

Fig S1. Specificity of CBX2 and HA antibodies against endogenous and HA-tagged CBX2

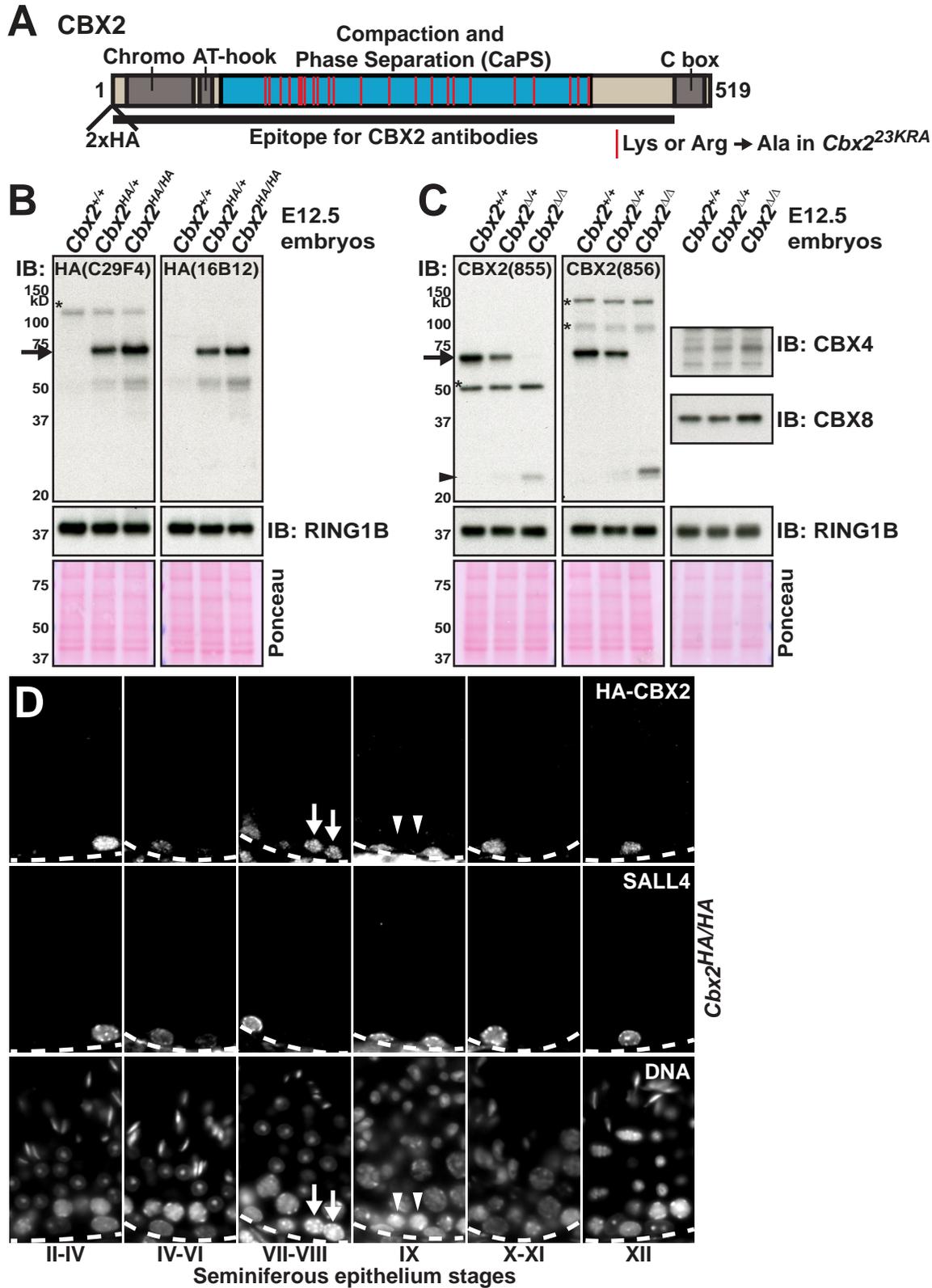


Figure S1. Specificity of CBX2 and HA antibodies against endogenous and HA-tagged CBX2

(A) Schematic domain structure of CBX2. The positions of Lys and Arg replaced with Ala in the *Cbx2*^{23KRA} mutant are marked by red bars.

(B and C) Western blot of CBX2 using (B) HA antibodies with wild-type, *Cbx2*^{HA/+}, *Cbx2*^{HA/HA} and (C) affinity purified CBX2 antibodies with wild-type, *Cbx2*^{Δ/+}, *Cbx2*^{Δ/Δ} embryonic day 12.5 (E12.5) embryo lysates. The HA knock-in mouse (*Cbx2*^{HA/HA}) was homozygous-viable and fertile.

RING1B was blotted to show the total PRC1 complex amount. Ponceau staining was used as the loading control. (C) A full-length CBX2 band was detected at right below the 75 kD size marker (arrow), and a truncated CBX2 band was detected at slightly above the 20 kD size marker from *Cbx2* deletion embryos (arrowhead). CBX4 and CBX8 were blotted to show potential paralogous compensation in *Cbx2* deletion embryos. Asterisks (*) denote non-specific background bands.

(D) Co-immunofluorescence staining of testis sections of a *Cbx2*^{HA/HA} animal with HA (C29F4) and SALL4 antibodies. Different seminiferous tubule stages were denoted at the bottom. Dotted lines represent the basement membrane. Arrows indicate pre-leptotene spermatocytes (CBX2+) and arrowheads indicate leptotene spermatocytes (CBX2-).

