KRAB zinc-finger proteins regulate endogenous retroviruses to sculpt germline transcriptomes and genome evolution

Kai Otsuka, Akihiko Sakashita, So Maezawa, et al.

Genome Res. published online March 12, 2025
Access the most recent version at doi:10.1101/gr.279924.124

KRAB zinc-finger proteins regulate endogenous retroviruses to sculpt germline transcriptomes and genome evolution

Kai Otsuka,[1,2] Akihiko Sakashita,[3,4] So Maezawa,[2] Richard M. Schultz,[1,5] and Satoshi H. Namekawa[1,3]

[1]*Department of Microbiology and Molecular Genetics, University of California, Davis, California 95616, USA;* [2]*Department of Applied Biological Science, Faculty of Science and Technology, Tokyo University of Science, Noda, Chiba 278-8510, Japan;* [3]*Reproductive Sciences Center, Division of Developmental Biology, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio 45229, USA;* [4]*Department of Molecular Biology, Keio University School of Medicine, Tokyo 160-8582, Japan;* [5]*Department of Biology, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA*

As transposable elements (TEs) coevolved with the host genome, the host genome exploited TEs as functional regulatory elements of gene expression. Here we show that a subset of KRAB domain–containing zinc-finger proteins (KZFPs), which are highly expressed in mitotically dividing spermatogonia, repress the enhancer function of endogenous retroviruses (ERVs) and that the release from KZFP-mediated repression allows activation of ERV enhancers upon entry into meiosis. This regulatory feature is observed for independently evolved KZFPs and ERVs in mice and humans, suggesting evolutionary conservation in mammals. Further, we show that KZFP-targeted ERVs are underrepresented on the sex chromosomes in meiosis, suggesting that meiotic sex chromosome inactivation (MSCI) may antagonize the coevolution of KZFPs and ERVs in mammals. Our study uncovers a mechanism by which a subset of KZFPs regulate ERVs to sculpt germline transcriptomes. We propose that epigenetic programming during the transition from mitotic spermatogonia to meiotic spermatocytes facilitates the coevolution of KZFPs and TEs on autosomes and is antagonized by MSCI.

[Supplemental material is available for this article.]

Transposable elements (TEs) are mobile genetic elements that account for a large fraction (~40%–50%) of the mammalian genome (Koito and Ishizaka 2013; Trono 2015). Retrotransposons, a major group of TEs, function by a copy-and-paste mechanism using an RNA intermediate that can cause DNA damage upon integration into the genome, thereby leading to genome instability. In response to TEs, the host genome coevolved various defense mechanisms to silence and control retrotransposons, which include DNA methylation, histone modifications, and piRNA pathways, all of which tightly control expression of TEs in the germline (Zamudio and Bourc'his 2010; Bao and Yan 2012; Di Giacomo et al. 2013; Crichton et al. 2014; Fu and Wang 2014; Ku and Lin 2014; Elbarbary et al. 2016; Zhou et al. 2023). The host genome, however, exploits TEs' requirement to be expressed in the germline by co-opting them to serve as regulatory elements to gain reproductive fitness. Thus, as TEs coevolved with the host genome, the germline arose as the prominent site to maximize mutual fitness (Zamudio and Bourc'his 2010; Elbarbary et al. 2016; Zhou et al. 2023). Nevertheless, there remain major questions about how the activities of TEs are regulated and how TEs coevolved with the host in the germline.

Many TE-derived sequences function as gene regulatory elements, such as promoters and enhancers, to drive tissue- and cell type–specific gene expression (Peaston et al. 2004; Pi et al. 2004; Rebollo et al. 2012; Erwin et al. 2014; Friedli and Trono 2015; Garcia-Perez et al. 2016; Thompson et al. 2016; Chuong et al. 2017; Huang et al. 2017). Genome-wide studies demonstrate that a significant portion of transcription factor (TF)-binding sites is derived from TEs (Rebollo et al. 2012; Sundaram et al. 2014). Comprising ~10% of the mammalian genome, endogenous retroviruses (ERVs) are subfamilies of TEs and remnants of retroviruses integrated into the genome (Mouse Genome Sequencing Consortium 2002). Most ERVs in the mammalian genome are truncated but remain as long terminal elements (LTRs).

Testis-specific expression of ERVs in humans and mice was initially reported over 40 years ago (Del Villano and Lerner 1976), and recent studies reveal regulatory functions for TEs in male meiosis. These functions include post-transcriptional regulation of mRNA and long noncoding RNAs (lncRNAs) via the piRNA pathway (Watanabe et al. 2015), as well as promoter functions for ERVs that drive lncRNA expression (Davis et al. 2017; Gill et al. 2023). Our recent study demonstrated that ERVs function as active enhancers to drive expression of species-specific germline genes during the transition from mitotic spermatogonia (SGs) to meiotic spermatocytes (Sakashita et al. 2020), a critical developmental transition during which both the chromatin state and epigenomic modifications are altered (Sin et al. 2015; Maezawa et al. 2018b; Alavattam et al. 2019; Patel et al. 2019; Wang et al. 2019). How the activities of ERVs are dynamically regulated during spermatogenesis is poorly understood.

Corresponding author: snamekawa@ucdavis.edu
Article published online before print. Article, supplemental material, and publication date are at https://www.genome.org/cgi/doi/10.1101/gr.279924.124.

In this study, we address the regulatory mechanism of ERV enhancers to generate germline transcriptomes by focusing on rapidly evolving ERVs and Krüppel-associated box (KRAB) domain–containing zinc-finger proteins (KZFPs), a group of TFs that have coevolved to bind and repress expression of ERVs (Tadepally et al. 2008; Ecco et al. 2017; Yang et al. 2017; Bruno et al. 2019; Wolf et al. 2020; Senft and Macfarlan 2021). KZFPs bind DNA using tandem arrays of C2H2 zinc-finger domains, each recognizing two to four specific nucleotides, thereby recognizing specific DNA sequences with high affinity (Patel et al. 2018). There are hundreds of ERV subfamilies and KZFPs in mammals, with many pairwise functional interactions between them (Imbeault et al. 2017), suggesting an evolutionary arms race between the host genome and viral-derived elements (Ecco et al. 2017; Yang et al. 2017; Bruno et al. 2019; Senft and Macfarlan 2021). Although previous studies revealed that KZFPs function to repress ERV activity in cultured cells by establishing repressive epigenetic marks, such as H3K9me3 via TRIM28 (also known as KAP1) recruitment (Imbeault et al. 2017; Kumar et al. 2020), in vivo functions of KZFPs remain elusive (Wolf et al. 2020; Pontis et al. 2022). In this study, we address how the activities of ERVs and TEs are dynamically regulated during spermatogenesis and how they have coevolved through epigenetic programming during spermatogenesis and meiotic sex chromosome inactivation (MSCI), an essential process in the male germline (Turner 2015; Alavattam et al. 2022).

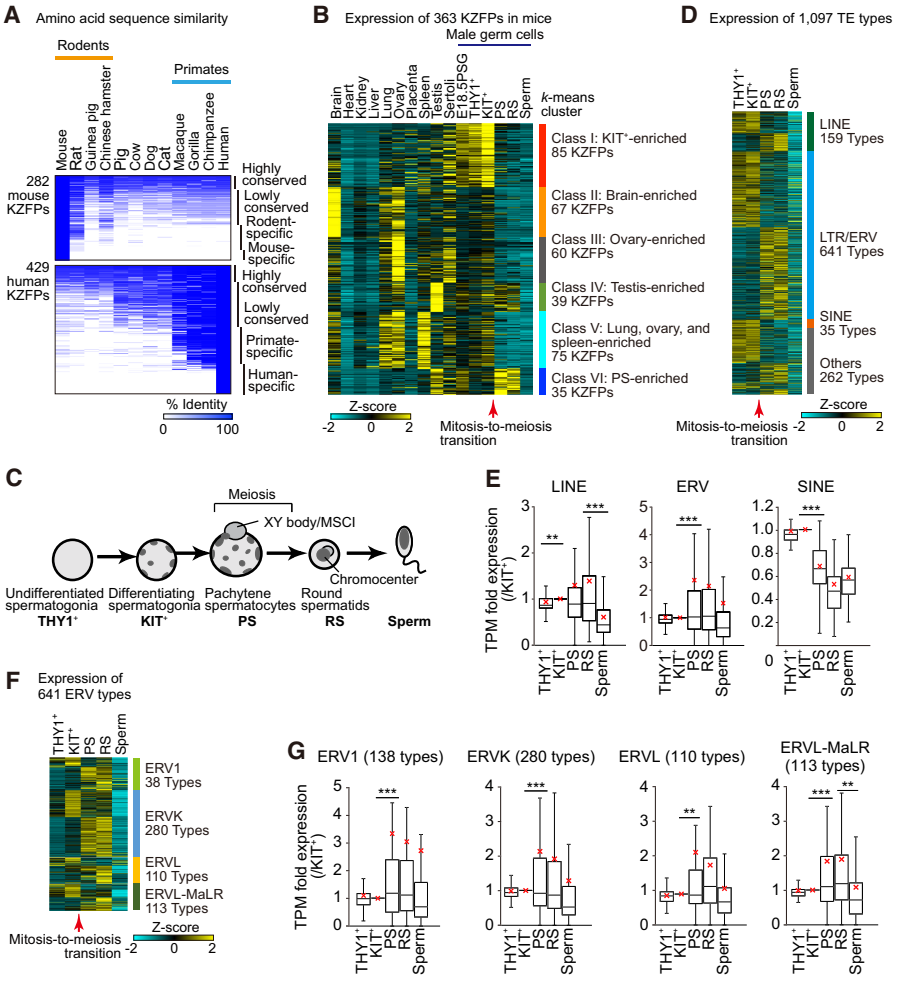

Figure 1. Expression of KZFPs and TEs during mouse spermatogenesis. (*A*) Amino acid sequence identities of 282 murine KZFPs (*top*) and 429 human KZFPs (*bottom*) among 12 mammalian species. All mouse and human KZFPs registered in the BioMart database (282 for mice and 429 for humans) were analyzed. (*B*) A heatmap showing a *k*-means clustering analysis of expression (RNA-seq data) for all protein-coding KZFPs in various tissues and testicular cell types in mice. (E18.5PGC) Prospermatogonia from E18.5 testis, (THY1$^+$) THY1$^+$ undifferentiated spermatogonia, (KIT$^+$) KIT$^+$ differentiating spermatogonia, (PS) pachytene spermatocytes, and (RS) round spermatids. (*C*) Schematic of mouse spermatogenesis and the five representative stages. (*D*) A heatmap showing the expression (RNA-seq data) of all detected TE types in male germ cells at the representative five stages. (*E*,*G*) Relative expression (RNA-seq data) of each TE class (*E*) and each ERV subclass (*G*) during spermatogenesis. Box-and-whisker 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* whiskers are extended from the hinge to the largest value no further than the 1.5× inter-quartile range (IQR) from the hinge. The values of KIT$^+$ were set to one. The red crosses represent the average values. Pairwise permutation ANOVA test (nonparametric test). (**) $P<0.01$, (***) $P<0.001$. (*F*) A heatmap showing the expression (RNA-seq data) of all ERV types in mouse spermatogenesis.

