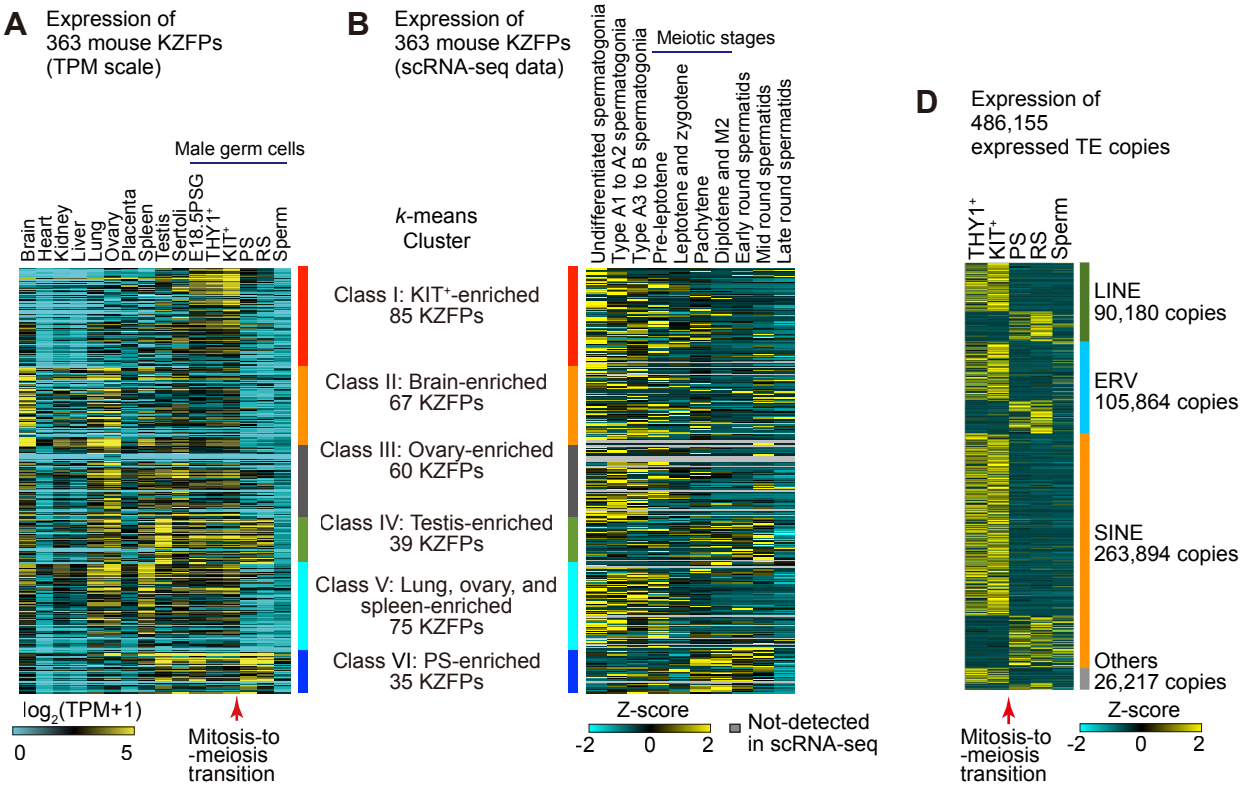
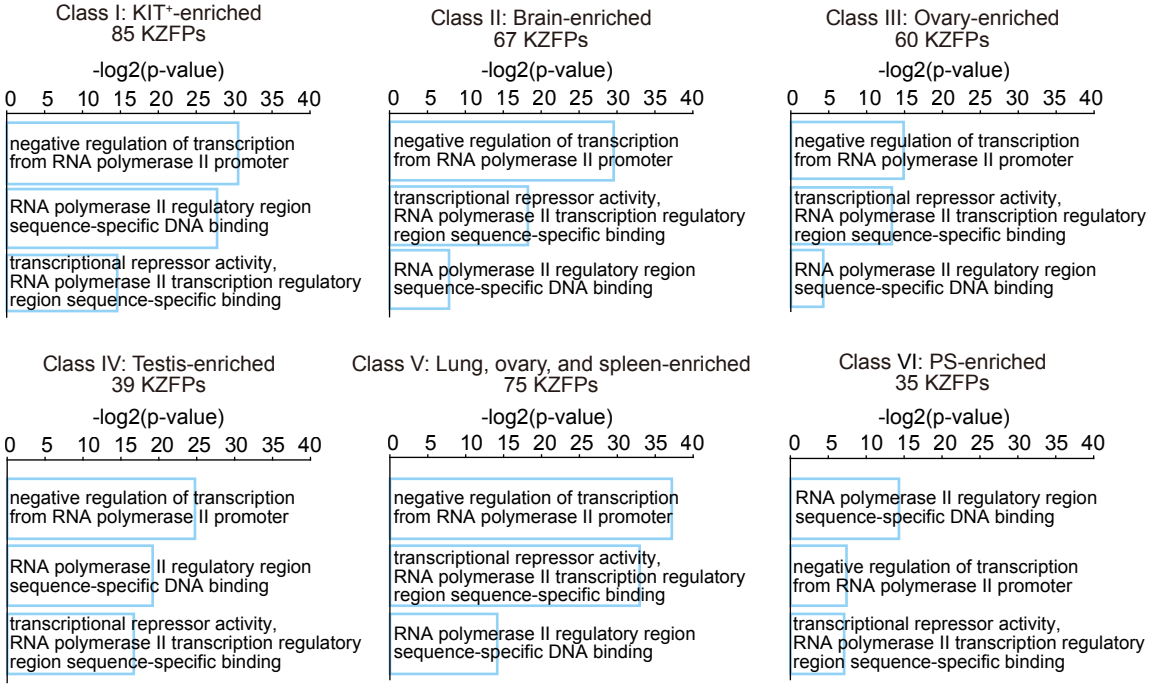