Fig S2. CBX2 is upregulated as spermatogonial stem cells differentiate

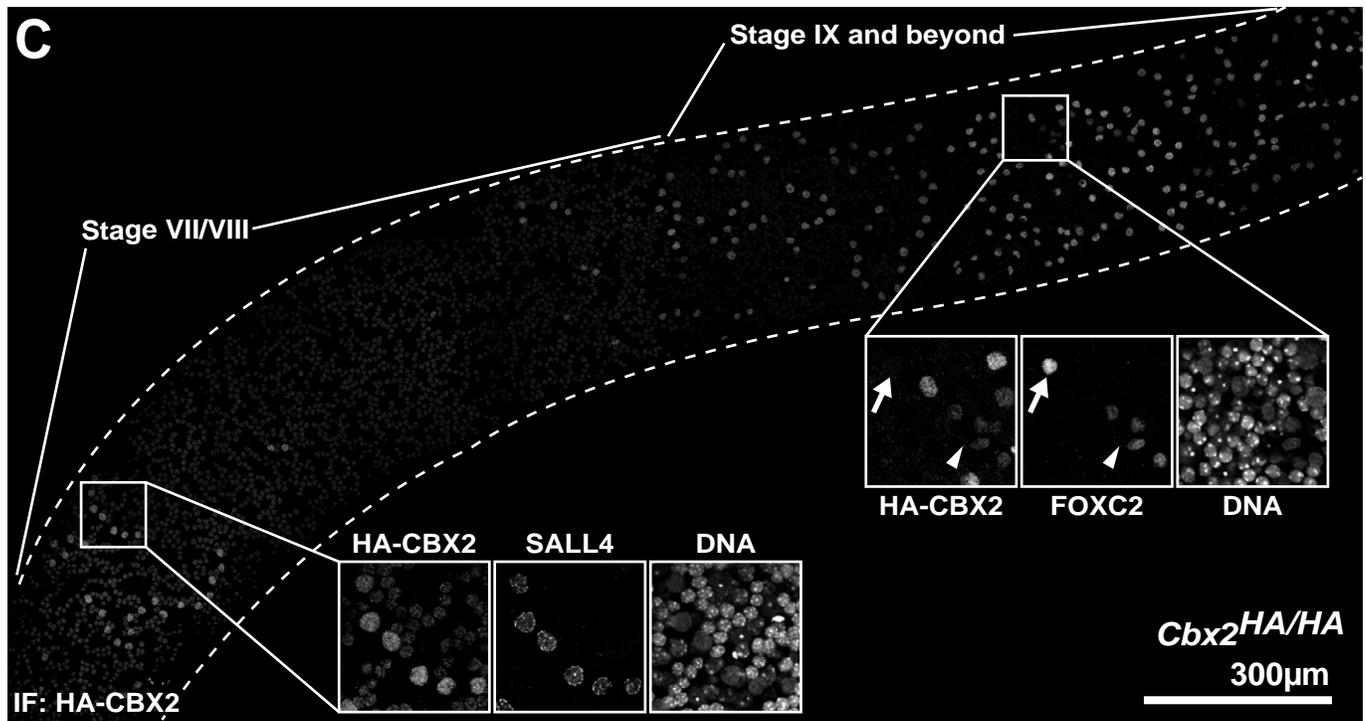
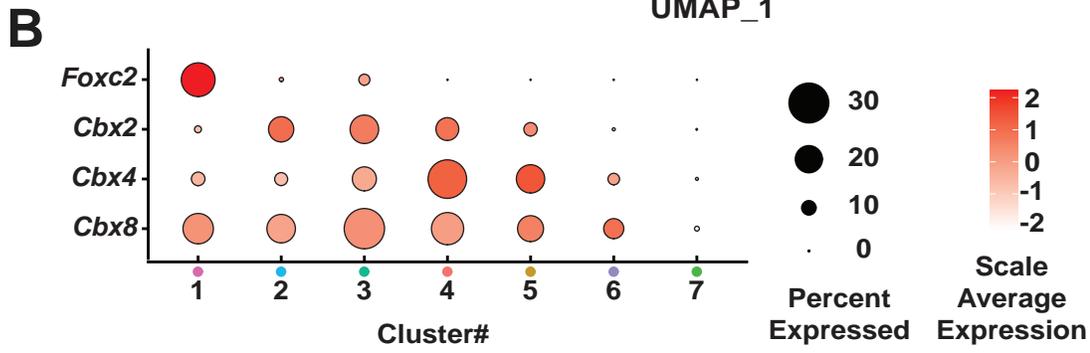
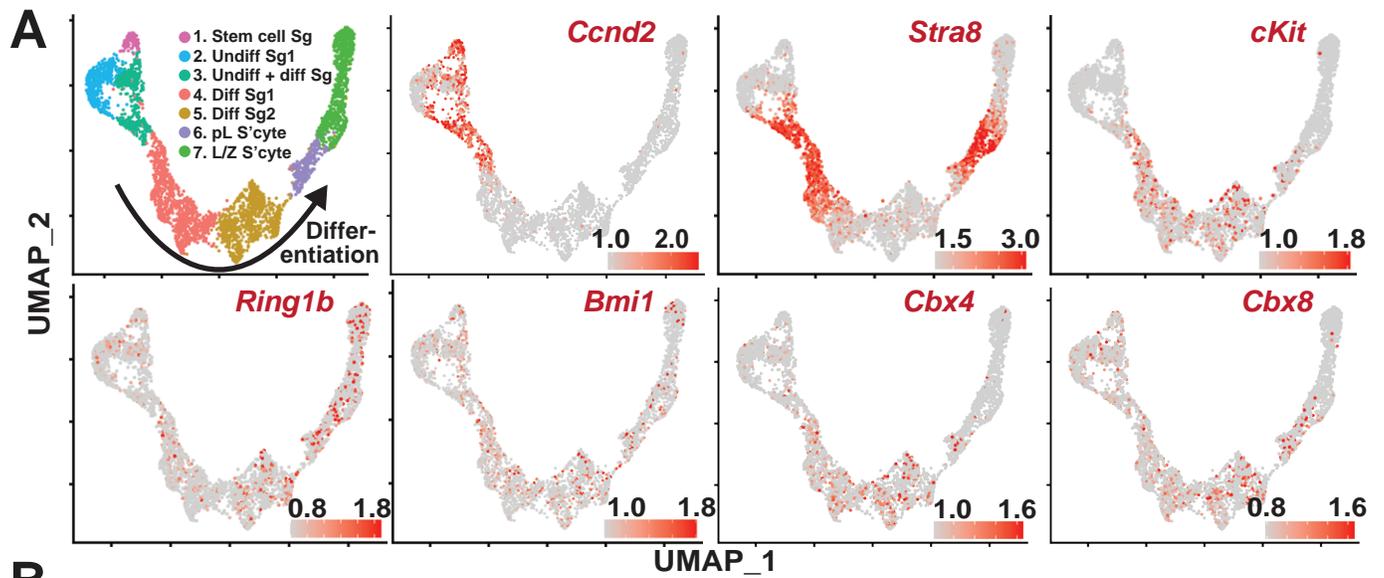