Results

Expression of KZFPs and ERVs changes dynamically during spermatogenesis in mice

Because the KZFP family of proteins evolved rapidly in mammals and contributed to the evolution of gene regulatory networks in mammals (Imbeault et al. 2017; Senft and Macfarlan 2021), we first evaluated the amino acid sequence similarity of mouse and human KZFP proteins across mammalian species using the BioMart web tool. Notably, a subset of mouse KZFPs was present only in mice, and the other subset of mouse KZFPs were rodent specific, whereas others shared low sequence homologies with their counterparts in other mammals (Fig. 1A). In addition to these rapidly evolved KZFPs, a small group of KZFPs was highly conserved among mammals (Fig. 1A). The same trend was also observed for human KZFPs, including human-specific, primate-specific, lowly conserved, and highly conserved KZFPs (Fig. 1A). These analyses, together with a recent review (Senft and Macfarlan 2021), further highlight the rapid evolution of KZFPs, as well as a subset of KZFPs specific to evolutionary orders.

Because rapid gene evolution is a hallmark of germline genes expressed in testis (Soumillon et al. 2013), we suspected that rapidly evolving KZFP genes are expressed during spermatogenesis. Furthermore, our recent study demonstrated dynamic expression of ERVs during spermatogenesis (Sakashita et al. 2020), raising the possibility that changes in expression of their counterpart KZFPs occur during spermatogenesis. Accordingly, we analyzed

expression of 363 mouse protein-coding KZFPs using RNA sequencing (RNA-seq) data sets from nine tissues, Sertoli cells, and germ cells at representative stages of spermatogenesis (Fig. 1B). For the analysis of male germ cells, we examined representative stages such as THY1$^+$ undifferentiated SGs, which contain a stem cell population; KIT$^+$ differentiating SGs, which have undergone spermatogenic differentiation; pachytene spermatocytes (PSs), which are in meiotic prophase I; and haploid round spermatids (RS) (Fig. 1C). We were able to classify the 363 murine protein-coding KZFPs into six major classes based on nonhierarchical *k*-means clustering analysis (Fig. 1B; Supplemental Table 1). Class I KZFPs include 85 KZFPs that were highly expressed in mitotically dividing male germ cells but suppressed upon entry into meiosis. We also detected classes of KZFPs that were highly enriched in specific tissues (Class II: brain-enriched 67 KZFPs; Class III: ovary-enriched 60 KZFPs, which were also highly expressed in Sertoli cells; Class IV: testis-enriched 39 KZFPs; and Class V: lung, ovary, and spleen-enriched 75 KZFPs). Classes II, III, and V were downregulated during the transition from mitotic SGs to meiotic spermatocytes. In contrast, Class VI KZFPs, comprising 35 PS-enriched KZFPs, were upregulated when germ cells reached the pachytene stage. Thus, nearly 90% of KZFPs undergo dynamic expression changes during the transition from SGs to spermatocytes in the male germline. We further evaluated the expression of KZFPs by absolute expression (TPM)–based analysis and confirmed the tissue-specific expression in each class (Supplemental Fig. S1A).

Next, we examined changes in KZFP expression during different stages of meiosis using published single-cell RNA-seq data (Supplemental Fig. S1B; Hermann et al. 2018). Class I, II, III, and V KZFPs were downregulated by the leptotene and zygotene stages of meiotic prophase I, whereas expression of Class VI KZFPs was observed from the pachytene stage. Thus, the transition occurs during early meiotic prophase I. Although the expression patterns of these KZFPs are dynamic, Gene Ontology (GO) term analysis revealed that all classes of KZFPs are enriched with gene functions that share common features, such as negative regulation of transcription and transcriptional repressor activity, as expected from previous studies (Supplemental Fig. S1C; Imbeault et al. 2017; Kumar et al. 2020). A limitation of GO analysis, however, is that it only represents known functions of genes and, therefore, cannot be used to infer novel functions.

Because KZFPs are thought to regulate expression of TEs (Imbeault et al. 2017; Senft and Macfarlan 2021), we next evaluated the dynamics of TE expression at representative stages of spermatogenesis by analyzing the expression of each unique TE copy. We detected 486,155 uniquely expressed TE copies during spermatogenesis (Supplemental Fig. S1D), and these unique copies were categorized into 1097 TE types (representative of TE copies, such as RLTR10B), including 159 types of long interspersed nuclear elements (LINEs), 641 types of LTR/ERVs (hereafter abbreviated as ERVs), and 35 types of short interspersed nuclear elements (SINEs) (Fig. 1D). Consistent with our recent report (Sakashita et al. 2020), expression of TEs changed dynamically during the transition from mitotic SGs to meiotic spermatocytes (Fig. 1D). Notably, we found that overall expression of ERVs was significantly increased in PSs compared with KIT$^+$ SGs (Fig. 1E), whereas expression of SINEs was largely downregulated during the transition from mitotic SGs to meiotic spermatocytes (Fig. 1D,E). The number of differentially expressed TE types was increased during the transition from KIT$^+$ SGs to PSs, and the number of upregulated ERV types exceeded the number of downregulated ERV types (Supplemental Fig. S1E). The overall upregulation of ERVs in PSs mirrors the overall downregulation of KZFPs in PSs, suggesting a mechanistic link. Therefore, we further investigated expression of ERV types in four representative ERV families: ERV1, ERVK, ERVL, and ERVL-MaLR. We found that expression of ERV types in all ERV families was significantly increased during the transition from mitotic SGs to meiotic spermatocytes (Fig. 1F,G). Among ERV families, the number of upregulated ERVK types exceeded the number of downregulated ERVK types (Supplemental Fig. S1F). These results suggest that ERVs are generally suppressed during the premeiotic stage and are then expressed during the meiotic stage.

KIT$^+$ SG-enriched KZFPs target-specific TEs in mice

Because KZFPs are highly expressed in premeiotic cells and because TEs, particularly ERVs, become active in meiotic cells, we hypothesized that premeiotic KZFPs suppress expression of TEs and that downregulation of KZFPs leads to activation of TEs in meiotic cells. To test this hypothesis, we first determined possible interactions between KZFPs and TEs by reanalyzing ChIP-seq data for 61 murine KZFPs in cultured cell lines, including embryonic stem cells (ESCs) and embryonic carcinoma cells (Wolf et al. 2020). Because of the sequence specificity of KZFP binding—KZFP binding sites are considered to be conserved among various cell types (Imbeault et al. 2017; Wolf et al. 2020)—this strategy should identify common binding sites of KZFPs present in KIT$^+$ SGs. Among 61 KZFPs, 13 belong to KIT$^+$ SG-enriched Class I KZFPs. Accordingly, we examined their binding sites and determined their binding enrichment to TEs across the mouse genome. First, we detected peaks for the 13 KZFPs from the ChIP-seq data according to previously described criteria (Supplemental Fig. S2A; Wolf et al. 2020). Among the 13 KZFPs, several KZFPs, including GM14406, GM14401, and ZFP987, showed unique binding sites that are not shared by other members of this group (Fig. 2A). We next examined whether these peaks were enriched at TEs in the genome. Among the 13 KZFPs, GM14393, ZFP600, REX2, GM14406, and ZFP990 showed significant enrichment at all TE loci compared with random loci (Supplemental Fig. S2B). Enrichment to TEs, however, is class specific; GM14406 at LINEs (Fig. 2B), ZFP989, GM14401, and ZFP987 at ERVs (Fig. 2C), and GM14393, ZFP990, and ZFP986 at SINEs (Fig. 2D). These results demonstrate the target specificity of KIT$^+$-enriched KZFPs.

A general function of KZFPs is to repress transcription by promoting H3K9me3 formation at the KZFP binding sites (Schultz et al. 2002; Groner et al. 2010; Bruno et al. 2019). Therefore, we next examined whether KZFP binding sites are enriched with H3K9me3 in KIT$^+$ SGs. Consistent with KZFP enrichment at specific classes of TEs, we observed significant enrichment of H3K9me3 at GM14406-bound LINEs (Fig. 2E) and at ZFP989-bound ERVs (Fig. 2F). H3K9me3 enrichment was also observed at other KZFP-bound ERVs, including GM14406, GM14401, ZFP986, and ZFP985 (Fig. 2F). Such H3K9me3 enrichment was not observed on KZFP-bound SINEs, except for REX2-bound SINEs (Fig. 2G), noting that REX2 did not preferentially associate with SINEs in the genome (Fig. 2D). These results raise the possibility that KIT$^+$-enriched KZFPs independently recognize specific classes of TEs and establish H3K9me3 at the target loci before germ cells enter meiosis.