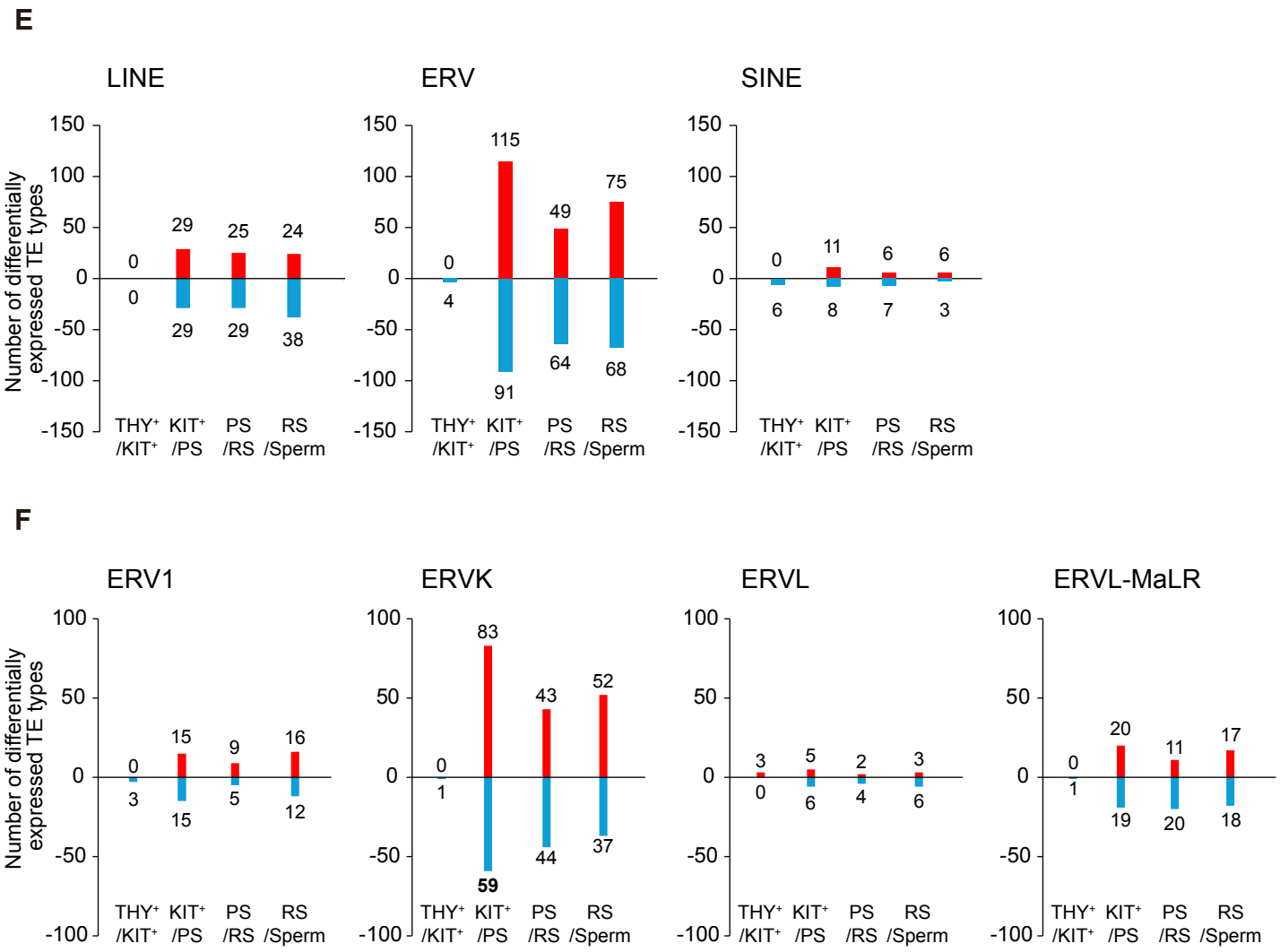
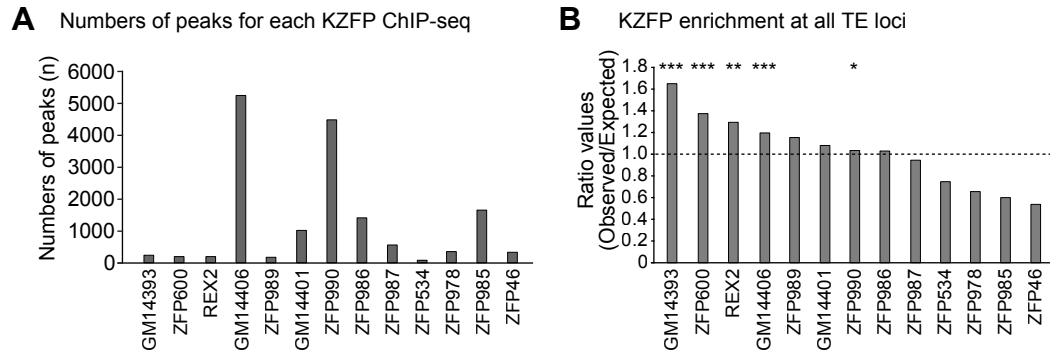