Figure S2. CBX2 is upregulated as spermatogonial stem cells differentiate

(A) UMAP representation of a subset of spermatogonia and early spermatocytes from scRNA-seq from p15 mouse testes shown in Fig. 2B (Ernst et al. 2019). Normalized expression levels of *Ccnd2* (enriched in a subset of undifferentiated spermatogonia and A1 spermatogonia), *Stra8* (upregulated first in A1 spermatogonia and later in pre-leptotene spermatocytes), *c-Kit* (enriched in differentiating spermatogonia), and *Ring1b/Bmi1/Cbx4/Cbx8* (components of cPRC1) are represented in red.

(B) Dot plots showing the percentage (dot size) of cells expressing *Foxc2*, *Cbx2*, *Cbx4* and *Cbx8* in each cluster and scaled average expression (shade of red). Data are scaled to have mean 0 with standard deviation 1.

(C) Co-immunofluorescence staining of a whole testis tubule of *Cbx2^{HA/HA}* animals with HA, SALL4, and FOXC2 antibodies. CBX2 signal was represented in contiguous tubule stages from VII to IX and beyond. The inset from stage VII/VIII is an example of SALL4(+) A1 spermatogonia, and smaller pre-leptotene spermatocytes with CBX2 signal. The inset from stage IX and beyond is an example of As FOXC2(+) spermatogonia without CBX2 signal (arrows) and Aaln4 FOXC2(+) spermatogonia with weak CBX2 signal (arrowheads).