Correlation between expression of target ERVs and adjacent genes in mice

To further elucidate how KIT$^+$-enriched KZFPs regulate target ERVs, we focused on three ERV-targeting KZFPs, ZFP989, GM14401, and

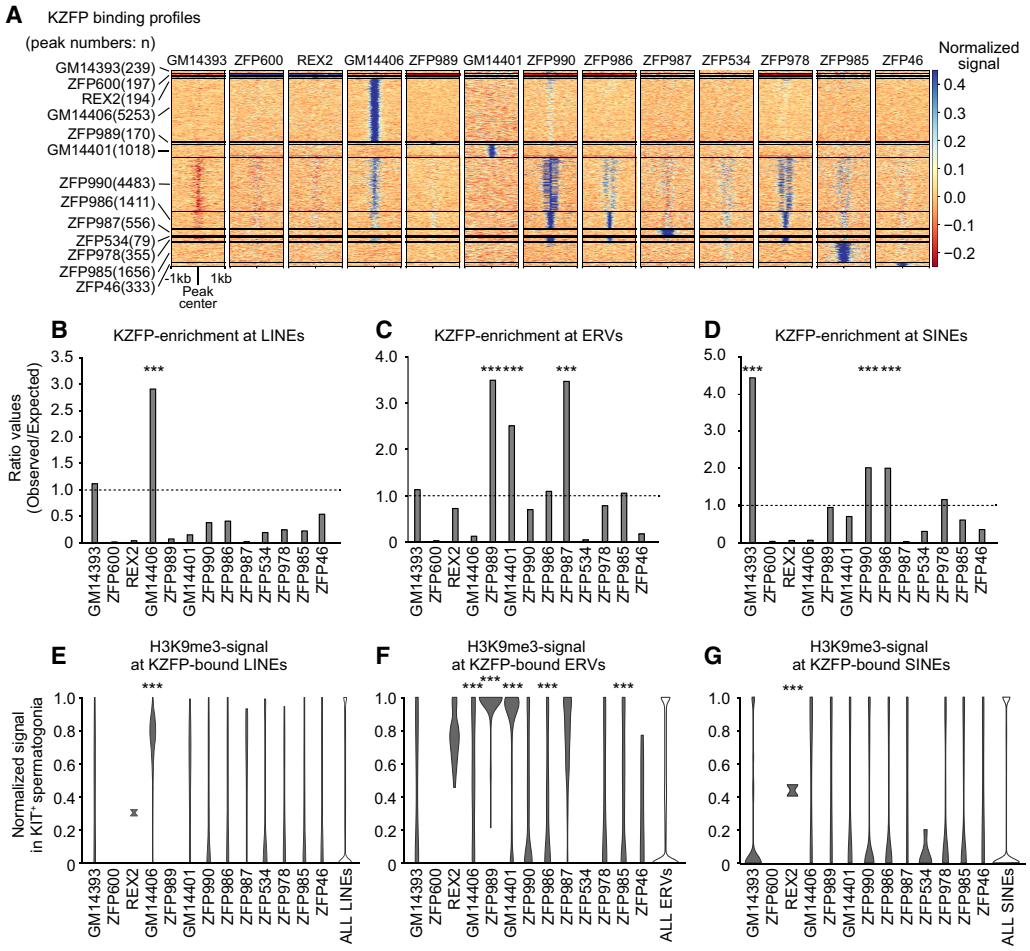

Figure 2. Binding preference of KIT$^+$ spermatogonia-enriched KZFPs and the H3K9me3 enrichment. (*A*) Heatmap of the binding profiles of KIT$^+$ spermatogonia-enriched KZFPs. The numbers in parentheses represent the numbers of detected peaks. (*B–D*) Binding preference of KZFPs at LINEs (*B*), LTRs (*C*), and SINEs (*D*). The bar plot represents the ratio value of the observed numbers of peaks versus the theoretically expected numbers of peaks. Binomial test. (***) $P < 0.001$. (*E–G*) Violin plots showing the data distribution of normalized H3K9me3 signals at each KZFP-bound LINEs (*E*), LTRs (*F*), and SINEs (*G*) compared with enrichment at all interspersed copies in KIT$^+$ spermatogonia. The width of the plot illustrates the data density. Mann–Whitney *U* test. (***) $P < 0.001$.

ZFP987, which showed significant enrichment at ERV loci (Figs. 2C, 3A). Because these KIT$^+$-enriched KZFPs bind to specific ERVs that are enriched with H3K9me3 in KIT$^+$ SGs, we suspected that the target ERVs would be activated as gene regulatory elements when expression of these KZFPs is repressed in the meiotic stage. To test this possibility, we first compared H3K9me3 levels at the target ERVs during the transition from mitotic SGs to meiotic spermatocytes. At the KIT$^+$-enriched KZFPs (ZFP989, GM14401, and ZFP987) target 1170 ERV loci (Fig. 3A), H3K9me3 decreased from KIT$^+$ SGs and PSs (Fig. 3B). The reduction of H3K9me3 was confirmed in track views of representative ERV loci that are targets of KIT$^+$-enriched KZFPs (Supplemental Fig. S3A). Notably, H3K27 acetylation (H3K27ac), a marker of active enhancers, was increased at these ERV loci from KIT$^+$ SGs to PSs (Fig. 3C). Further examination of substages during meiotic prophase I confirmed the progressive loss of H3K9me3 and gain of H3K27ac at these ERV loci (Supplemental Fig. S3B,C). We examined the overlap of target ERVs at which H3K9me3 levels decreased and H3K27ac levels increased (Fig. 3D). The data indicate that target ERVs acquiring H3K27ac in meiotic cells were predominantly marked with the repressive H3K9me3 in mitotic SGs with a reduction in these repressive marks upon entry into meiosis. These results suggest that H3K9me3 KIT$^+$-enriched KZFP-target ERVs are released from KZFP-mediated silencing and acquire the hallmark of active enhancers during the transition from mitotic SGs to meiotic spermatocytes.

Consistent with these epigenomic features, the KZFPs ZFP989, GM14401, and ZFP987 target 1170 ERVs that were predominantly located in intergenic regions (Fig. 3E). Given our previous finding that a subset of ERVs serve as active enhancers during meiosis to drive spermatogenic gene expression (Sakashita et al. 2020), we suspected that these ERVs also function as enhancers. To test this proposal, from the 1170 ERV loci potentially regulated by KZFPs, we first selected 286 ERV loci based on their transcripts and the adjacent gene transcripts being detected by RNA-seq either in KIT$^+$ SGs or in PSs. Whereas expression of intergenic ERVs was readily detected (Supplemental Fig. S4A–C), expression of ERVs located within introns was difficult to distinguish from the background level within introns (Supplemental Fig. S4D). To eliminate the possible confounding effect of intron-derived signals, we excluded seven ERV loci located in intronic regions, resulting in selecting 279 nonintronic ERVs as targets for KZFPs (Fig. 3A).

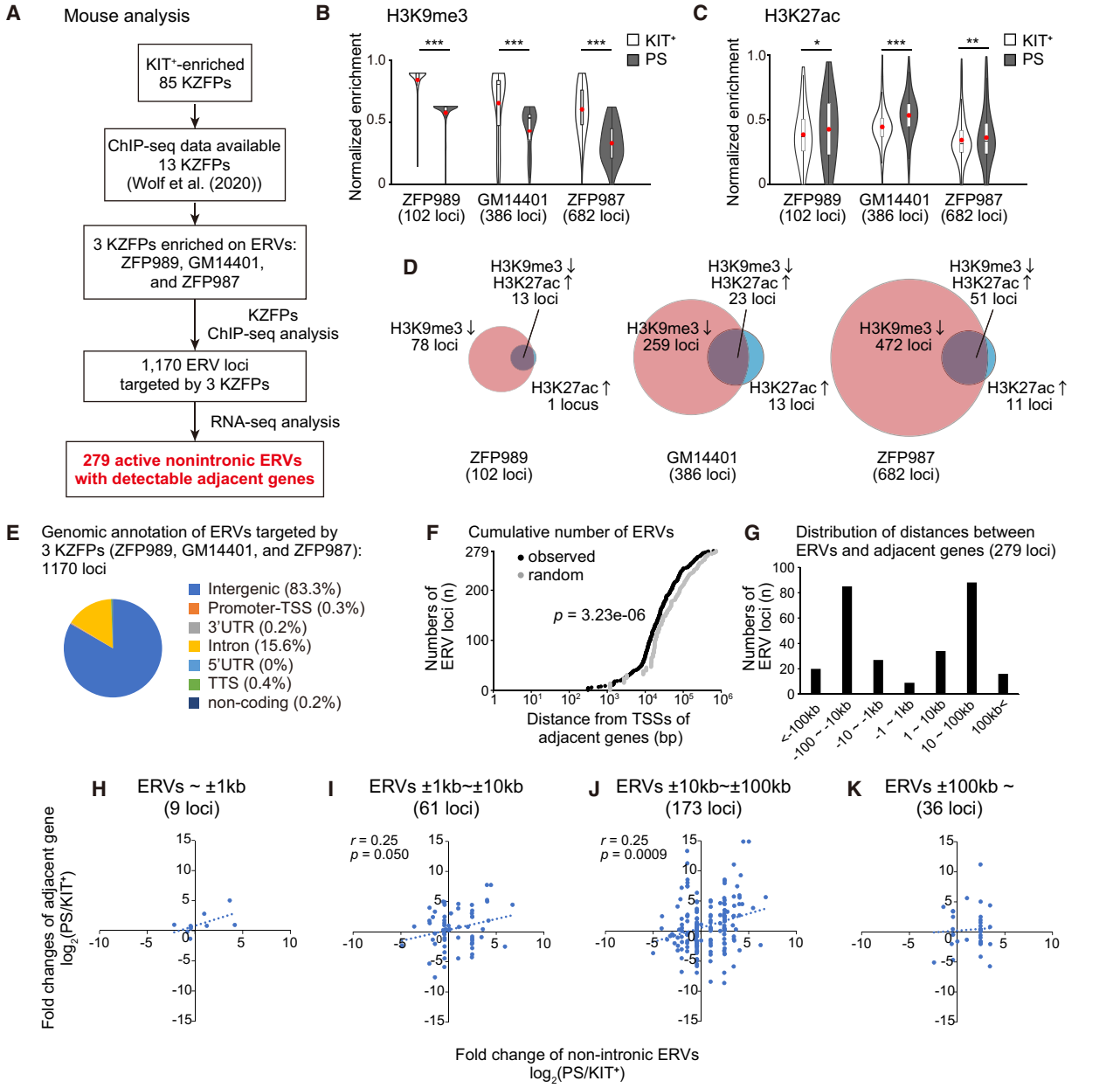

Figure 3. KZFPs target and suppress ERVs until the mitosis-to-meiosis transition. (*A*) Flowchart of analyses to identify target ERVs. (*B*,*C*) Normalized enrichment of H3K9me3 (*B*) and H3K27ac (*C*) at the target ERV loci in KIT$^+$ and PS. Violin plot with a box plot overlay of the normalized enrichment of each histone modification. 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 normalized enrichment. Paired *t*-test. (***) $P<0.001$, (**) $P<0.01$, (*) $P<0.05$. (*D*) Overlap between the target ERV loci at which H3K9me3 decreased and H3K27ac increased. (*E*) Genome annotation of the ERV loci targeted by ZFP989, GM14401, and ZFP987. Each proportion is shown in parentheses. (*F*) Cumulative curve of 279 active nonintronic ERVs targeted by ZFP989, GM14401, and ZFP987, based on the distances to the most adjacent TSSs. The gray dots represent the distance from randomly chosen ERVs to the most adjacent TSSs. *D* and *P*-value are calculated by the Kolmogorov–Smirnov test. (*G*) Distances between 279 active nonintronic ERVs targeted by ZFP989, GM14401, and ZFP987 and their adjacent genes. (*H–K*) Correlation between expression of the adjacent gene and nonintronic ERVs at the mitosis-to-meiosis transition (PS/KIT$^+$). All nonintronic ERVs were separated into four groups according to the distance to the most adjacent genes. The correlation coefficient (*r*) and *P*-value were calculated on panels with more than 50 plots (*I*,*J*).