C GO term analysis

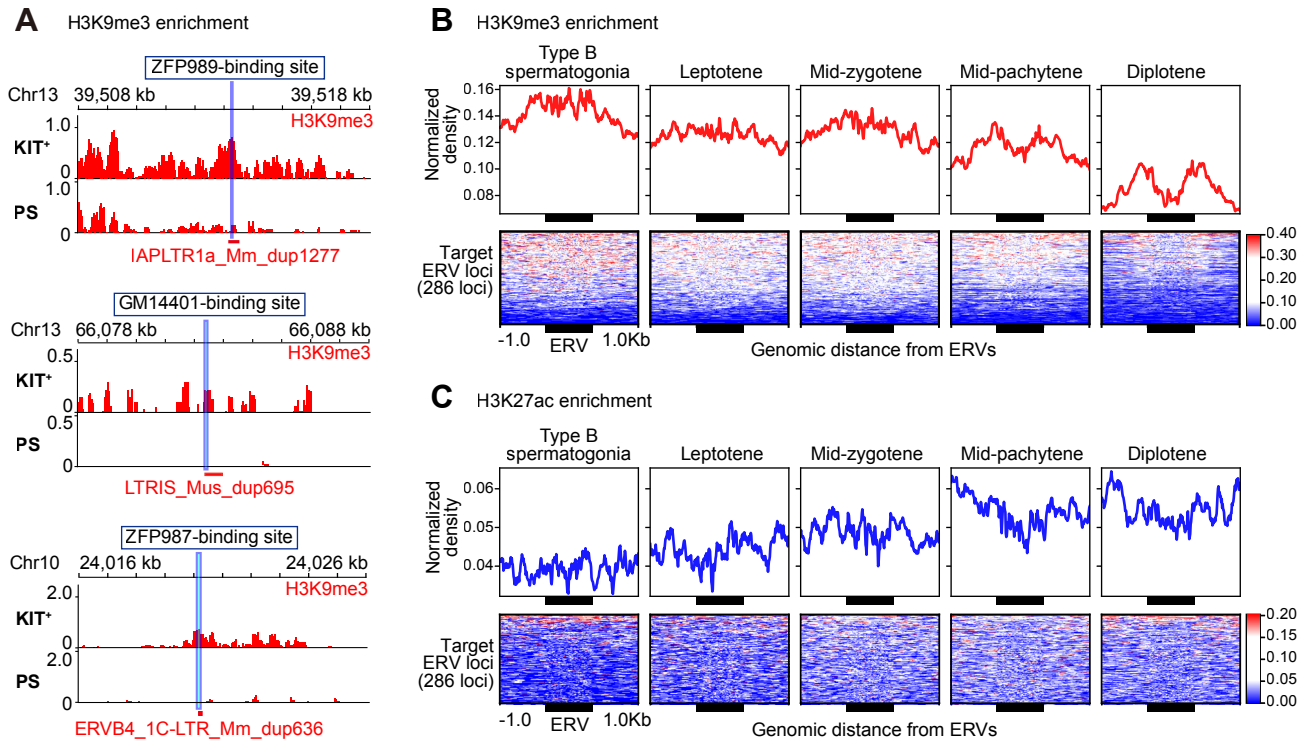




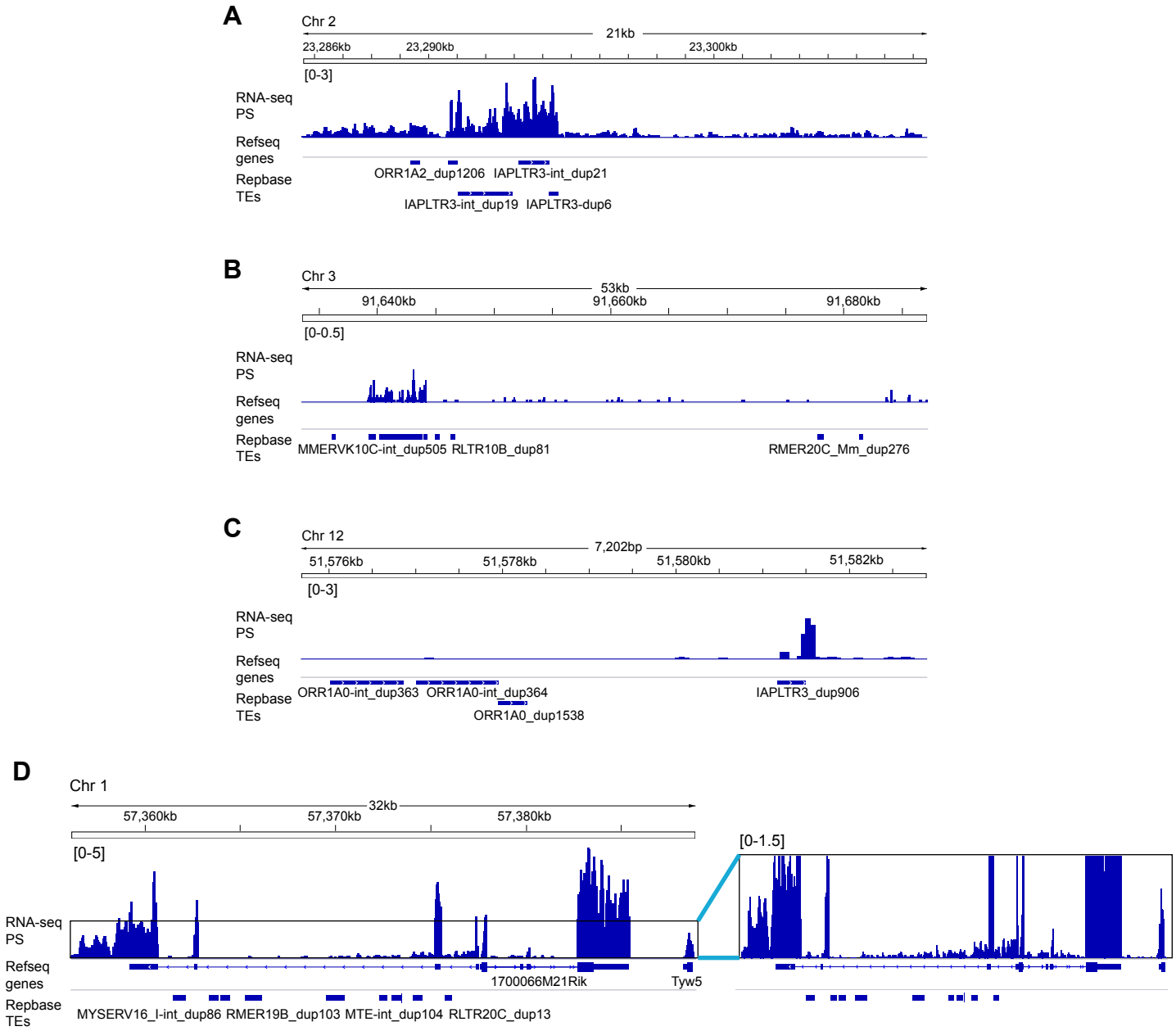
Supplemental_Fig. S1: Expression of KZFPs and TEs during mouse spermatogenesis and GO term analyses.
A. A heatmap showing a *k*-means clustering analysis of expression (RNA-seq data, TPM) for all protein-coding KZFPs in various tissues and testicular cell types in mice.
B. A heatmap showing the expression (single-cell RNA-seq) for KZFPs analyzed in Fig. 1B in various tissues and testicular cell types in mice. Data from adult testis were classified into 10 substages of spermatogenic cells. The expression level was calculated as the median-normalized average at the specific substage.
C. GO term of each KZFP cluster. The bar plot represents the -log2 p-value of each GO term.
D. A heatmap showing the expression (RNA-seq data) of all detected TE copies in male germ cells at the representative five stages.
E. A bar graph representing the number of TEs (LINE, ERV, and SINE) differentially expressed between two consecutive substages. TEs showing upregulation or downregulation are shown by red bars and blue bars, respectively.
F. A bar graph representing the number of ERV types (ERV1, ERVK, ERVL, and ERVL-MaLR) differentially expressed between two consecutive substages. ERVs types showing upregulation or downregulation are shown by red bars and blue bars, respectively.



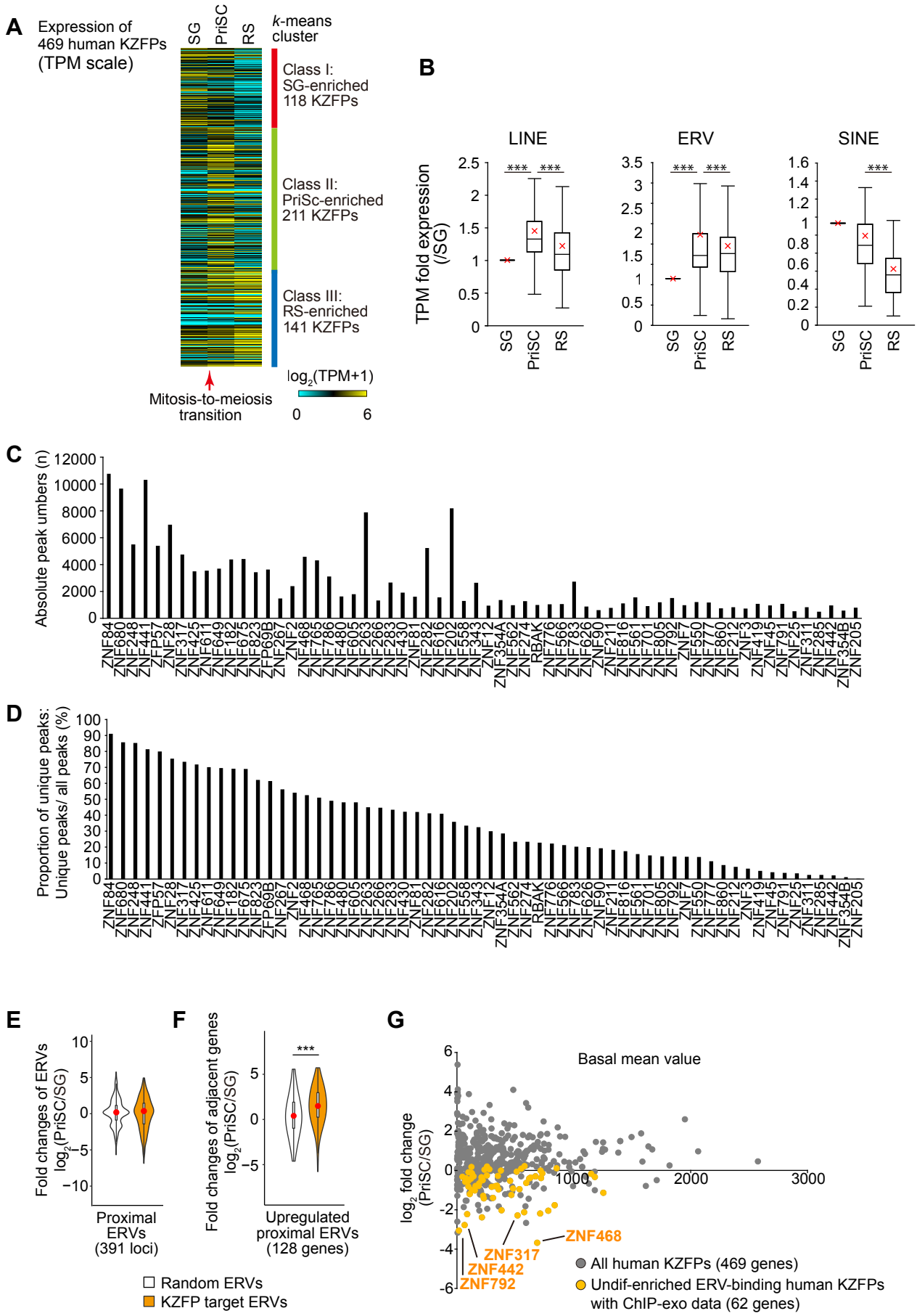
Supplemental_Fig. S2: ChIP-seq analyses of KIT$^+$ spermatogonia-enriched KZFPs in mice.
A. Numbers of all detected ChIP-seq peaks.
B. Binding preference of KZFPs at all TEs. The bar plot represents the ratio values of the observed peak number versus the theoretically expected number. Binomial test. ***: p<0.001, **: p<0.01, *: p<0.05.



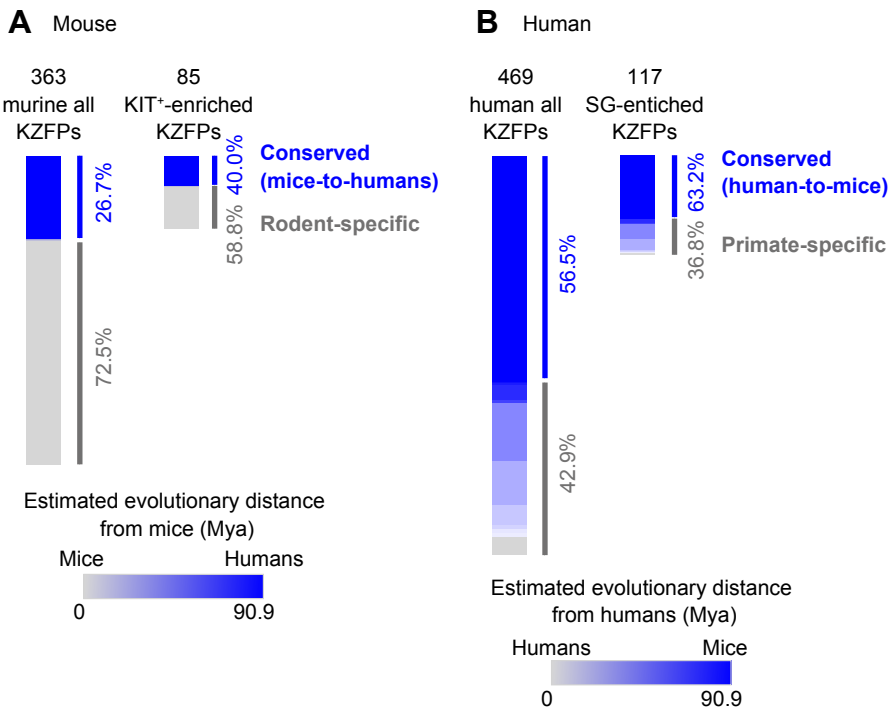
Supplemental_Fig. S3: H3K9me3 and H3K27ac enrichment at the KIT⁺ spermatogonia-enriched KZFP loci in mice.
A. Track views of H3K9me3 enrichment at a binding site of ZFP989 (Top), GM14401 (Middle), and ZFP987 (Bottom) in KIT⁺ spermatogonia and PS.
B, C. The top panels represent an average tag density of H3K9me3 (**B**) and H3K27ac (**C**) at the target ERVs ± 1kb regions. The bottom heatmaps represent normalized H3K9me3 signal intensity.



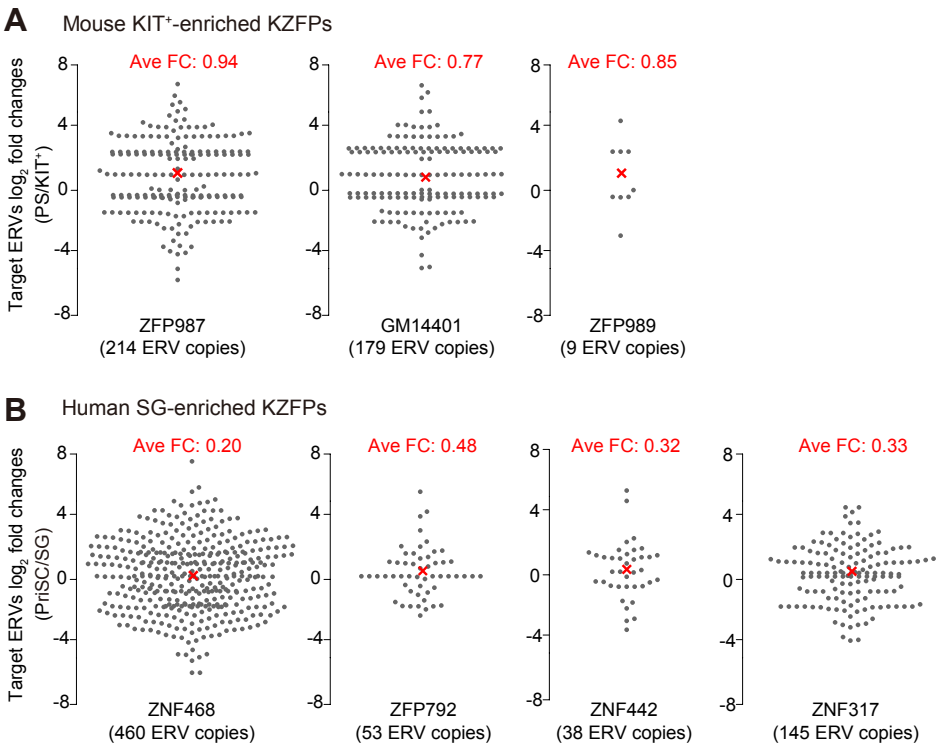
Supplemental_Fig. S4: Genome track view of RNA-seq around TE regions.
A-C. Track views of RNA-seq in PS at intergenic TE loci. Genomic regions around specific TEs (IAPLTR3-int_dup19 (**A**), MMERVK10C-int_dup505 (**B**), and IAPLTR3_dup906 (**C**)) are shown.
D. The track view of RNA-seq in PS at the intronic TE locus. The genomic region around specific TE (RLTR20C_dup13) is shown. The panel on the right side is an enlarged image of the region.


A
Expression of 469 human KZFPs (TPM scale)
SG
PriSC
RS
k-means cluster
Class I: SG-enriched 118 KZFPs
Class II: PriSc-enriched 211 KZFPs
Class III: RS-enriched 141 KZFPs
log2(TPM+1)
0
6
Mitosis-to-meiosis transition
B
LINE
ERV
SINE
TPM fold expression (/SG)
***
C
Absolute peak umbers (n)
D
Proportion of unique peaks: Unique peaks/ all peaks (%)
E
Fold changes of ERVs log2(PriSC/SG)
Proximal ERVs (391 loci)
F
Fold changes of adjacent genes log2(PriSC/SG)
Upregulated proximal ERVs (128 genes)
Random ERVs
KZFP target ERVs
G
Basal mean value
log2 fold change (PriSC/SG)
ZNF468
ZNF317
ZNF442
ZNF792
All human KZFPs (469 genes)
Undif-enriched ERV-binding human KZFPs with ChIP-exo data (62 genes)

Supplemental_Fig. S5: Expression and ChIP-exo analyses of SG-enriched KZFPs in humans.
A. A heatmap showing a *k*-means clustering analysis of expression (RNA-seq data, TPM) for all protein-coding KZFPs in various tissues and testicular cell types in mice.
B. Relative expression (RNA-seq data) of each TE class during human spermatogenesis. Box-and-whiskers plots showing distributions of RNA-seq read data. The central lines represent medians. The upper and lower hinges correspond to the 25th and 75th percentiles. The upper and lower whisker are extended from the hinge to the largest value no further than the 1.5× inter-quartile range (IQR) from the hinge. The red cross in the box plot is the average of the value. The values of SG were set to 1. The red crosses represent the average values. ANOVA followed by Tukey-HSD test. ***: p<0.001.
C. The bar plot represents the number of peaks from previous ChIP-exo data (Imbeault et al. 2017). 62 KZFPs enriched in SG are listed.
D. Unique peak frequency of SG-enriched KZFPs. The bar plot represents the percentage of uniquely observed peaks.
E. The violin plot with a box plot overlay of fold-expression change of target proximal ERVs (less than 1kb from TSSs) at the mitosis-to-meiosis transition (PriSC/SG). The central lines represent medians. The upper and lower hinges correspond to the 25th and 75th percentiles. The upper and lower whisker are extended from the hinge to the largest value no further than the 1.5× inter-quartile range (IQR) from the hinge. The red dots in the box plot are the average of the fold-expression change. Student's *t*-test, not significant.
F. The violin plot with a box plot overlay of fold-expression change of the genes adjacent to upregulated proximal ERVs at the mitosis-to-meiosis transition (PriSC/SG), as described in the legend for panel **E**. Student's *t*-test. ***: p<0.001.
G. Expression and fold change of human KZFPs at the mitosis-to-meiosis transition. 4 KZFPs showing the most drastic change (ZNF792, ZNF442, ZNF317, and ZNF468) were selected for further analysis.

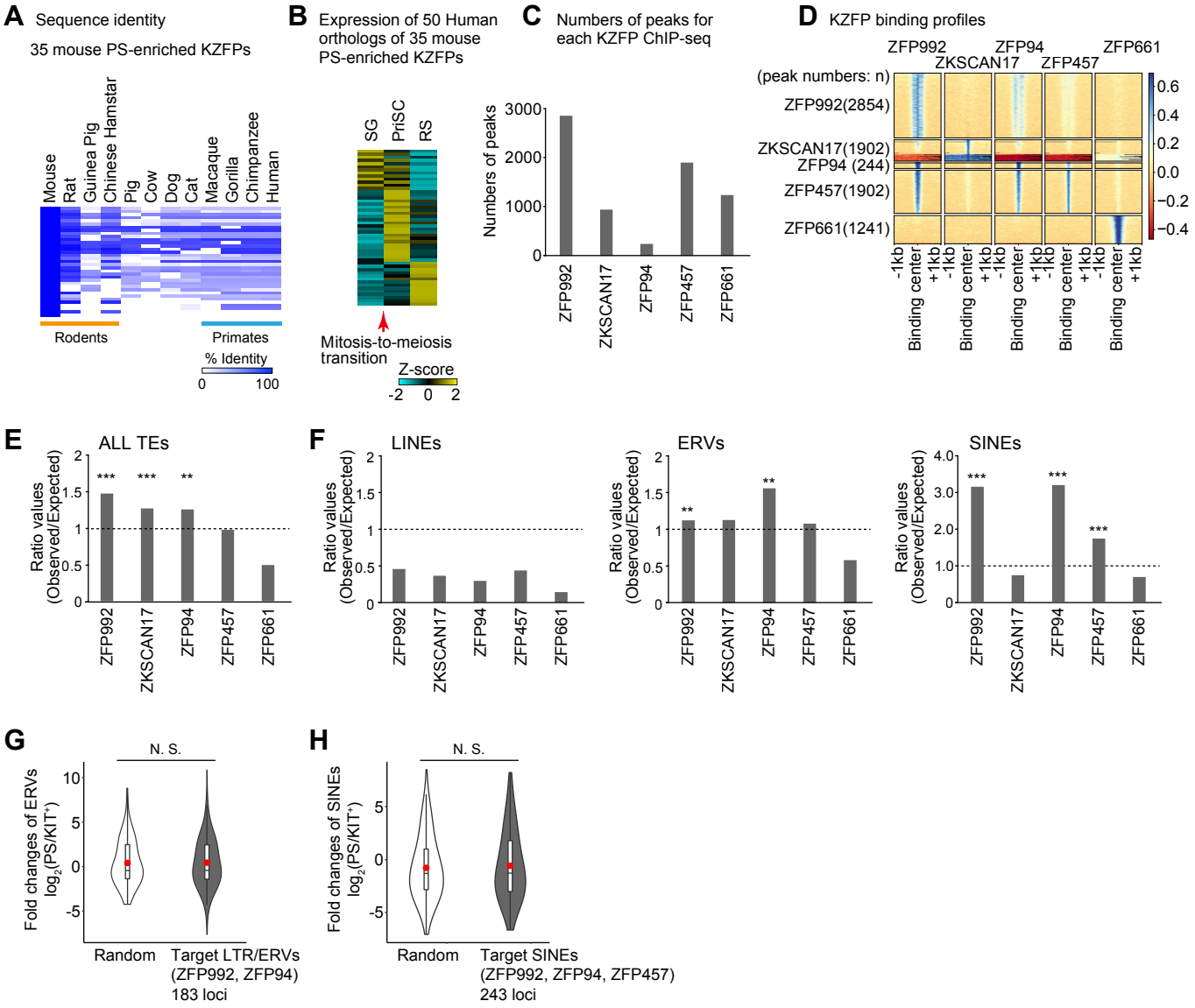


Supplemental_Fig. S6: The evolutionary age of KZFPs enriched in mitotic spermatogonia.
A. The color chart represents the estimated evolutionary distance (Million Years Ago: Mya, from mice) of all 363 murine KZFPs (left) and 85 KIT$^+$-enriched KZFPs (right). Evolutionary distance was calculated based on the previous study (Imbeault et al. 2017).
B. The color chart represents the estimated evolutionary distance (Mya, from humans) of all 469 human KZFPs (left) and 117 SG-enriched KZFPs (right). Evolutionary distance was calculated based on the previous study (Imbeault et al. 2017).



Supplemental_Fig. S7: Expression changes of spermatogonia-enriched KZFPs at the mitosis-to-meiosis transitions in mice and humans.

A, B. Log_2-fold changes of ERVs targeted by mouse (**A**) and human (**B**) spermatogonia-enriched KZFPs. The red crosses represent the average fold change values (shown on top of the panels). The numbers in parentheses represent the detected ERV copy numbers.



Supplemental_Fig. S8: Analyses of mouse PS-enriched KZFPs.

A. Heatmap represents the identity of 35 murine PS-enriched KZFPs compared among mammalian species.

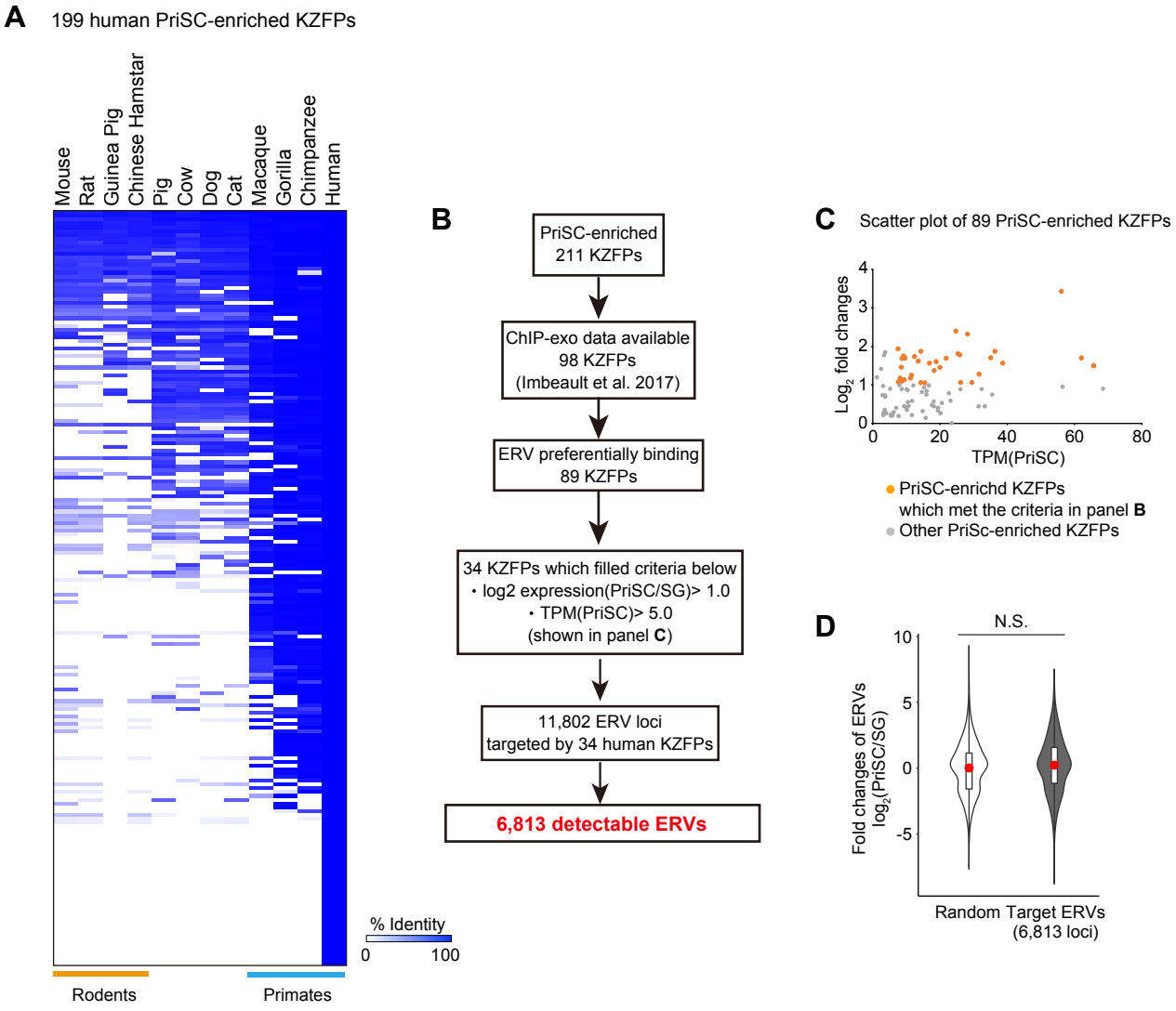
B. Heatmap represents the expression of 50 human orthologous genes.

C. Numbers of all detected ChIP-seq peaks.

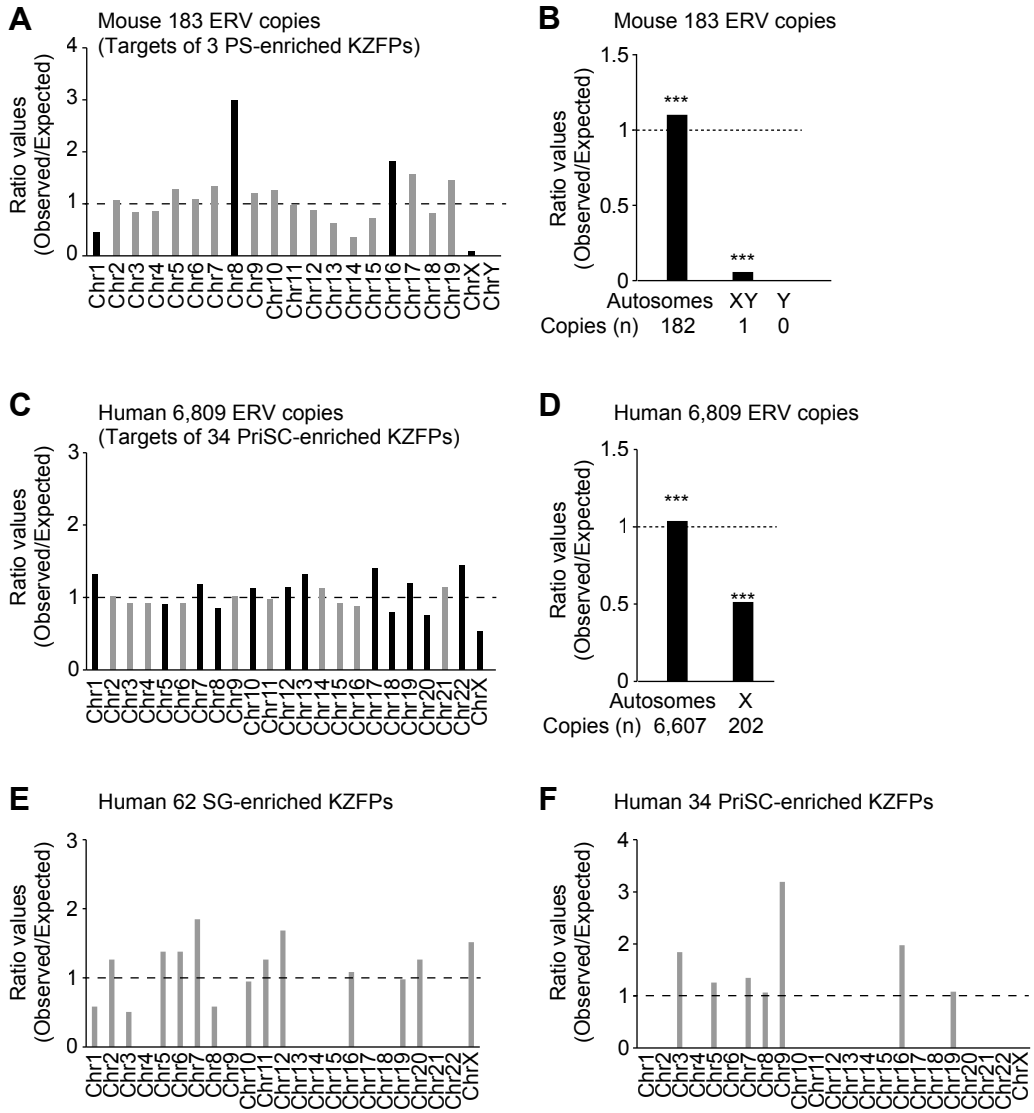
D. Heatmap showing the binding profiles of PS-enriched KZFPs. The numbers in parentheses represent the numbers of detected peaks.

E, F. Binding preference of KZFPs at all TEs (**E**) and each TE subclass (**F**). The bar plots represent the ratio values of the observed peak number versus the theoretically expected number. Binomial test. ***: p<0.001, **: p<0.01.

G, H. The violin plot with a box plot overlay of fold-expression change of ERVs targeted by ZFP992 and ZFP94 (**G**) and SINEs targeted by ZFP992, ZFP94, and ZFP457 (**H**) at the mitosis-to-meiosis transition (PS/KIT$^+$). The central lines represent medians. The upper and lower hinges correspond to the 25th and 75th percentiles. The upper and lower whiskers are extended from the hinge to the largest value no further than the 1.5× inter-quartile range (IQR) from the hinge. The red dots in the box plot are the averages of the fold-expression changes. Student's *t*-test. N.S.: not significant.



Supplemental_Fig. S9: Analyses of human ProSC-enriched KZFPs.
A. The heatmap represents the identity of 199 human KZFPs compared among mammalian species.
B. Flowchart of analyses to identify target ERVs.
C. Scatter plot represents the TPM value (x-axis) and the $\log_2$ fold changes (y-axis) of 89 PriSC-enriched KZFPs whose ChIP-exo data are available.
D. The violin plot with a box plot overlay of fold-changes in expression of target ERVs at the mitosis-to-meiosis transition (PriSc/SG). The central lines represent medians. The upper and lower hinges correspond to the 25th and 75th percentiles. The upper and lower whiskers are extended from the hinge to the largest value no further than the 1.5× inter-quartile range (IQR) from the hinge. The red dots in the box plot are the average of the fold-expression change. Student's *t*-test. N.S.: not significant.



Supplemental_Fig. S10: Chromosomal distributions of KZFPs and ERVs.

A. Chromosomal distribution of mouse ERV copies targeted by PS-enriched KZFPs. The bar plot represents the ratio values of the observed copy number of ERVs versus the theoretically expected number. The color of each bar represents the statistical significance (Black: $p<0.05$, Grey: not significant, binomial test).

B. Distribution of mouse ERV copies targeted by PS-enriched KZFPs on autosomes or sex chromosomes. Binomial test. ***: $p<0.001$.

C. Chromosome distribution of human ERV copies targeted by PriSC-enriched KZFPs. The bar plot is represented the same as panel **A**. Note that the HEK293T model lacks the Y Chromosome.

D. Distribution of human ERV copies targeted by PriSC-enriched KZFPs on autosomes or sex chromosomes. The bar plot is represented the same as panel **B**. Note that the HEK293T model lacks the Y Chromosome.

E. Chromosome distribution of human 62 SG-enriched KZFPs. The bar plot is represented the same as panel **A**.

F. Chromosome distribution of human 34 PriSC-enriched KZFPs. The bar plot is represented the same as panel **A**.