Fig S3. Hox loci overlap with CBX2 puncta

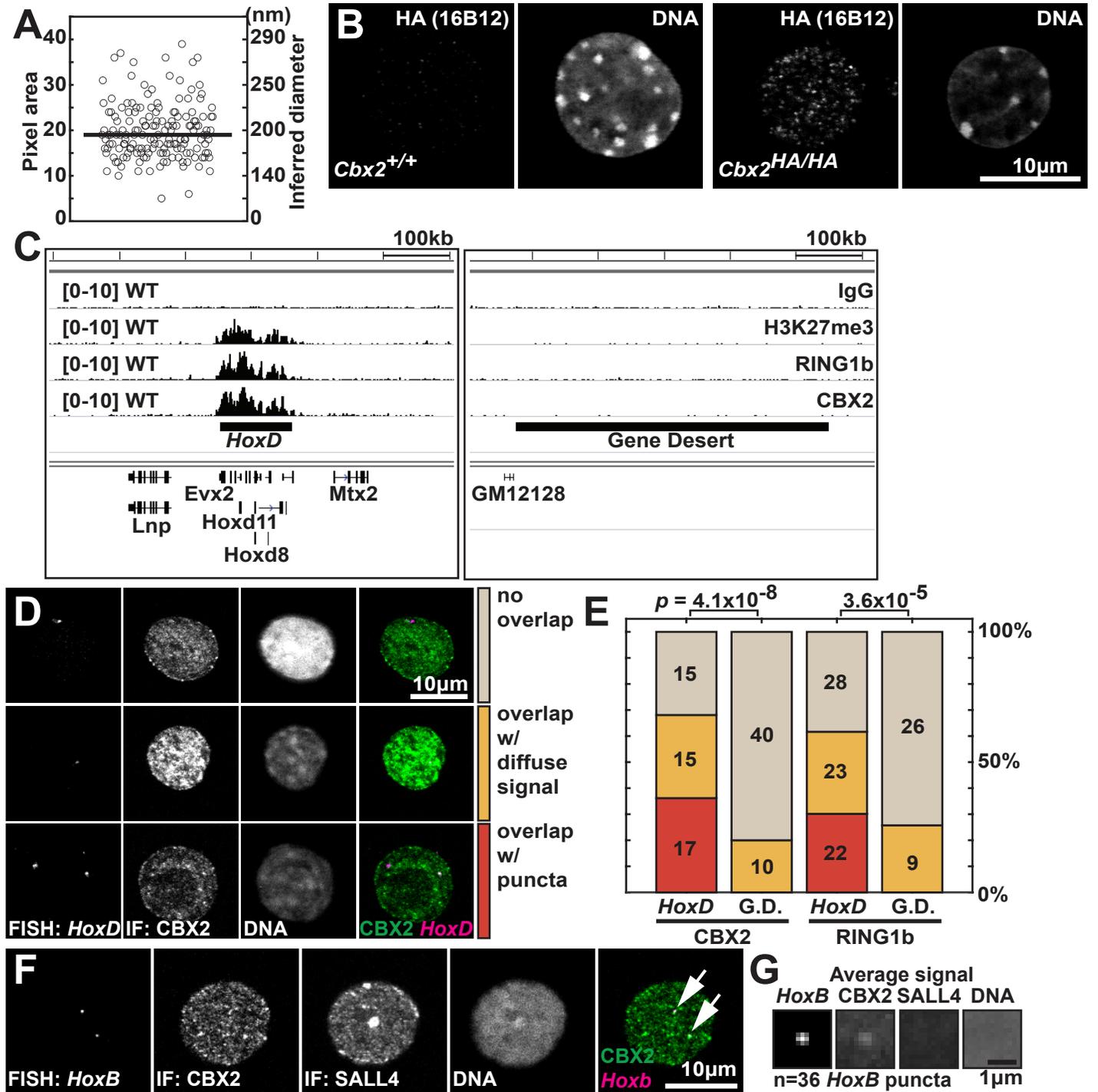


Figure S3. *Hox* loci overlap with CBX2 puncta

- (A) A dot plot representing the pixel size of CBX2 puncta in spermatogonia. Diameters of puncta were calculated based on the image acquisition setting of 1 pixel corresponding to 80 nm.
- (B) Immunostaining of HA epitope using FACS sorted c-KIT(+) spermatogonia from *Cbx2*^{+/+} and *Cbx2*^{HA/HA} mice.
- (C) Genome browser screenshots of CUT&RUN enrichment of IgG, H3K27me3, RING1B, and CBX2 in FACS sorted c-KIT(+) spermatogonia at *HoxD* and gene desert loci.
- (D) Examples with different patterns of overlap between *HoxD* locus and CBX2 by co-immuno-FISH.
- (E) Quantification of different overlap patterns between *HoxD* or gene desert loci with CBX2 or RING1B. *p*-values are based on Fisher's exact test.
- (F) Co-immuno-FISH of *HoxB* locus and CBX2 and SALL4 protein using FACS sorted c-KIT(+) spermatogonia from wild-type mice. A single optical section was represented. Arrows represent *HoxB* loci overlapping with CBX2 puncta.
- (G) Average signal projection of 2 μ m square centered at *HoxB* puncta. Average signal projection of CBX2, SALL4, and DNA at the corresponding locations were also represented.