Among these 279 nonintronic ERV loci, most loci were located 10 kb–100 kb from the gene TSS, and the distribution, when compared with randomly selected ERVs, was skewed toward relatively proximal regions (Fig. 3F,G), suggesting a potential role as distal enhancers. We further classified 279 loci into four categories according to their genomic distance to the nearest TSSs; nine loci were located within 1 kb, 61 loci were within 1–10 kb, 173 loci were within 10–100 kb, and 36 loci were located >100 kb from adjacent genes (Fig. 3H–K). Although we could not calculate the *P*-value owing to the small sample size in the groups of within 1

kb and of >100 kb, we observed a positive correlation between the increased expression of ERVs and increased expression of their adjacent genes from KIT$^+$ SGs to PSs in all groups. Taken together, we conclude that KIT$^+$-enriched KZFPs repress the activity of ERVs, and downregulation of KZFPs would subsequently activate the enhancer activity of ERVs and their adjacent genes in meiotic cells in mice.

Knockdown of KIT$^+$-enriched KZFPs results in derepression of target ERVs and adjacent genes

To ascertain whether KIT$^+$-enriched KZFPs suppress meiotic ERVs in nonmeiotic cells, we performed siRNA knockdown experiments for KIT$^+$-enriched KZFPs. We used ESCs as an example of nonmeiotic cells that undergo mitotic divisions to test the function of KZFPs in nonmeiotic cells, a strategy we previously used to activate meiotic enhancers in nonmeiotic cells (Maezawa et al. 2020; Sakashita et al. 2020). Because meiotic ERV enhancers and germline genes are not active in nonmeiotic cells, we examined whether, after KZPF knockdown, meiotic ERVs and their adjacent genes were derepressed in ESCs. ZFP987 binds and suppresses one type of meiotic ERV, *ERVB4_1C-LTR*, and its adjacent germline genes, *Or1f19* and *Vmn2r100* (Fig. 4A). Targeting *Zfp987* reduced expression of *Zfp987* and derepressed *ERVB4_1C-LTR*, *Or1f19*, and *Vmn2r100* (Fig. 4B). In line with this observation, an overall decrease in H3K9me3 signals and an increase in H3K27ac signals were observed in adjacent regions of ZFP987 binding sites near the *Or1f19* and *Vmn2r100* gene loci at the KIT$^+$-to-PS transition (Fig. 4C,D), although the increase in RNA-seq signals was only observed at the *Or1f19* locus. These results suggest that ZFP987 represses *ERVB4_1C-LTR* via H3K9me3 in KIT$^+$ and that *ERVB4_1C-LTR* and its target gene *Or1f19* are activated in PSs after the loss of ZFP987. We also performed siRNA knockdown experiments for another KIT$^+$-enriched KZFP, *Zfp989*, which targets a type of ERV, *IAPLTR1a*, adjacent to the *Ube3a* gene (Fig. 4E). siRNA knockdown of *Zfp989* resulted in derepression of *IAPLTR1a* and the *Ube3a* gene (Fig. 4F). During the KIT-to-PS transition, a decrease in H3K9me3 signals and an increase in RNA-seq and H3K27ac signals were observed in adjacent regions of ZFP989 binding sites near the *Ube3a* gene locus (Fig. 4G), leading to the upregulation of the *Ube3a* gene in PSs. Thus, the results observed in siRNA knockdown experiments likely reflect in vivo events in the transition from mitotic SGs to meiotic spermatocytes. Taken together, these results suggest that KIT$^+$-enriched KZFPs repress target ERVs and that the release from KZFP-mediated repression allows activation of ERV enhancers and their target genes, which occurs during the transition from SGs to spermatocytes during spermatogenesis (Fig. 4H).

Expression of human KZFPs at the mitotic stage is associated with ERV suppression in male germ cells

We next examined whether human KZFPs, which have undergone distinct evolution from mouse KZFPs, also regulate expression of meiotic ERVs. Like mouse ERVs, human ERVs have undergone rapid evolution. Some human ERVs act as enhancers to drive species-specific genes in meiosis (Sakashita et al. 2020), raising the possibility that KZFPs repress meiotic ERVs in premeiotic cells in humans. To test this possibility, we reanalyzed previous RNA-seq data of human spermatogenic cells (Zhu et al. 2016). Human KZFPs showed a drastic change in expression during the transition from SGs to primary spermatocytes (PriSCs), which include PSs (Fig. 5A). Human KZFPs undergo a further change from PriSCs to RSs (Fig. 5A). Four hundred sixty-nine human KZFPs were categorized into three classes, including 118 SG-enriched KZFPs (Class I), 211 PriSC-enriched KZFPs, and 141 RS-enriched KZFPs (Fig. 5A; Supplemental Table 2). Further, the stage-specific expression of each Class was confirmed with a TPM-based expression analysis (Supplemental Fig. S5A). Likewise, human TEs, including LINEs and ERVs, undergo dynamic expression changes during spermatogenesis (Fig. 5B), consistent with our previous analysis (Sakashita et al. 2020). LINEs and ERVs were significantly upregulated from SGs to PriSCs (Supplemental Fig. S5B). Of note, like mouse, human SINEs were predominantly expressed in SGs but were suppressed in PriSCs and RSs (Fig. 5B; Supplemental Fig. S5B).

To examine the interactions between human KZFPs and TEs, we reanalyzed a previous data set of ChIP-exo (ChIP with lambda exonuclease and sequencing) experiments of 222 human KZFPs in HEK293T cells (Imbeault et al. 2017). Among 118 SG-enriched KZFPs, 65 KZFPs have ChIP-exo data, and 62 KZFPs were highly enriched on ERVs based on the previous analysis (Fig. 5C; Imbeault et al. 2017). Using these data, we detected binding peaks that have varied frequencies of unique peaks that do not overlap with peaks of other KZFPs (Supplemental Fig. S5C,D). These 62 KZFPs bind 16,998 ERV loci, of which 70.1% were located in intergenic regions and 28.6% were located in introns (Fig. 5D).

Among 16,998 ERV loci targeted by SG-enriched KZFPs, we first selected 9812 ERV loci, whose transcripts and the adjacent gene transcripts were detected by RNA-seq in either SGs or PriSCs. After excluding 2233 ERVs located in intronic regions, 7579 nonintronic ERVs remained as targets for KZFPs (Fig. 5C). The majority of the nonintronic ERV loci were located within 10 kb–100 kb from the gene TSSs, and the distribution, when compared to randomly selected ERVs, was skewed toward relatively proximal regions (Fig. 5E,F), again suggesting a potential role as distal enhancers. We further classified 7579 loci into four categories according to their genomic distance to the nearest TSSs; 300 loci were located within 1 kb, 2283 loci were within 1–10 kb, 4412 loci were within 10–100 kb, and 584 loci were located >100 kb from adjacent genes (Fig. 5G–J). In all categories, we observed a positive correlation between the increased expression of human ERVs and increased expression of their adjacent genes from SGs to PriSCs. These results suggest that in human SGs, a subset of SG-enriched KZFPs bind and repress the expression of ERVs that later function as enhancers. In addition, a subset of proximal ERVs (<1 kb from TSSs) were activated from SGs to PriSCs and correlated with adjacent gene expression (Supplemental Fig. S5E,F), suggesting that a subset of proximal ERVs act as promoters to drive meiotic gene expression.

The transition from mitosis to meiosis underlies coevolution of KZFPs and ERVs in mammals

Our analysis suggests a molecular mechanism by which a group of KZFPs controls ERVs to regulate meiotic gene expression in mammals. Because KZFPs and ERVs have evolved rapidly through an arms race (Imbeault et al. 2017; Senft and Macfarlan 2021), we determined the evolutionary history of SG-enriched KZFPs and their target ERVs that are activated in meiosis. In mice, 66 KIT$^+$-enriched KZFPs were relatively conserved KZFPs and also formed a rodent-specific clade (Fig. 6A,B). This finding suggests that the KIT$^+$-enriched KZFPs also evolved rapidly and function in closely related species such as mice and rats.

We also evaluated the evolutionary distance of murine KZFPs to other mammalian species based on previous data sets (Imbeault et al. 2017). We found that KIT$^+$-enriched KZFPs were enriched

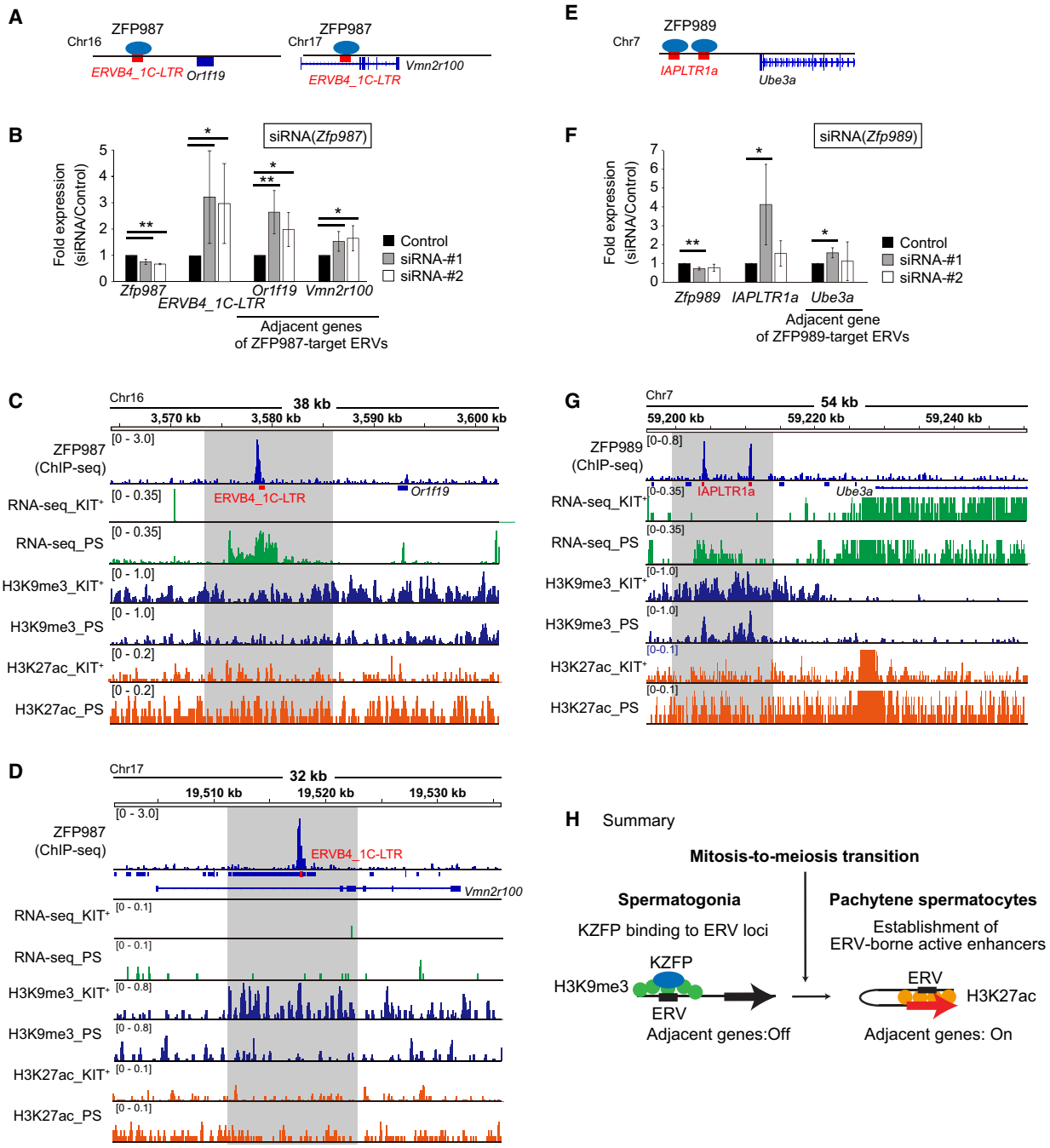

Figure 4. Knockdown of KIT$^+$-enriched KZFPs results in derepression of ERVs and their adjacent genes. (*A*,*E*) Schematic of ERV enhancers and their target loci. (*B*,*F*) Results of qRT-PCR for KZFP gene-targeted ERV and ERV-adjacent genes under *Zfp987* knockdown (*B*) and *Zfp989* knockdown (*F*). Data values for controls were normalized to 1.0. Student's *t*-test. (**) $P<0.01$, (*) $P<0.05$. n = 4. (*C*,*D*,*G*) Track views of KZFP, RNA-seq, H3K9me3, and H3K27ac enrichment in KIT$^+$ and PSs at the target ERV loci. Genomic regions around target genes (*Or1f19* [*C*], *Vmn2r100* [*D*], and *Ube3a* [*G*]) are shown. Target ERVs are represented as red-colored letters and quadrangles. Adjacent regions of ZFP987 or ZFP989 binding sites are highlighted with gray-colored areas. (*H*) Summary: In spermatogonia, spermatogonia-enriched KZFPs bind and suppress ERVs that later function as enhancers. At the ERV loci targeted by spermatogonia-enriched KZFPs, after entering meiosis, KZFPs were displaced and H3K9me3 decreased, whereas H3K27ac increased to become active enhancers.

with relatively conserved KZFPs (Supplemental Fig. S6A). Among mouse KIT$^+$-enriched KZFPs, we specifically focused on three ERV-targeting KZFPs, ZFP989, GM14401, and ZFP987, because they preferentially bound ERVs (Fig. 2C). These three KZFPs were identified as *Mus* genus–specific (Imbeault et al. 2017), and we found that these three KZFPs consistently recognized a specific subset of *Mus* genus–specific ERVKs; such pairs include ZFP989-IAPLTR1a_Mm, GM14401-IAPEz-int, and ZFP987-ERVB4_1C-LTR (Fig. 6C,D). The ERVK family showed the most significant upregulation after the transition from mitotic SGs to meiotic spermatocytes in mice (Fig. 1F), and indeed, the target ERVs of these three KZFPs were upregulated in PS (Supplemental Fig. S7A). Thus, these

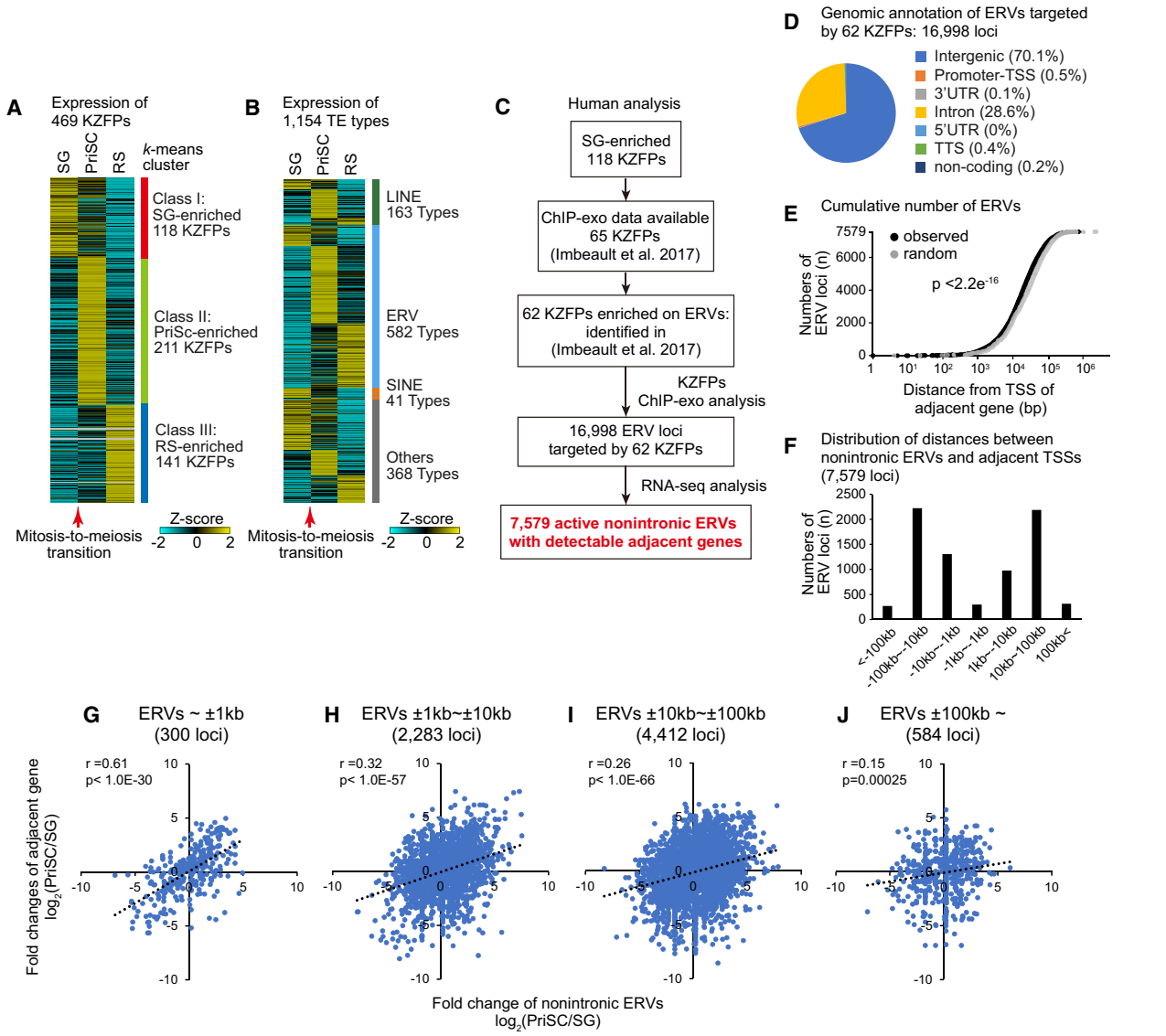

Figure 5. Human premeiotic KZFPs can act as suppressors of meiotic enhancer ERVs. (*A*) Heatmaps showing a *k*-means clustering analysis of expression (RNA-seq data) for all protein-coding KZFPs in human testicular germ cells. (*B*) Heatmaps showing the expression (RNA-seq data) of all detected TE types in human testicular germ cells. (*C*) Flowchart of analyses to identify target ERVs. (*D*) Genome annotation of the human ERV loci targeted by 62 human SG-enriched KZFPs. (*E*) Cumulative curve of active nonintronic ERVs targeted by 62 human SG-enriched KZFPs, based on the distance to the most adjacent TSSs. The gray dots represent the distance from randomly chosen ERVs to the most adjacent TSSs. D and *P*-value are calculated by the Kolmogorov–Smirnov test. (*F*) Distances between ERVs targeted by 62 human SG-enriched KZFPs and their adjacent genes. (*G–J*) Correlation between expression of the adjacent gene and nonintronic ERVs at the mitosis-to-meiosis transition (PriSC/SG). All nonintronic ERVs were separated into four groups according to the distance to the most adjacent genes. The correlation coefficient (*r*) and *P*-value are shown.

murine-specific KZFPs (ZFP989, GM14401, and ZFP987) target newly evolved ERVKs in SGs, and these targeted ERVKs become active in meiosis, showing an evolutionary relationship with the target ERVs.

Next, we examined the evolutionary aspects of human KZFPs. Overall, 113 SG-enriched KZFPs had relatively high sequence homologies among primates and were less enriched with human-specific KZFPs (Fig. 6E). Consistent with this feature, SG-enriched KZFPs were relatively well conserved between humans and chimpanzees (Fig. 6F), a feature consistent with murine KZFPs. We also evaluated the evolutionary distance of human KZFPs and found that SG-enriched KZFPs were enriched with relatively conserved KZFP genes (Supplemental Fig. S6B). We then evaluated the evolutionary relationship between human SG-enriched KZFPs and their target ERVs. We focused on four representative SG-enriched KZFPs (ZNF468, ZNF792, ZNF442, and ZNF317) as models because these KZFPs showed a decrease in expression during the transition from mitotic SGs to meiotic spermatocytes (Supplemental Fig. S5G). These four human SG-enriched KZFPs targeted various ERV families with a bias toward the ERV1 and ERVL-MaLR families (Fig. 6G). The evolution of these four KZFPs corresponds to the evolution of target ERVs; Catarrhini-specific ZNF468-targeted Catarrhini-specific MER11A and LTR17, and Eutheria-specific ZNF792 and ZNF317 primarily

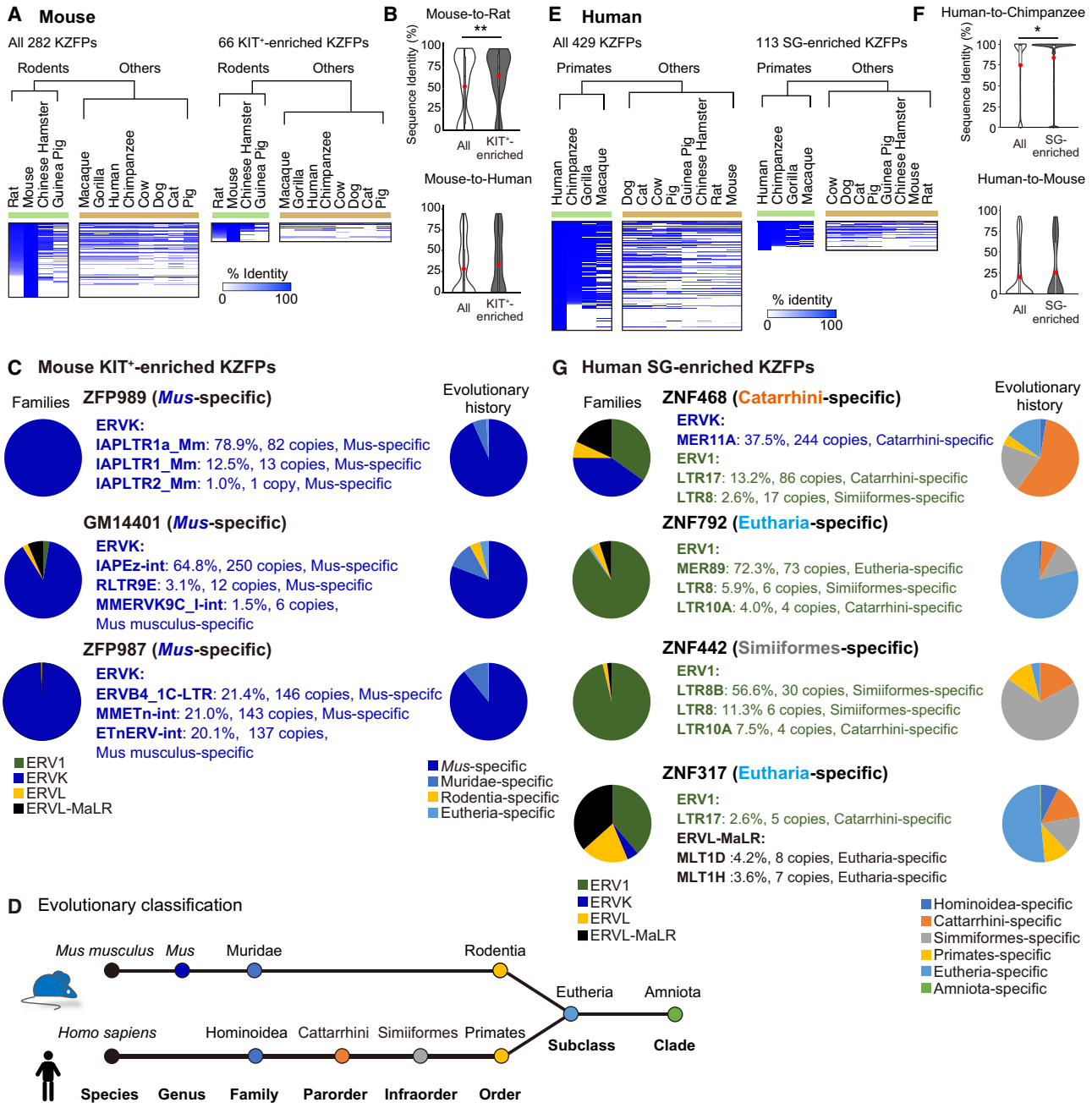

Figure 6. The evolutional aspects of KZFPs enriched in the mitotic spermatogonia. (*A*) The heatmap represents the identity of all 282 murine KZFPs (*left*) and 66 KIT$^+$-enriched KZFPs (*right*) compared among mammalian species. The dendrogram represents the result of the hierarchical analysis, which categorizes all species into two groups. (*B*) The violin plot with a box plot overlay of the sequence identity of ALL KZFPs (white) and KIT$^+$-enriched KZFPs (gray) between mice and rats (*top*) and mice and humans (*bottom*). 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 1.5X inter-quartile range (IQR) from the hinge. The red dots in the box plot are the average of the normalized enrichment. Mann–Whitney *U* test. (**) $P < 0.01$. (*C*) The *left* pie chart represents the proportion of subfamilies of target ERVs. The names of the ERV types shown in the *middle* represent the top three ERV types targeted by each KZFP. The *right* pie chart represents the proportion of evolutional ages of target ERVs. The evolutional history of each KZFP is shown in parentheses. (*D*) Schematics of evolutional clades for mice (*top*) and humans (*bottom*). (*E*) The heatmap represents the identity of all 429 murine KZFPs (*left*) and 113 KIT$^+$-enriched KZFPs (*right*) compared among mammalian species. The dendrogram represents the result of the hierarchical analysis, which categorizes all species into two groups. (*F*) The violin plot with a box plot overlay of the sequence identity of All KZFPs (white) and KIT$^+$-enriched KZFPs (gray) between humans and chimpanzees (*top*) and humans and mice (*bottom*). 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 1.5X IQR from the hinge. The red dots in the box plot are the average of the normalized enrichment. Mann–Whitney *U* test. (*) $P < 0.05$. (*G*) The evolutional age of four representative human SG-enriched KZFPs and their target ERVs. The *left* pie chart represents the proportion of subfamilies of target ERVs. The names of the ERV types shown in the *middle* represent the top three ERV types targeted by each KZFP. The *right* pie chart represents the proportion of evolutional ages of target ERVs.

targeted Eutheria-specific ERVs. Furthermore, Simiiformes-specific ZNF442-targeted Simiiformes-specific ERVs. These target ERVs were upregulated in PriSCs (Supplemental Fig. S7B). These results suggest concomitant evolution of SG-enriched KZFPs and their target meiotic ERVs in mice and humans. Such coevolution is consistent with an arms race between these subsets of KZFPs and ERV pairs that are associated with the transition from mitotic SGs to meiotic spermatocytes in mammals.

Meiotic KZFPs are unlikely to be associated with ERV repression in meiosis

In contrast to SG-enriched KZFPs, groups of KZFPs were highly upregulated after the transition from mitotic SGs to meiotic spermatocytes in both mice and humans. In mice, 35 Class VI KZFPs were highly expressed in PSs (Fig. 1B), raising the possibility that they function in meiotic cells. Of note, these 35 KZFPs are relatively conserved compared with KIT$^+$-enriched KZFPs among mammals (Supplemental Fig. S8A), and their human orthologs tend to be expressed in meiosis or in postmeiotic RSs (Supplemental Fig. S8B). To characterize these PS-enriched KZFPs, we first analyzed the binding preference of five PS-enriched KZFPs for which ChIP-seq data were available (Supplemental Fig. S8C,D). Among these five KZFPs, ZFP992, ZKSCAN17, and ZFP94 showed a significant enrichment at TE loci (Supplemental Fig. S8E). ZFP992 and ZFP94 were significantly enriched at ERVs, and ZFP992, ZFP94, and ZFP457 were significantly enriched at SINEs, whereas none of the five KZFPs were preferentially bound to LINEs (Supplemental Fig. S8F). However, these target ERVs and SINEs did not show expression changes during the transition from mitotic SGs to meiotic spermatocytes (KIT$^+$ to PS) (Supplemental Fig. S8G,H), suggesting that PS-enriched KZFPs are not associated with repression of target ERVs in meiosis in mice.

We next examined whether human KZFPs that are highly expressed in meiosis are associated with repression of target ERVs in meiosis. In contrast to mouse meiotic KZFPs, which were well conserved among mammals (Supplemental Fig. S8A), human PriSC-enriched KZFPs were not well conserved among mammals (Supplemental Fig. S9A). Of the 211 PriSC-enriched KZFPs, 89 of the 99 KZFPs with ChIP-exo data preferentially bind ERVs (Supplemental Fig. S9B). We further selected 34 KZFPs that were significantly expressed in PriSCs (Supplemental Fig. S9C). However, expression of the target ERV loci of these 34 KZFPs was not significantly changed during the transition from mitotic SGs to meiotic spermatocytes (SGs to PriSC) (Supplemental Fig. S9D). Together with the mouse analysis, our results suggest that meiotic KZFPs are unlikely to be associated with ERV repression in meiosis in mammals.

MSCI may antagonize the coevolution of KZFPs and ERVs in mammals

During meiosis, sex chromosomes undergo MSCI, which is an essential process in the male germline (Turner 2015; Alavattam et al. 2022). MSCI silences expression of most sex chromosome–linked genes during the pachytene and diplotene stages and largely continues to postmeiotic silencing, with a subset of genes associated with spermiogenesis escaping postmeiotic silencing (Namekawa et al. 2006; Mueller et al. 2008). Of note is that MSCI impacts sex chromosome evolution (Potrzebowski et al. 2008). For example, the X Chromosome is enriched with genes expressed before meiosis (Khil et al. 2004) and with spermatid-specific genes that escape postmeiotic silencing of the sex chromosomes (Sin et al. 2012; Mueller et al. 2013). Therefore, we sought to determine whether MSCI impacted the coevolution of KZFPs and ERVs in mammals. Because we found that SG-enriched KZFPs regulate meiotic ERVs, we examined the chromosome distribution of these ERVs in mice and humans. Notably, meiotic ERVs were significantly enriched on autosomes and underrepresented on the sex chromosomes in mice (Fig. 7A,B). Nevertheless, when focusing solely on the X Chromosome, we found a similar trend in humans (Fig. 7C,D). The previous human KZFP ChIP-exo experiments were performed using HEK293T cells from a female donor, which precludes evaluating KZFP-binding sites on the Y Chromosome. We next examined the chromosomal distribution of ERVs targeted by meiotic KZFPs and found that these ERVs were predominantly located on autosomes and underrepresented on the sex chromosomes (Supplemental Fig. S10A–D). In humans, although PriSC-enriched KZFPs were not located on the X Chromosome, the presence of SG-enriched KZFPs was observed on the X Chromosome (Supplemental Fig. S10E,F). This distribution pattern is consistent with the enrichment of genes expressed before meiosis on the X Chromosome (Khil et al. 2004). These analyses suggest that MSCI may

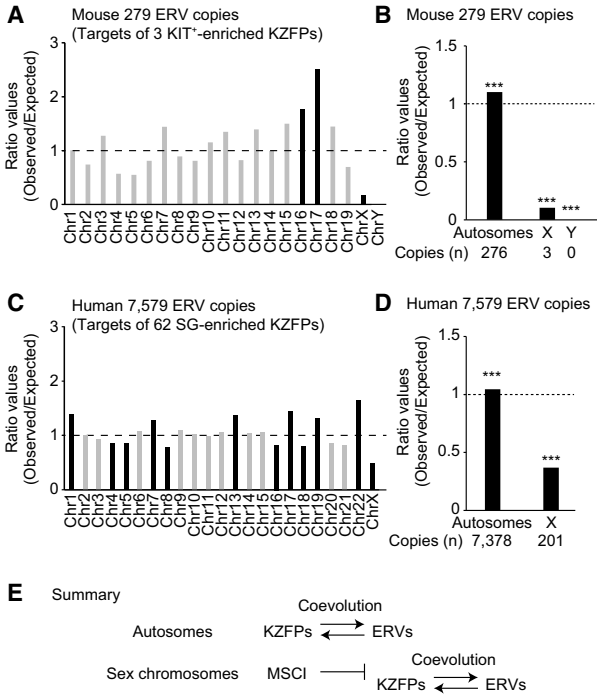

Figure 7. Chromosomal distribution of ERVs targeted by mouse and human spermatogonia-enriched KZFPs. (*A*) Chromosomal distribution of mouse ERV copies targeted by KIT$^+$-enriched KZFPs. The bar plot represents the ratio values of the observed copy number of ERVs versus the theoretically expected number. The color of each bar represents the statistical significance: black indicates $P<0.05$; gray, not significant; binomial test. (*B*) Distribution of mouse ERV copies targeted by KIT$^+$-enriched KZFPs on autosomes or sex chromosomes. Binomial test. (***) $P<0.001$. (*C*) Chromosomal distribution of human ERV copies targeted by SG-enriched KZFPs. The bar plot is represented the same way as panel *A*. The Y Chromosome data are absent because the HEK293T cells used in the ChIP-exo experiments to determine human KZFP binding were derived from a female donor and, therefore, lacked Y Chromosomes. (*D*) Distribution of human ERV copies that are targeted by PriSC-enriched KZFPs on autosomes or sex chromosomes. The bar plot is represented the same as panel *B*. (*E*) Summary. MSCI antagonizes the coevolution of KZFPs and ERVs on the sex chromosomes in mammals.

antagonize the coevolution of KZFPs and ERVs on the sex chromosomes in mammals (Fig. 7E).

Discussion

In this study, by focusing on KZFPs, one of the most rapidly evolving families of TFs during mammalian radiation (Tadepally et al. 2008; Yang et al. 2017; Bruno et al. 2019; Wolf et al. 2020; Senft and Macfarlan 2021), we describe a molecular logic by which a large number of ERVs are controlled and specifically activated after the transition from mitotic SGs to meiotic spermatocytes. The expression of KZFPs changes drastically during the same transition during spermatogenesis in both mice and humans, and these changes accompany expression changes of a variety of TEs. At the ERV loci targeted by a subset of SG-enriched KZFPs, active enhancers likely form after entering meiosis, as KZFPs are lost, H3K9me3 decreases, and H3K27ac increases (Fig. 4H).

Because retrotransposons require transcription to integrate into the germline, retrotransposons are activated in the germline, where their expression is tightly controlled to minimize genome instability. What remains largely unknown is when retrotransposons are activated in the germline. Our study suggests that the transition from mitosis to meiosis is a critical time window for the coevolution of rapidly expanding TEs and rapidly evolving suppressive KZFPs as reflected in the reciprocal pattern of expression of KZFPs and their target ERVs. One consequence of such coevolution would be lineage-specific germline transcriptomes. Notably, this window of coevolution coincides with genome-wide changes in chromatin state and epigenomic modifications from mitosis to meiosis (Sin et al. 2015; Maezawa et al. 2018b; Alavattam et al. 2019; Patel et al. 2019; Wang et al. 2019). Results reported here provide further support for such coevolution. Coevolution of KZFPs and ERVs has also been suggested in the context of mouse and human embryonic development (Pontis et al. 2019, 2022; Seah et al. 2019; Iouranova et al. 2022). Thus, there may be multiple windows in the germline cycle to drive genomic evolution.

We also sought to understand the functions of KZFPs that are enriched at the meiotic stage. Although some PS-enriched KZFPs show a binding preference for ERVs or SINEs, we did not find any association with gene expression changes. Therefore, meiotic KZFPs may have functions other than transcriptional repression, may still be evolving toward TE regulation, or may have lost some of their targets owing to the arms race. For example, PRDM9, a well-conserved, and the most ancient, type of KZFP, functions as the determinant of meiotic DNA double-strand break sites during meiotic recombination but is not involved in gene regulation. Although human and mouse PRDM9 protein has a KRAB-A domain, it does not interact with an essential regulator of heterochromatin, TRIM28, and thus is not involved in transcriptional repression (Baudat et al. 2010; Myers et al. 2010; Patel et al. 2016; Imai et al. 2017). In addition, a previous study revealed that an evolutionarily ancient group of KZFPs, termed variant KZFPs, do not interact with TRIM28 and may have different roles other than TRIM28-dependent TE suppression (Helleboid et al. 2019); the study identified 35 human KZFPs as variant KZFPs, and there are 26 orthologs in the mouse genome. Notably, among 26 orthologous genes, nine KZFP genes (*Zfp202*, *Zkscan5*, *Zkscan6*, *Zkscan17*, *Zfp212*, *Zfp786*, *Zfp473*, *Zfp446*, *Zfp78*) are classified as PS-enriched KZFPs in our current study and therefore may not act as transcriptional repressors.

We find that KZFPs-targeted ERVs are underrepresented on the X Chromosome, presumably because of MSCI. A previous study reported that autosomes carry many retrotransposed genes that originated from the X Chromosome to avoid MSCI; these genes are expressed later during spermatogenesis (Emerson et al. 2004). These cases of retrotransposition coincide with the timing of the dynamic changes in expression of KZFP and retrotransposons reported here. Therefore, extensive genomic evolution may be driven in part by MSCI during the transition from mitotic SGs to meiotic spermatocytes when extensive epigenetic programming takes place. During meiosis, H3K9me3 is established on the X Chromosome by SETDB1 and its germline-specific partner ATF7IP2, but this action of SETDB1–ATF7IP2 is regulated downstream from the DNA damage response (DDR) pathway (Hirota et al. 2018; Alavattam et al. 2024), which is the master regulator of MSCI (Alavattam et al. 2022). Both X-linked protein-coding genes and TEs are also highly enriched with H3K9me3 in PSs (Liu et al. 2019; Alavattam et al. 2024), likely by the action of DDR-directed H3K9me3, which can be KZFP independent. Of note, we previously showed enrichment of ERV enhancers marked with H3K27ac on the X Chromosome in PSs (Sakashita et al. 2020), which may reflect that these X-linked ERVs do not require KZFPs as a regulatory mechanism. In addition to H3K9me3, H3K27ac is also established by the DDR pathway on the sex chromosomes in meiosis (Adams et al. 2018). Our observation suggests that distinct mechanisms regulate the activities of ERVs on autosomes and sex chromosomes during spermatogenesis.

The binding-site information generated in previous studies (Imbeault et al. 2017; Wolf et al. 2020) restricted our analysis to a limited number of KZFPs and, therefore, limited our ability to make broad generalizations of the KZFP function. Nevertheless, we predict not only that there is a much broader network of interactions taking place between KZFPs and ERVs in the mammalian male germline but also that the germline functions of KZFPs are much broader than previously thought, such as the cases for PRDM9 and other KZFPs required for imprinting control (Li et al. 2008; Takahashi et al. 2019).

Our results provide the first evidence that an evolutionary arms race between virus-derived sequences, ERVs, and endogenous KZFP defense machinery takes place during the transition from mitotic SGs to meiotic spermatocytes in germ cells. Expression of KZFPs mirrors expression of ERVs, and the target ERVs bound by a subset of SG-enriched KZFPs are highly covered with H3K9me3 in SGs. After germ cells enter meiosis, H3K9me3 disappears, and the ERV loci acquire an active enhancer mark H3K27ac that is accompanied by the loss of KZFPs we assessed, leading to activation of adjacent genes in meiotic PSs. This trend is found in mice and humans, suggesting that mammalian species have evolved regulatory machinery consisting of rapidly evolving KZFPs and ERV activity in the male germline. Furthermore, KZFP-dependent ERV suppression mainly operates on autosomes and is underrepresented on the sex chromosomes. We propose that epigenetic programming in the mammalian germline during the transition from mitotic SGs to meiotic spermatocytes facilitates the coevolution of KZFPs and TEs on autosomes and is antagonized by MSCI.

Methods

Sequence comparison of KZFPs among mammalian species

The identity scores of KZFPs were calculated using the BioMart web tool (https://www.ensembl.org/biomart/martview). We first organized the gene names of KZFPs in mice or humans as a gene list and

then chose the data set "GRCm39" or "GRCh38.p14" for mice and humans, respectively. The gene list was then entered as an input. As attributes, we chose the percentage of identities of target species (including 11 mammalian species) gene identical to query gene (mice or human) to obtain the sequence identity of the KZFP homolog.

RNA-seq data analysis

Raw reads of previous RNA-seq reads were first trimmed using Trim Galore! (version 3.3; https://github.com/FelixKrueger/TrimGal ore) and then aligned to either the mouse (GRCm38/mm10) or human (GRCh38/hg38) genomes using STAR alignment software (version 2.5.3a) (Dobin et al. 2013). Multiple mapped reads were used to quantify gene expression. All aligned reads were then sorted and indexed by SAMtools (version 1.11) (Li et al. 2009) function sort and index, respectively. The gene annotation file was obtained from the GENCODE database (v41 for humans and vM25 for mice; https ://www.gencodegenes.org). The list of KZFPs was obtained from the previous study (Imbeault et al. 2017).

To quantify TE expression, only uniquely aligned reads were used. All aligned reads were sorted and indexed by SAMtools function sort and index, respectively. TEs and repeat sequence annotation files were downloaded from the UCSC Genome Browser RepeatMasker Track (https://genome.ucsc.edu/cgi-bin/hgTrackUi ?g=rmsk), which is based on the Repbase (Bao et al. 2015) library.

After alignment and sorting, the reads mapped on genes were counted by the function featureCounts using the Subreads software with the -M option(multiple-mapped reads count mode) (Liao et al. 2014). To quantify RNA-seq reads per TE copy, we performed the function featureCounts without -M option (uniquely mapped reads count mode) using RepeatMasker track. After the read count, unexpressed TE copies (fewer than two raw read counts) in the data sets were removed and served as a cutoff. To obtain an overview of TE expression levels, we recalculated the read counts by aggregating the reads of each copy according to the type it belongs to, thereby obtaining the read counts at the type level.

To determine differentially expressed genes and TE copies, read counts were used as input data for the DESeq2, the R package (version 1.34.0) (Love et al. 2014). To perform *k*-means clustering analysis for the subset of genes and TEs and draw heatmaps, we used the web tool Morpheus distributed by the Broad Institute (https://software.broadinstitute.org/morpheus). Each TE copy was classified as we did in our previous report, briefly, starting with the largest categories, class, and then in the order of subclass > order > superfamily > family > type > copy (Sakashita et al. 2020). To perform GO enrichment analysis for the subset of KZFPs, we used the web tool DAVID (Huang et al. 2009; Sherman et al. 2022).

ChIP-seq and ChIP-exo data analysis

Raw reads of previous ChIP-seq and ChIP-exo reads were first trimmed using Trim Galore! and then aligned to either the mouse (GRCm38/mm10) or human (GRCh38/hg38) genomes using Bowtie 2 software (version 2.4.4) (Langmead and Salzberg 2012). All aligned reads were then sorted and indexed using the SAMtools function sort and index, respectively. Peak calling for ChIP-seq and ChIP-exo data was performed using MACS2 (version 2.2.7.1) (Zhang et al. 2008) following the setting described in a previous study (Imbeault et al. 2017; Wolf et al. 2020). The expected background was estimated by randomly generating and calculating numbers of background genomic regions equal to the numbers of ChIP-seq peak regions determined by BEDTools (version 2.30.0) (Quinlan and Hall 2010) function random.

Target TEs of each KZFP were determined by examining the overlap between TE annotations and detected peaks by BEDTools function intersect. The normalized signal of H3K9me3 and H3K27ac was calculated using BEDTools function coverage using corresponding BAM files and genome coordinates (BED file) obtained by intersect. This coverage function enabled us to evaluate the extent to which H3K9me3 or H3K27ac occupies TE regions. The genomic annotation of target TEs was analyzed by HOMER software (version 4.11) (Heinz et al. 2010) function annotatePeaks.pl using a genome coordinate file (BED file) that contains the genomic location of target ERVs as input data. The quantitative analysis of H3K9me3 and H3K27ac levels was performed with replicates supplied by the previous studies (Adams et al. 2018; Liu et al. 2019; Chen et al. 2020; Maezawa et al. 2020). The *P*-value was calculated by the paired *t*-test. Because H3K27ac ChIP-seq data were deposited with only two biological replicates, we employed $P < 0.1$ as a threshold to evaluate the significance.

The program deepTools (version 3.5.1) (Ramirez et al. 2016) was used to draw tag density plots and heatmaps for read enrichment (H3K9me3, H3K27ac, and each KZFP ChIP-seq read). To detect genes adjacent to target ERVs, we used ChIPpeakAnno, the R package (version 3.28.1) (Zhu et al. 2010), using a genome coordinate file (BED file) that contains the genomic location of target ERVs as input data. To visualize read enrichment over representative genomic loci, bigWig files were created from sorted BAM files using deepTools function bamCoverage and visualized by using IGVTools (Robinson et al. 2011). To determine the evolutional age of target TEs, we referred to the database of the research community Dfam (Storer et al. 2021).

ESC culture and KZFP knockdown experiment

The J1 male ESCs were cultured in ESC medium (15% FBS [Hyclone], 1× GlutaMAX [Invitrogen], 1× MEM nonessential amino acids solution [Invitrogen], 1× penicillin-streptomycin [Invitrogen], and 55 μM β-mercaptoethanol in knockout DMEM [Invitrogen]) containing 2i (1 μM PD0325901, LC Laboratories; 3 μM CHIR99021, LC Laboratories) and LIF (1300 U/mL, Merck) on cell culture plates coated with 0.2% gelatin under feeder-free conditions.

One day before transfection, 1.0×10^5 cells were seeded onto a 24-well plate coated with 0.2% gelatin. For transfection, Lipofectamine RNAiMAX (Invitrogen) was used following the manufacturer's guidelines with 10 pmol siRNA. Six days after transfection, cells were lysed, and total RNA was isolated using the RNeasy plus micro kit (Qiagen). One hundred nanograms of total RNA was reverse-transcribed by using SuperScript IV (Invitrogen). cDNA was subjected to qRT-PCR using PowerUp SYBR Green master mix (Applied Biosystems) with the primers listed in Supplemental Table 3.

Evaluation of amino acid sequence similarities across mammalian species

To calculate amino acid sequence similarities and detect orthologous KZFP proteins to mouse and human KZFPs across mammalian species, we applied a list of mouse and human KZFP genes to the BioMart web tool from the Ensembl database (Kinsella et al. 2011) to compute amino acid sequence similarities, that is, percentage identities of target genes in other species in comparison to the respective query genes. Among all 363 mouse KZFPs and 469 human KZFPs registered in the GENCODE gene annotation, 282 mouse KZFPs and 429 human KZFPs are registered in the BioMart database. Thus, 282 mouse KZFPs and 429 human KZFPs were analyzed in Figure 1A. The divergence time of each

evolutionary clade was determined by referring to the web tool TimeTree (Kumar et al. 2022).

Statistical analyses

All the statistical analyses were performed using R software (version 4.1.3) (R Core Team 2018; https://www.R-project.org/).

Data sets

The raw reads of RNA-seq, ChIP-seq, and ChIP-exo from previous studies were downloaded under the accession numbers as Supplemental Table 4. Germ cell isolation for NGS data was performed in previous studies using different methods. For the RNA-seq data, THY1$^+$ and KIT$^+$ SGs were isolated by magnetic-activated cell sorting (MACS), and PSs and RSs were isolated by the BSA-gradient method using STA-PUT (Hasegawa et al. 2015; Maezawa et al. 2018a). For the ChIP-seq (GSE137744), KIT$^+$ SGs were isolated by MACS, and PSs were isolated by fluorescence-activated cell sorting (FACS) (Liu et al. 2019). For the ChIP-seq (GSE107398), KIT$^+$ SGs were isolated by MACS, and PSs were isolated by the BSA-gradient method (Adams et al. 2018; Maezawa et al. 2020). For the ChIP-seq (obtained from the Gene Expression Omnibus [GEO; https://www.ncbi.nlm.nih.gov/geo/] under accession number GSE132446), the germ cell differentiation was first synchronized by a RA-synthesis inhibitor, WIN 18446, and then each spermatogenic cell type was collected after RA induction followed by FACS (Chen et al. 2020). Single-cell RNA-seq data from the previous study (Hermann et al. 2018) were downloaded from the Mendeley Data Deposition (https://data.mendeley.com/datasets/kxd5f8vpt4/1) and were processed using 10x Genomics Loupe Browser (version 6.0.0, https://www.10xgenomics.com/products/loupe-browser).

Competing interest statement

The authors declare no competing interests.

Acknowledgments

We thank the members of the Namekawa laboratory for the discussion and helpful comments regarding this manuscript, and Neil Hunter and Todd Macfarlan for discussion. This study was funded by National Institute of General Medical Sciences R35 GM141085, UC Davis startup fund to S.H.N., and a Japan Society for the Promotion of Science postdoctoral fellowship to K.O. (24KJ01990).

Author contributions: K.O. and S.H.N. designed the study. K.O. and A.S. designed and interpreted the computational analyses. All authors interpreted the results. K.O., R.M.S., and S.H.N. wrote the manuscript with critical feedback from all other authors. S.H.N. supervised the project.

References

Adams SR, Maezawa S, Alavattam KG, Abe H, Sakashita A, Shroder M, Broering TJ, Sroga Rios J, Thomas MA, Lin X, et al. 2018. RNF8 and SCML2 cooperate to regulate ubiquitination and H3K27 acetylation for escape gene activation on the sex chromosomes. *PLoS Genet* **14:** e1007233. doi:10.1371/journal.pgen.1007233

Alavattam KG, Maezawa S, Sakashita A, Khoury H, Barski A, Kaplan N, Namekawa SH. 2019. Attenuated chromatin compartmentalization in meiosis and its maturation in sperm development. *Nat Struct Mol Biol* **26:** 175–184. doi:10.1038/s41594-019-0189-y

Alavattam KG, Maezawa S, Andreassen PR, Namekawa SH. 2022. Meiotic sex chromosome inactivation and the XY body: a phase separation hypothesis. *Cell Mol Life Sci* **79:** 18. doi:10.1007/s00018-021-04075-3

Alavattam KG, Esparza JM, Hu M, Shimada R, Kohrs AR, Abe H, Munakata Y, