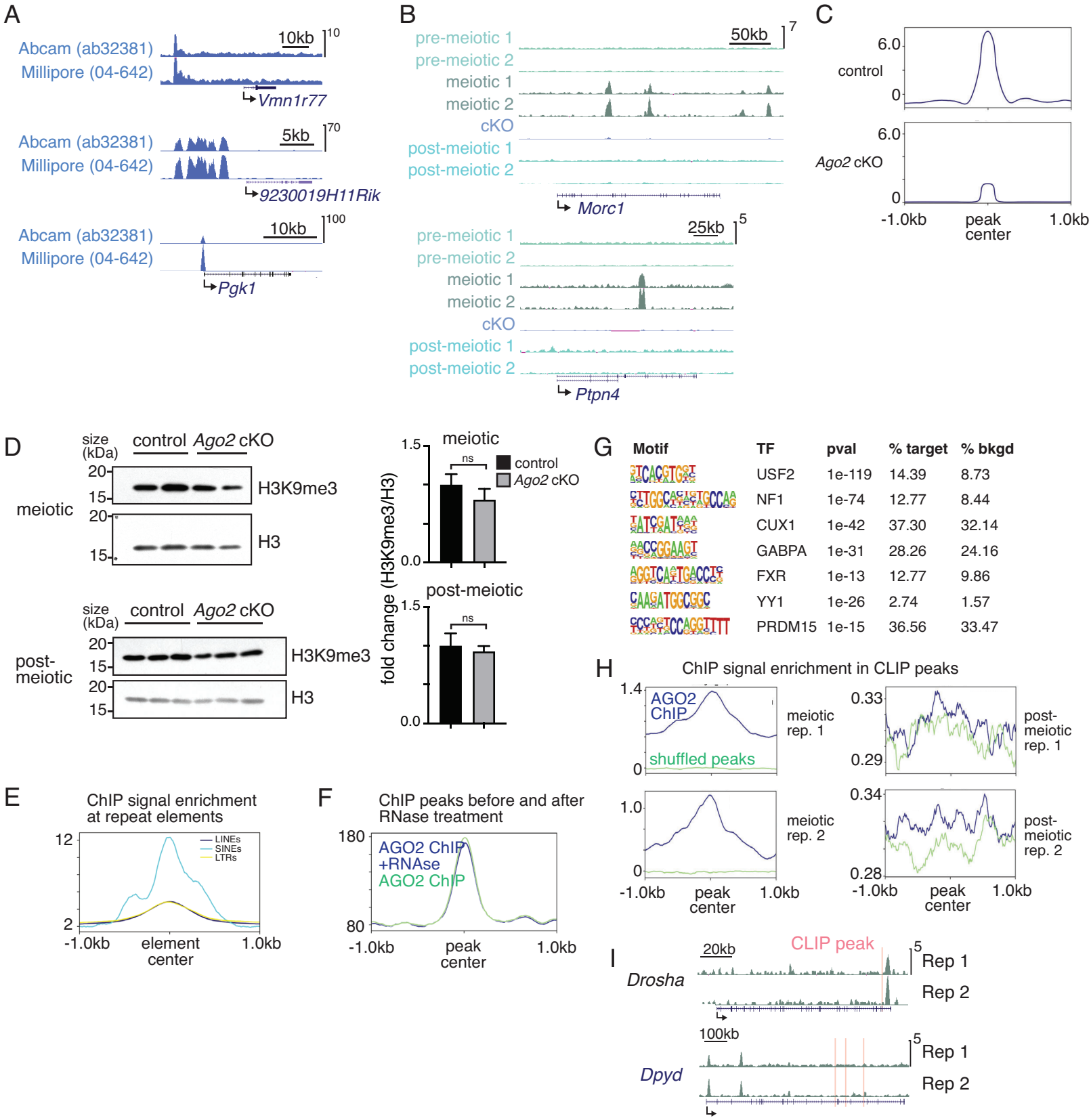




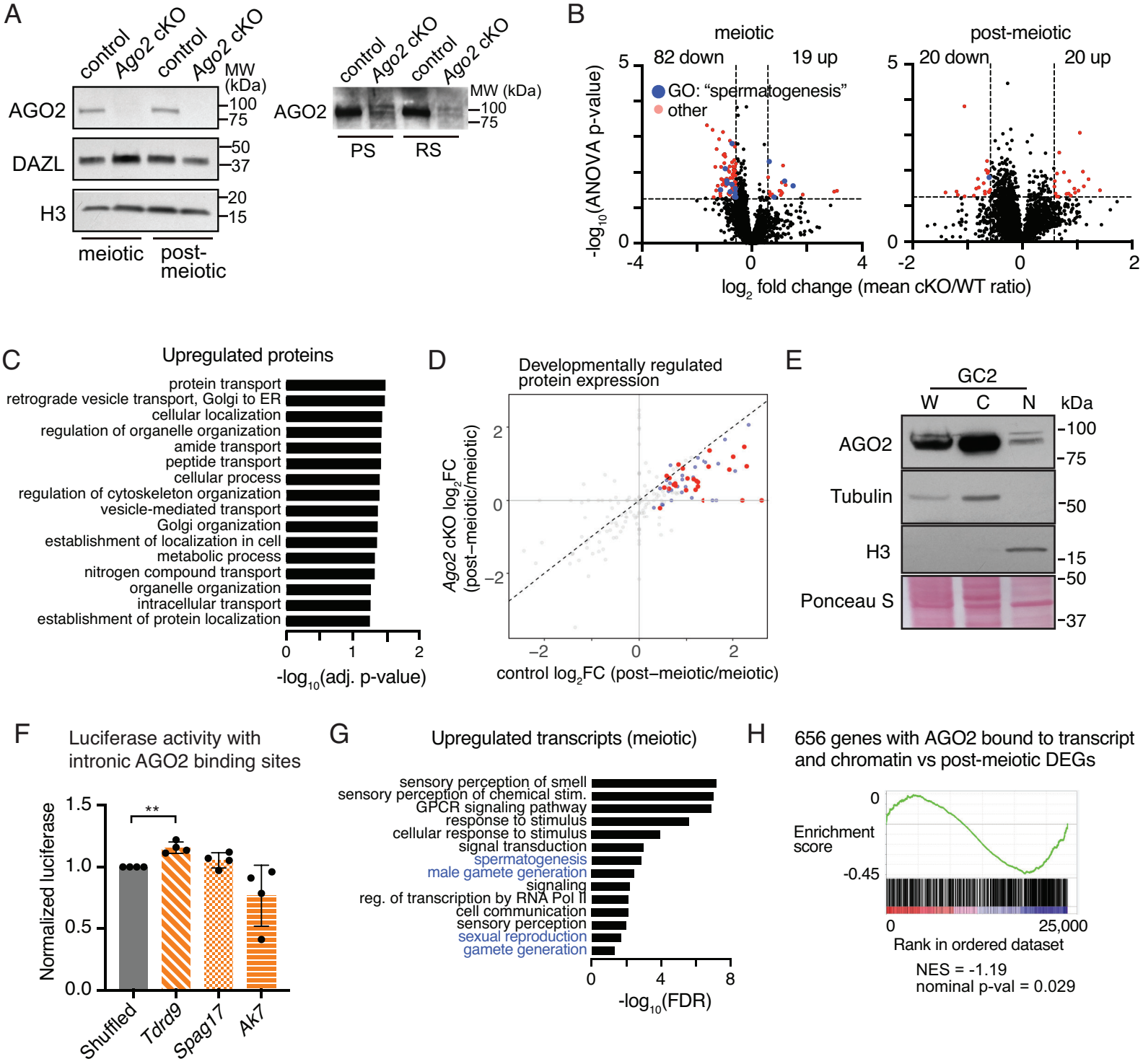
Supplemental Figure S1. AGO2 immunoprecipitation and mass spectrometry (IP-MS) in germ cell nuclear and cytoplasmic fractions. A, Brightfield and Hoechst-stained images of meiotic (pachytene spermatocytes) and post-meiotic (round spermatid) cell populations isolated by STA-PUT. **B,** Western blot for three replicates of meiotic (pachytene spermatocyte) and post-meiotic (round spermatid) populations isolated by STA-PUT from control and *Ago2* cKO males. Specificity of the meiotic marker SYCP3 and post-meiotic marker ACRV1 is consistent with >85% purity of the cell fractions. **C**, Number of unique peptides detected in AGO2 IP-MS in wild type germ cells for proteins associated with splicing, nuclear export, and nucleosome positioning. **D**, Validation of AGO2 nuclear and cytoplasmic interactors by co-immunoprecipitation in wild type whole testis. GAPDH and histone H3 are specific to cytoplasmic (C) and nuclear (N) fractions, respectively. Top, immunoprecipitation of SMARCA5 recovers AGO2, but no band is seen for the reciprocal experiment (AGO2 immunoprecipitation followed by blotting for SMARCA5). Bottom, high concentration protein loading recovers AGO2 interactions with PTBP2 (nuclear and cytoplasmic) and CNOT1 (cytoplasmic only). The SMARCA5 nuclear-specific isoform, but not the cytoplasmic isoform, is present in AGO2 IP from whole cell lysate (WCL), supporting a nuclear interaction between AGO2 and SMARCA5.



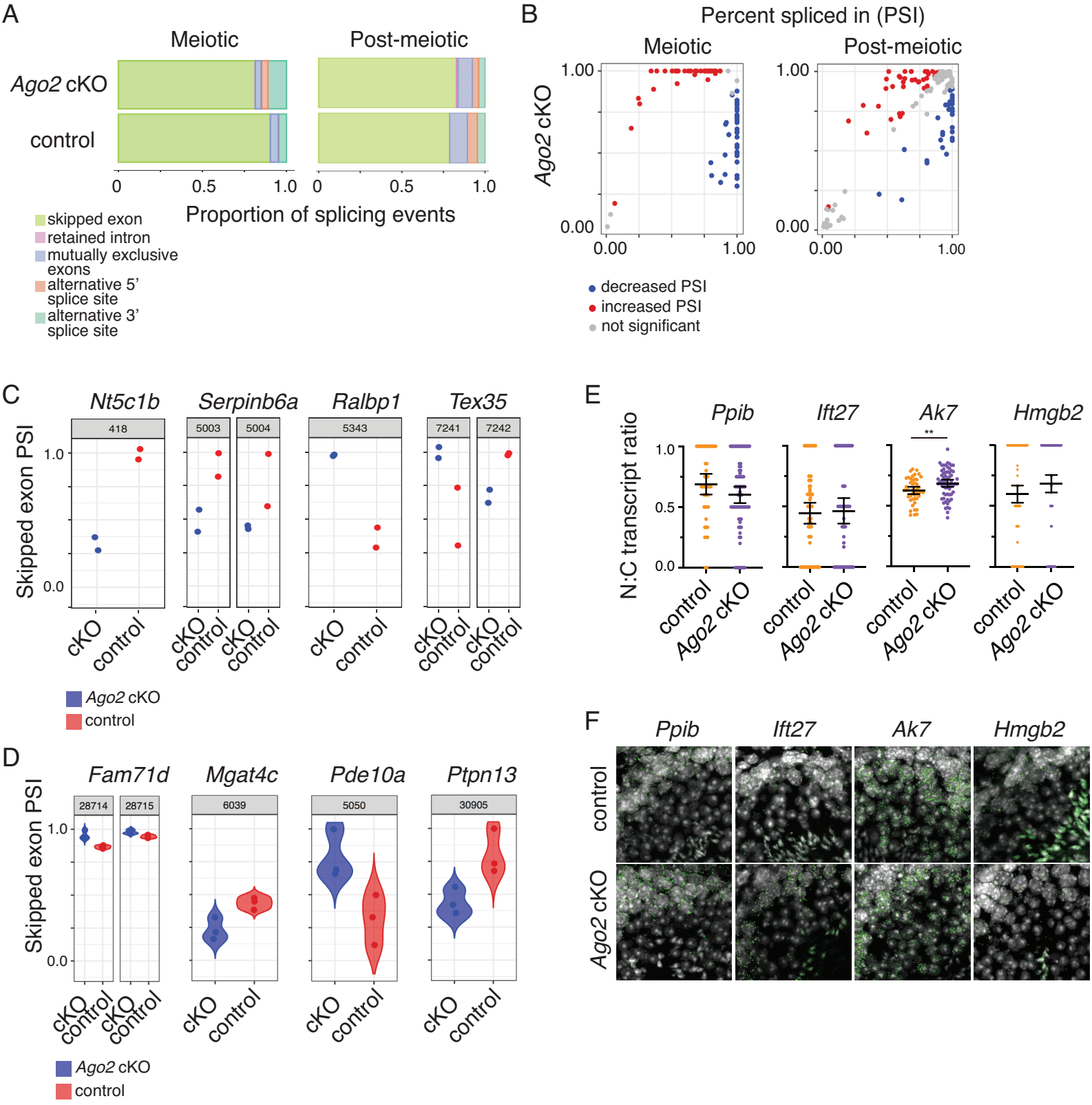
Supplemental Figure S2. Characterization of nuclear and cytoplasmic AGO2-bound transcripts identified by eCLIP. A, Gene Ontology (GO) analysis of post-meiotic nuclear consensus targets. **B**, GO analysis of targets bound in nuclei of both meiotic and post-meiotic cells. **C,** Western blot analysis of input and AGO2 IP samples used for cytoplasmic eCLIP. H3 is specific to the nuclear fraction, indicating efficient separation of cytoplasmic and nuclear compartments. **D,** Overlapping gene targets from cytoplasmic eCLIP replicates. **E,** Genomic distribution of cytoplasmic AGO2 eCLIP binding regions. For the shuffled control, peaks were randomized across all gene regions.



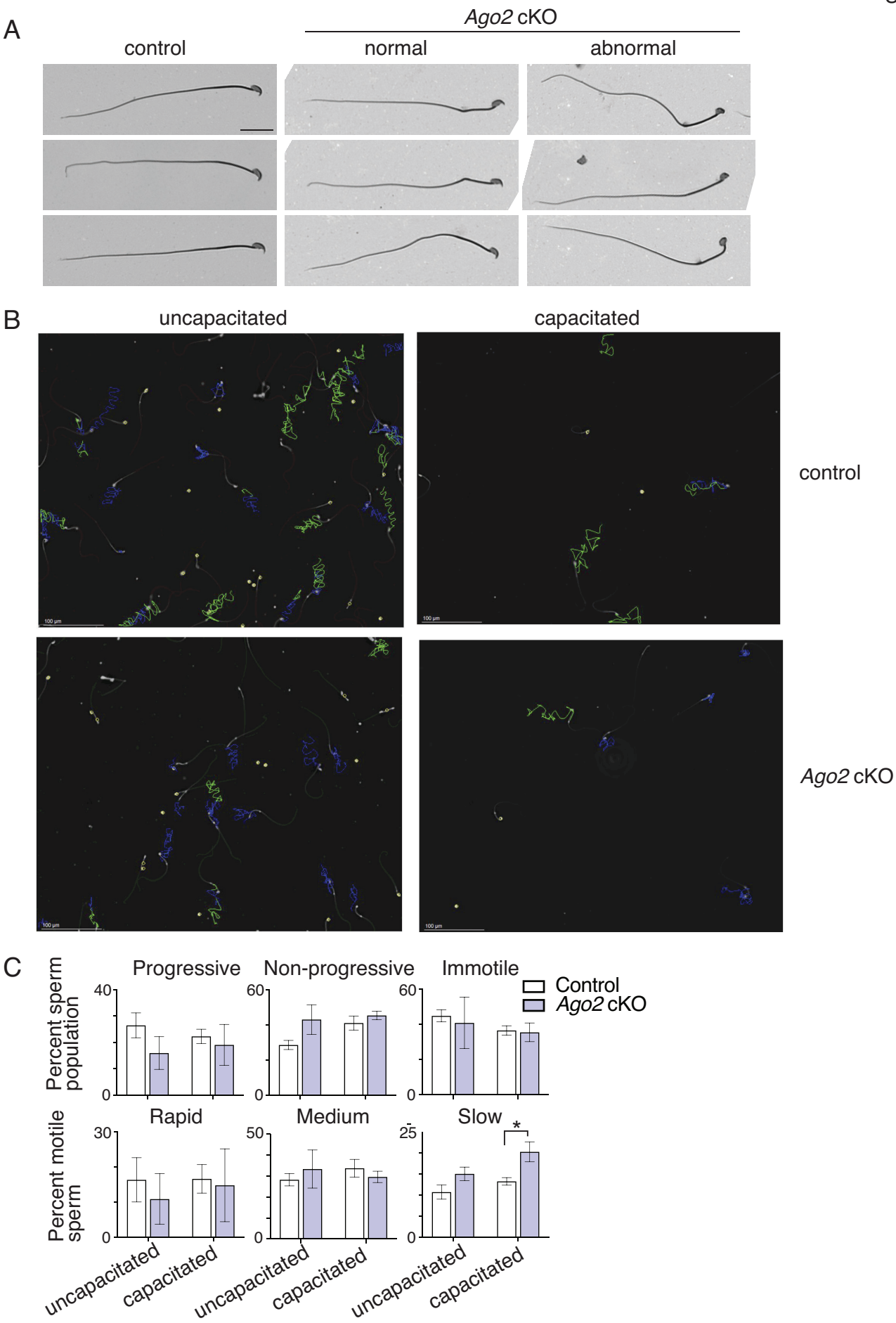
Supplemental Figure S3. AGO2 ChIP-seq in male germ cells. A, Genome browser tracks from AGO2 ChIP-seq in mESCs using two different antibodies. **B**, Additional representative genome browser tracks showing meiotic cell-specific AGO2 ChIP peaks in two biological replicates of pre-meiotic, meiotic, and post-meiotic wild type cells, and from *Ago2* cKO mixed meiotic and post-meiotic cells. **C**, Metagene plots comparing ChIP signal in control and *Ago2* cKO cells. **D**, Western blots showing global H3K9me3 levels in *Ago2* cKO cells. Quantitation is shown at right. Differences are not significant at a threshold of $p<0.05$, paired *t*-test. **E,** Metagene plot showing AGO2 ChIP signal at repeat element consensus sequences. **F**, Metagene plot comparing AGO2 ChIP signal in mESC nuclei treated with RNase to an RNase-negative control. **G**, Additional motifs enriched in meiotic cell AGO2 ChIP-seq peaks, where the corresponding transcription factors are expressed in meiotic male germ cells (see also **Figure 3F**). **H**, Separate metagenes for each biological replicate for AGO2 ChIP enrichment signal at eCLIP peaks(see also **Figure 3G**). **I**, Additional representative genome browser tracks for genes associated with both ChIP and CLIP peaks (see also **Figure 3H**).



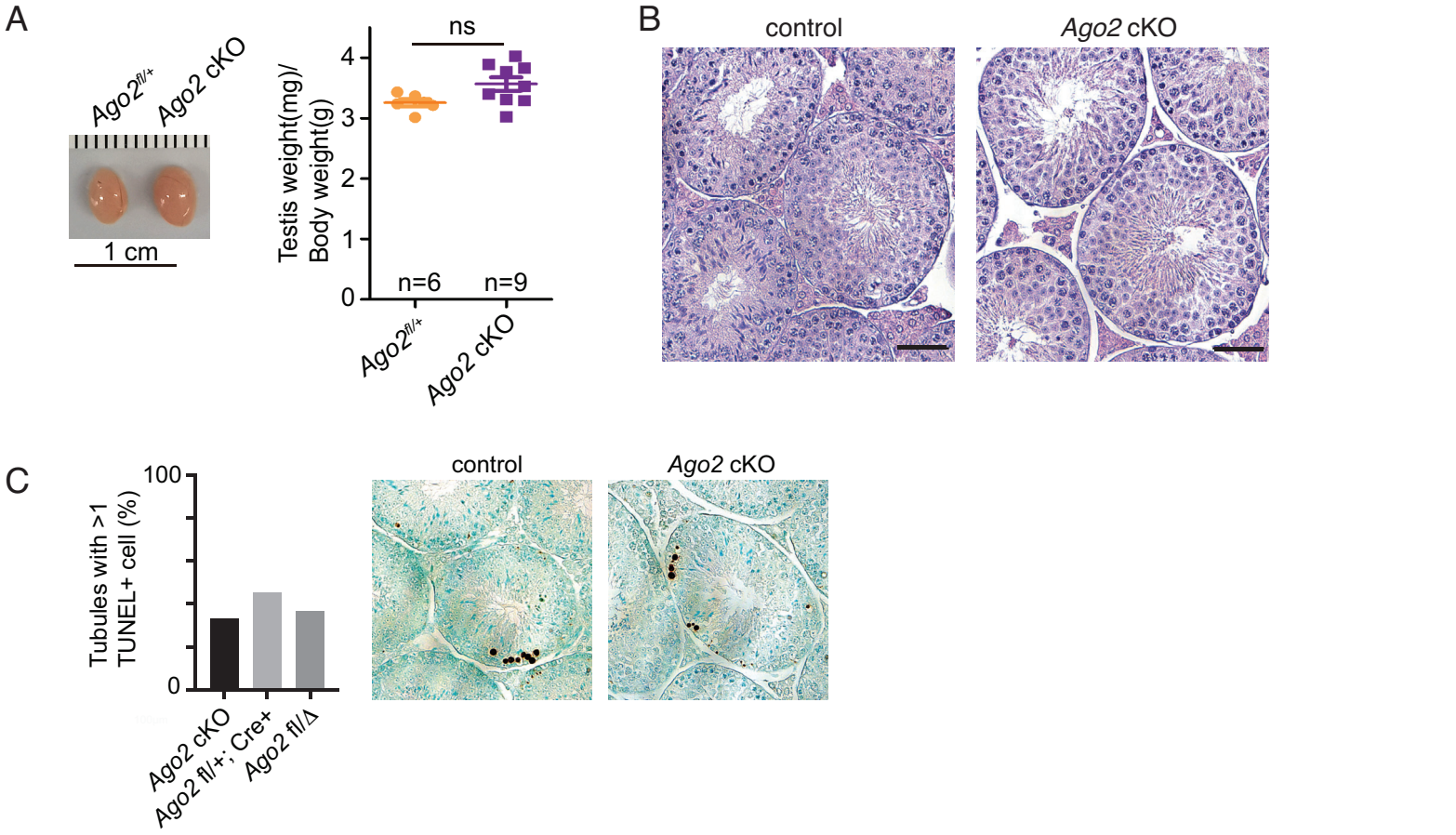
Supplemental Figure S4. Transcript and protein changes in *Ago2* cKO germ cells. A, Western blot showing loss of AGO2 in meiotic and post-meiotic germ cells of *Ago2* cKO males. DAZL marks germ cells; H3 is a loading control. Right-hand blot is a long exposure of the left blot. **B**, Volcano plots showing global protein level changes in meiotic and post-meiotic male germ cells from *Ago2* cKO males. Statistically significant differentially expressed proteins ($p \leq 0.05$ and $\log_2FC > 1.5$) are shown by a dashed line and highlighted in red. Differentially expressed proteins belonging to the Gene Ontology (GO) category "spermatogenesis" (GO:0007283) are highlighted in blue. **C**, Statistically enriched GO terms (FDR ≤ 0.05) associated with upregulated proteins in *Ago2* cKO post-meiotic cells. **D**, Changes in protein expression of meiotic AGO2 CLIP targets between meiotic and post-meiotic stages. Significantly upregulated proteins encoded by transcripts bound by AGO2 exclusively in the nucleus are in red. **E,** Western blot for AGO2 in nuclear and cytoplasmic fractions of GC2 cells used for luciferase assay in (F), showing that AGO2 is present in nuclei of GC2 cells. **F**, Luciferase activity following addition of predicted AGO2 intronic binding sites from three eCLIP-identified targets to the luciferase transcript. Data were normalized to a control containing shuffled binding site sequences. **p<0.01, paired *t*-test. **G**, Selected enriched GO terms associated with upregulated transcripts in *Ago2* cKO meiotic cells. Terms relevant to male reproduction are highlighted in blue. **H,** Gene set enrichment analysis (GSEA) for nuclear AGO2 targets (Fig. 3I) among differentially expressed transcripts in post-meiotic cells. NES, normalized enrichment score.



Supplemental Figure S5. Altered splicing and nuclear transcript retention in *Ago2* cKO germ cells. A, Overall proportion of different alternative splicing events in control and *Ago2* cKO meiotic and post-meiotic germ cells. **B,** Change in exon percent spliced in (PSI) for individual transcripts with a significant difference in skipped exons in *Ago2* cKO compared to control germ cells. **C**, Examples of transcripts with differential exon skipping in control and *Ago2* cKO meiotic cells. **D**, Examples of transcripts with differential exon skipping in control and *Ago2* cKO post-meiotic cells. **E,** Nuclear:cytoplasmic (N:C) ratios in post-meiotic cells for two transcripts bound (*Ak7*, *Hmgb2*) and two transcripts not bound (*Ift27*, *Ppib*) by nuclear AGO2. Localization was assessed by single-molecule RNA in situ hybridization (RNAscope) in *Ago2* cKO compared to control germ cells. Data points show ratios for individual cells from at least three tubules. Bars show mean and 95% confidence interval. **p<0.01, Welch's *t*-test. **F**, Sample RNAscope images in control and *Ago2* cKO testes.



Supplemental Figure S6. Sperm function in *Ago2* cKO males. A, Brightfield images showing the full length of the spermatozoa with heads shown in **Figure 5B**. Scale bar, 20μm. **B**, Representative images of control and *Ago2* cKO sperm motility tracks from computer-aided sperm analysis (CASA). **C**, CASA results showing motility parameters for control and *Ago2* cKO sperm before and after capacitation. *p<0.05, Student's *t*-test.



Supplemental Figure S7. Spermatogenesis in *Ago2* cKO males. A, Gross images of testes from *Ago2* cKO males and littermate controls (left) and quantitation of testis/body weight ratios (right). **B,** Hematoxylin & eosin (H&E) stained testicular sections from *Ago2* cKO and control adult males. Scale bar, 50 μm. **C**, Quantitation (left) and sample images (right) of TUNEL staining in *Ago2* cKO and littermate control (*Ago2 fl/+*; *Cre+* or *Ago2 fl/Δ*) testes. **D**, RT-qPCR for LINE and SINE elements in *Ago2* cKO and control post-meiotic cells. **E**, RT-qPCR for all *Ago* family members in whole testes and post-meiotic cells of *Ago2* cKO and control males. *p<0.05, Welch's *t*-test; all other comparisons not significant.