Fig S4. CBX2 and H3K27me3 are depleted at genes with strong H4K3me3 enrichment

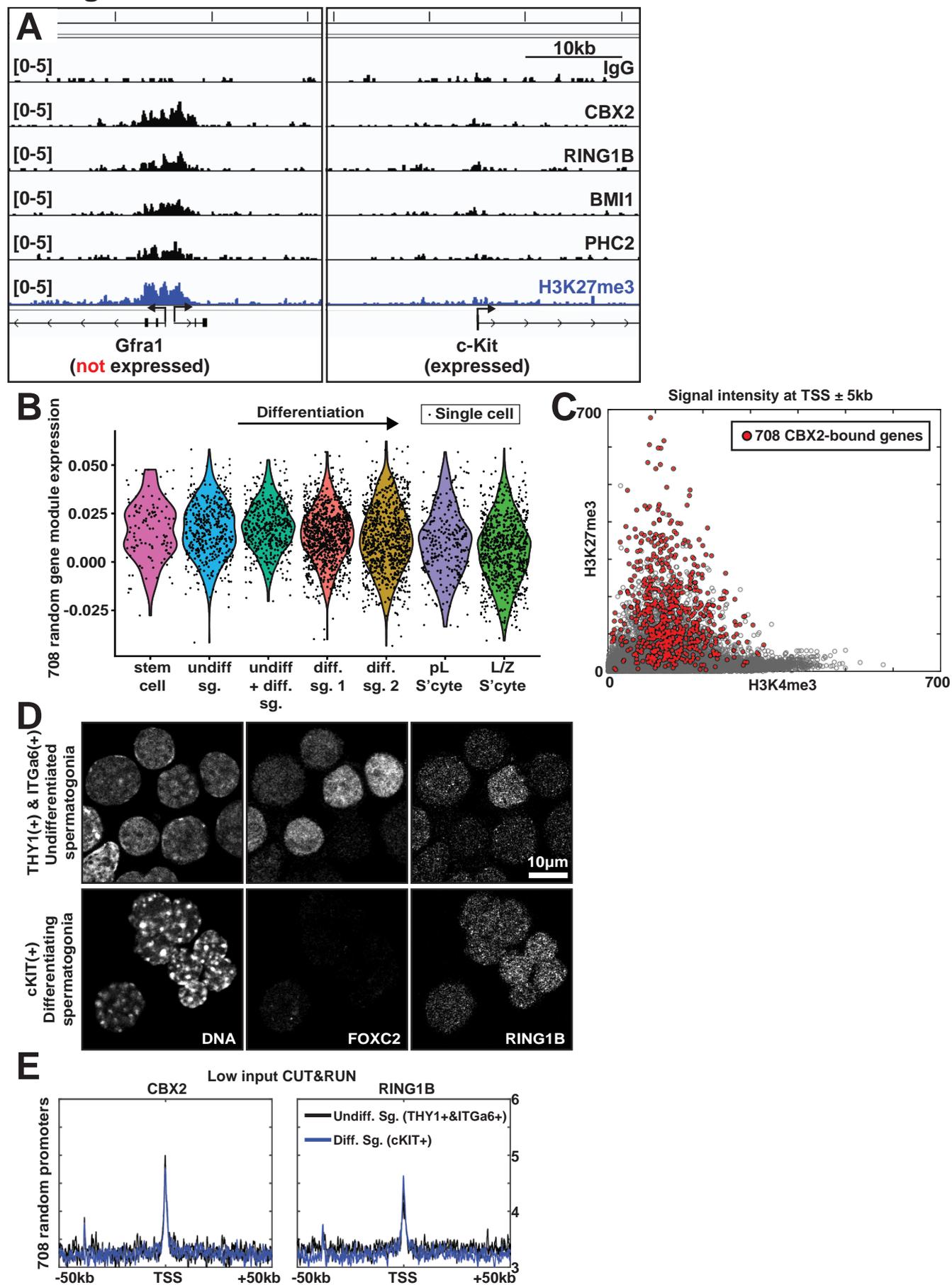


Figure S4. CBX2 and H3K27me3 are depleted at genes with strong H4K3me3 enrichment

(A) Additional example genome browser screenshots of CUT&RUN enrichment of IgG, CBX2, cPRC1 components (RING1B, BMI1, PHC2), and H3K27me3 from FACS sorted c-KIT(+) spermatogonia. Enrichment profiles at *Gfra1* and *c-Kit* genes are shown as representative examples. *Gfra1* is highly expressed in spermatogonial stem cells and downregulated in c-KIT(+) differentiating spermatogonia. In contrast, *c-Kit* is robustly expressed in c-KIT(+) spermatogonia.

(B) Average expression levels of a random module (genes expressed from 708 randomly chosen promoters) per cell in single cell clusters represented in Fig. 2B.

(C) A scatter plot showing the intensity relationship between H3K4me3 and H3K27me3 at 10kb regions encompassing 28,656 transcription start sites (open gray circles). CBX2-bound genes were represented with red-filled circles.

(D) Co-immunostaining of FOXC2 and RING1B using FACS-sorted THY1(+) & ITGa6(+) undifferentiated spermatogonia and c-KIT(+) differentiating spermatogonia from wild-type mice.

(E) Average low-input CUT&RUN enrichment of CBX2 and RING1B at 708 random target promoters used in (B). Profiles for each protein in FACS-sorted undifferentiated (THY1+ & ITGa6+, black lines, two replicates) and differentiating (c-KIT+, blue lines, two replicates) spermatogonia are shown. Signal intensities for the same antibody were TMM-normalized using values from all 27,848 promoters ± 5 kb centered at TSSs.

Fig S5. CBX2 is required for differentiation to CCND2(+) A1 spermatogonia

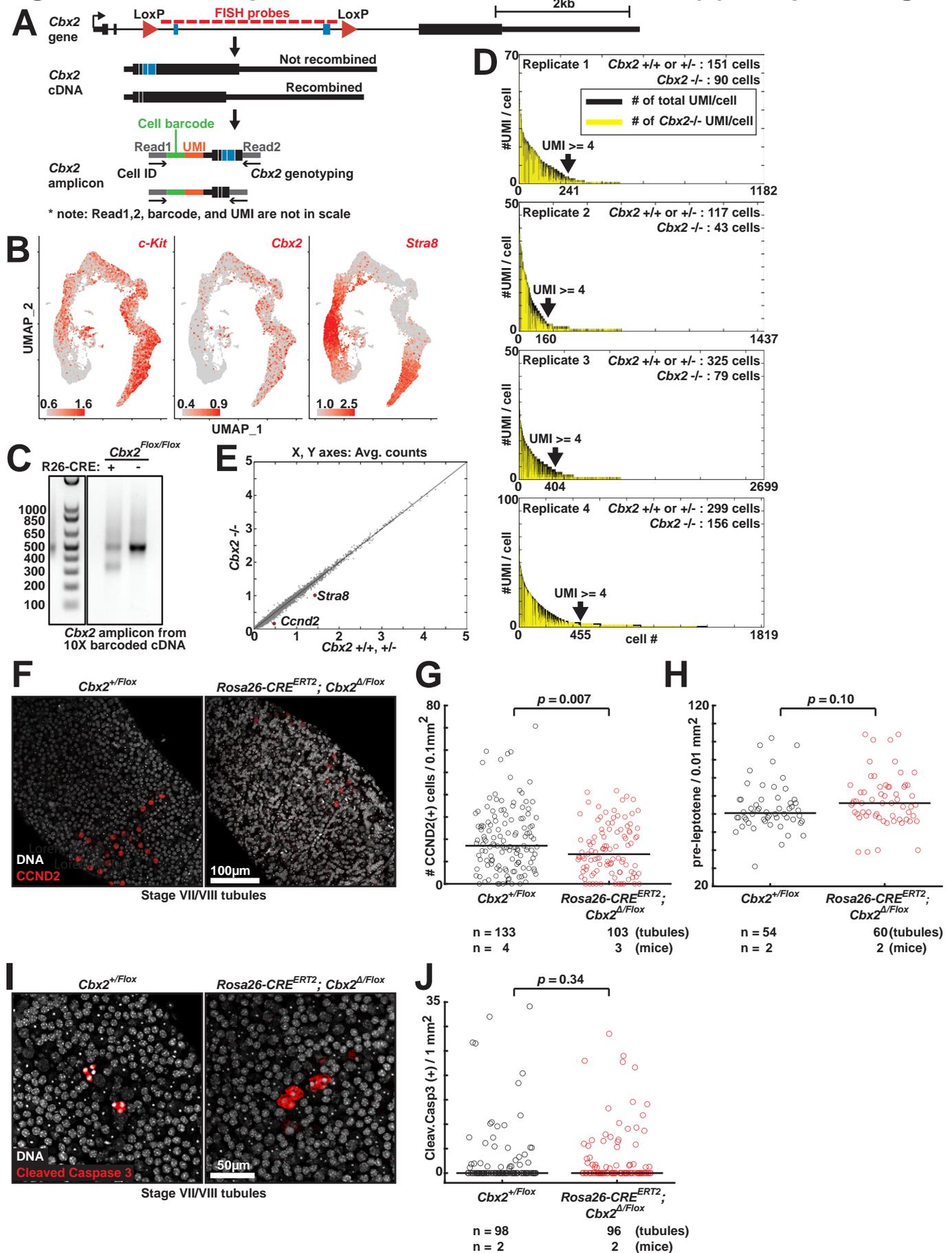


Figure S5. CBX2 is required for differentiation to CCND2(+) A1 spermatogonia

(A) Schematic of genotyping strategy by sequencing *Cbx2* cDNA amplicons with cell barcodes. In the *Cbx2* gene schematic, thick lines represent five *Cbx2* exons (based on NM_007623.3), and thin lines represent introns. Blue lines represent exon3 and exon4 that are deleted when recombined by CRE.

Cbx2 amplicons were amplified using *Cbx2*-specific primers such that the final product included the *Cbx2* exon5 junctions and the cell barcodes.

(B) UMAP representation showing normalized expression levels of *c-Kit*, *Cbx2*, and *Stra8* in FACS sorted c-KIT(+) spermatogonia. The UMAPs represent combined cell profiles from control and *Cbx2* inducible mutant animals (15,362 cells).

(C) A representative DNA gel image of *Cbx2* amplicons generated from 10X cell-barcoded cDNAs from control and *Cbx2* inducible mutant animals.

(D) Distribution of the number of total UMIs detected per cell (black lines) and the number of UMIs associated with *Cbx2* mutation (yellow lines). Data are sorted by the decreasing amount of total UMIs per cell from left to right. Cellular genotypes are assigned only when ≥ 4 UMIs are detected per cell.

(E) A scatter plot comparing averaged gene expression levels of *Cbx2*-expressing cells (population#1-6 in Fig. 5B) between *Cbx2*^{+/+, +/-} and *Cbx2*^{-/-} cells. Wilcoxon rank sum tests (with Bonferroni correction) between 892 *Cbx2*^{+/+, +/-} cells versus 368 *Cbx2*^{-/-} cells identified *Stra8* ($p = 0.0001$) and *Ccnd2* ($p = 0.04$) as the top two downregulated genes, and they are marked by red dots.

(F) Immunofluorescence staining of whole testis tubules of control (*Cbx2*^{+/Flox}) and *Cbx2* inducible mutant (*Rosa26-CRE*^{ERT2}; *Cbx2* ^{Δ /Flox}) animals with CCND2 (red) antibody. Only stage VII/VIII tubules where CCND2(+) A1 spermatogonia reside are analyzed.

(G) Quantification of the number of CCND2(+) A1 spermatogonia per 0.1mm² segments of stage VII/VIII tubule sections. Because A1 spermatogonia are not uniformly distributed on testis tubules, at least 30 different tubule segments were imaged per animal. p -value was calculated by two-tailed t-test.

(H) Quantification of the number of pre-leptotene spermatocytes per 0.01mm² segments of stage VII/VIII tubule sections. p -value was calculated by two-tailed t-test.

(I) Immunofluorescence staining of whole testis tubules of control (*Cbx2*^{+/Flox}) and *Cbx2* inducible mutant (*Rosa26-CRE*^{ERT2}; *Cbx2* ^{Δ /Flox}) animals with cleaved Caspase 3 (red) antibody to mark apoptotic cells. Only stage VI to VIII tubules (where A aligned to A1 transition happens) are analyzed.

(J) Quantification of the number of apoptotic cells with cleaved Caspase 3 signal per 1mm² segments of stage VI to VIII tubule sections. p -value was calculated by Wilcoxon rank sum test.

Fig S6. Estimation of *Cbx2* mutation rate and selection of a expression matched random control gene set

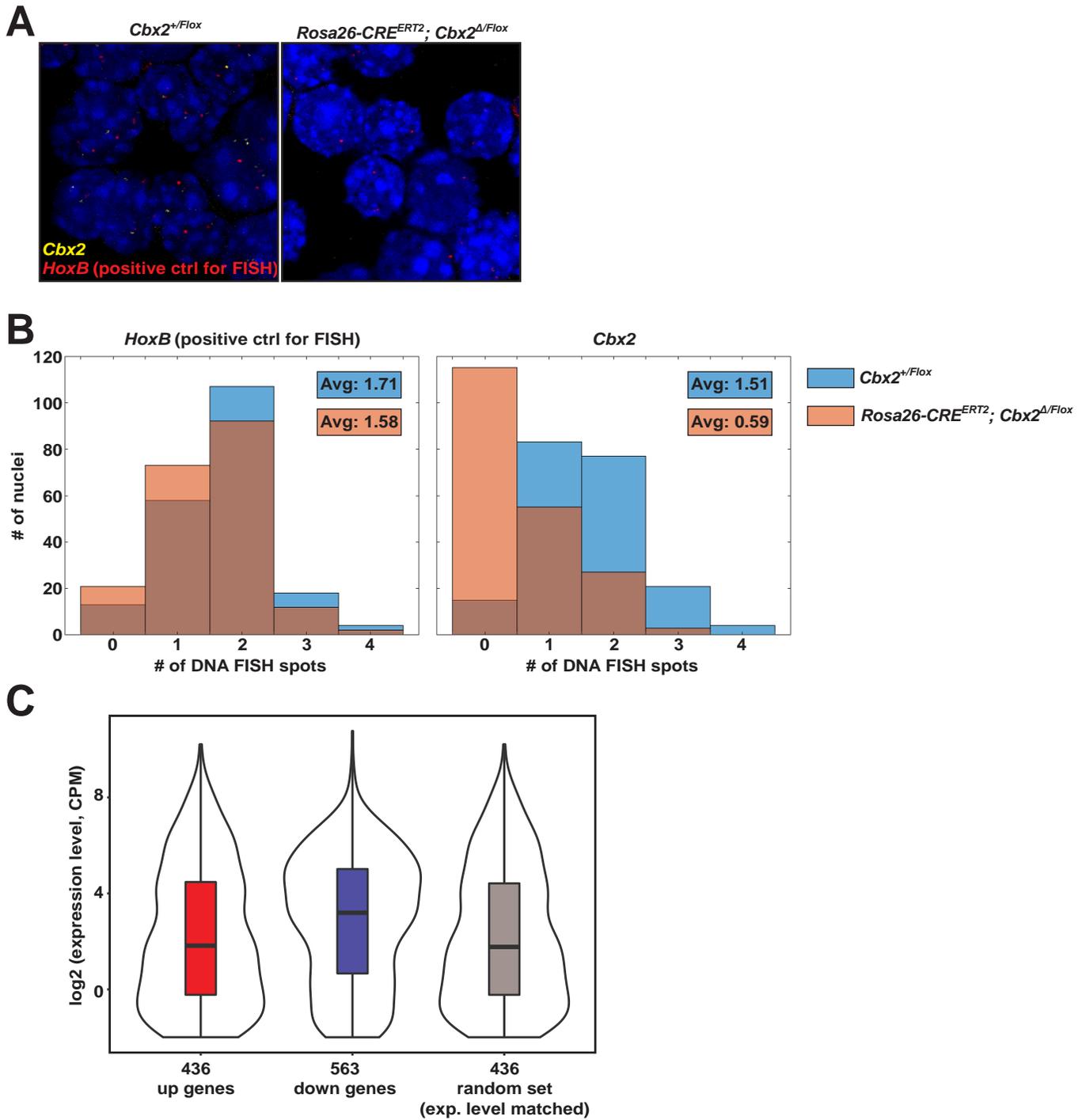


Figure S6. Estimation of *Cbx2* mutation rate and selection of an expression matched random control gene set

(A) SABER-DNA-FISH of *Cbx2* (yellow) and *HoxB* (red, positive control for FISH) using FACS-sorted c-KIT(+) differentiating spermatogonia from control (*Cbx2*^{+/*Flox*}) and *Cbx2* inducible mutant (*Rosa26-CRE^{EERT2}; Cbx2^{Δ/*Flox*}*) mice.

(B) Quantification of (A) by histograms of number of *Cbx2* and *HoxB* puncta per nucleus. 200 nuclei were analyzed per condition. Wild-type diploid cells will have two *Cbx2* spots, and *Cbx2* null cells will have one *Cbx2* spot because FISH can still detect *Cbx2*^Δ allele. Because we do not know the exact detection rate by DNA-FISH and cellular ploidy, we based our analyses on the relative comparison. We observed about a 60% reduction in the average number of *Cbx2* spots after Tamoxifen administration, which we estimate as the *Cbx2* null rate (as the other *Cbx2*^Δ allele is already mutant before injecting Tamoxifen).

(C) Violin plots showing Counts Per Million (CPM) value distributions of 436 upregulated and 563 downregulated genes defined in Fig. 6A. CPM distribution of 436 randomly chosen genes matching expression levels of 436 upregulated genes is also represented.

Fig S7. No significant differences in nuclear and chromatin distribution of CBX2 between wild-type and *Cbx2* CAPS region mutant mice

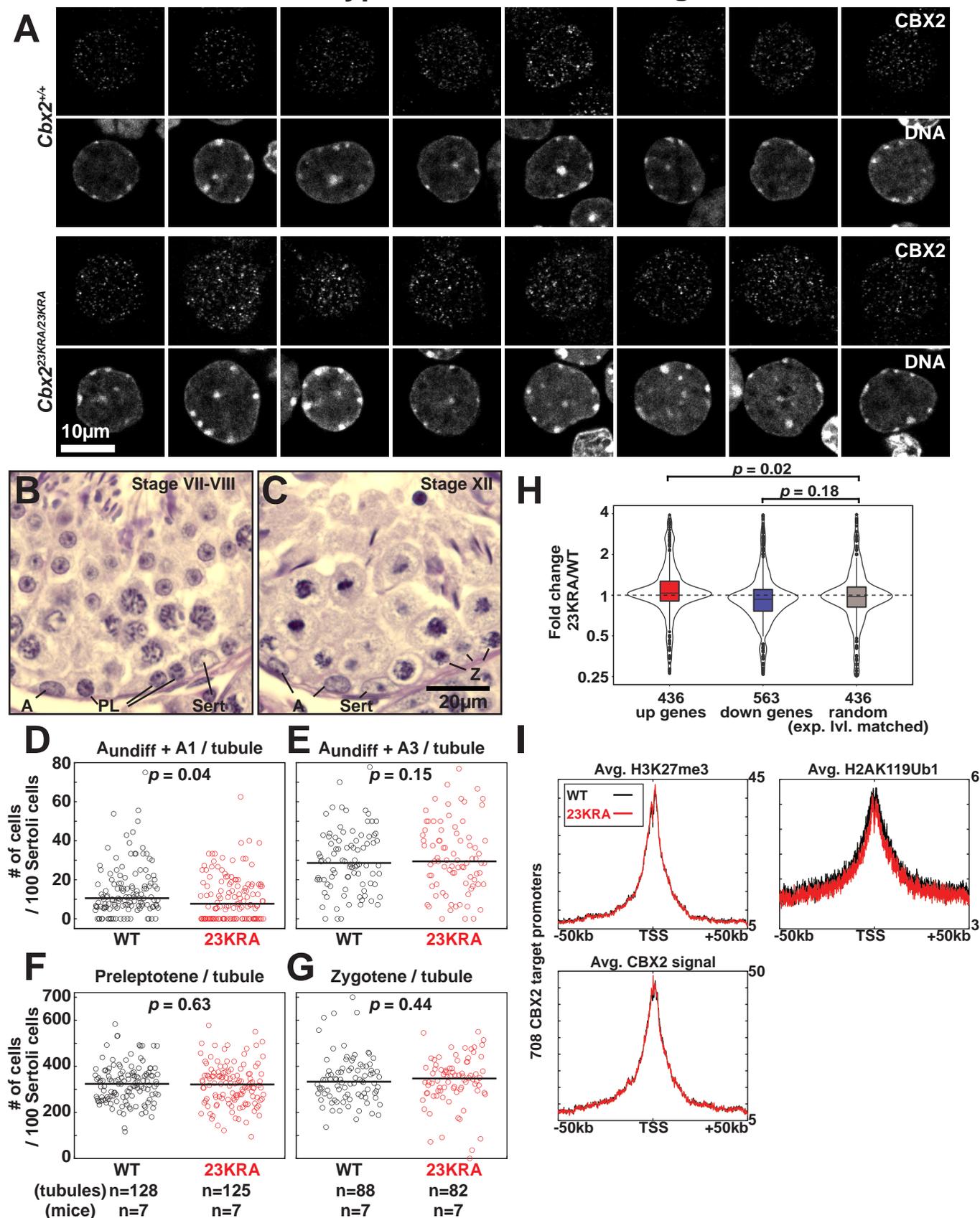


Figure S7. No significant differences in nuclear and chromatin distribution of CBX2 between wild-type and *Cbx2* CAPS region mutant mice

(A) Immunostaining of CBX2 using FACS sorted c-KIT(+) spermatogonia from wild-type and *Cbx2*^{23KRA/23KRA} mice. Spermatogonia at a similar stage between wild-type and *Cbx2*^{23KRA/23KRA} were chosen based on nuclear morphology. Eight representative nuclei per condition were shown.

(B and C) Representative Hematoxylin-PAS staining showing part of tubule sections at (B) stage VII/VIII and (C) stage XII. A (at stage VII/VIII): undifferentiated and A1 spermatogonia; A (at stage XII): undifferentiated and A3 spermatogonia; PL: pre-leptotene spermatocytes; Z: zygotene spermatocytes; Sert: Sertoli cells.

(D-G) Quantification of the number of (D) undifferentiated and A1 spermatogonia, (E) undifferentiated and A3 spermatogonia, (F) pre-leptotene spermatocytes, (G) zygotene spermatocytes in (D and F) stage VII/VIII or (E and G) XII tubules in wild-type and *Cbx2*^{23KRA/23KRA} mice after normalizing cell numbers to the number of somatic Sertoli cells. *p*-value was calculated by Wilcoxon rank sum test. Note, the *p*-value of A_{undiff}+A1 comparison between wild type and *Cbx2*^{23KRA/23KRA} is 0.14 when cell numbers are normalized to tubule circumference.

(H) Violin plots showing gene expression fold changes between wild-type and *Cbx2*^{23KRA/23KRA} mice. Gene sets defined in Fig. 6A as up- and downregulated in *Cbx2*^{KO} mice are represented in comparison to a randomly chosen 436 gene set matching expression levels of upregulated genes. *p*-values are based on Wilcoxon rank sum test.

(I) Average CUT&RUN enrichment of CBX2, H3K27me3 and H2AK119Ub1 using FACS-sorted c-KIT(+) spermatogonia. Signals centered at promoters of 708 CBX2 target promoters in wild-type (black lines) and *Cbx2*^{23KRA/23KRA} mutant (red lines) spermatogonia are represented. Signal intensities for the same antibody were TMM-normalized using values from all 27,848 promoters ±5kb centered at TSSs.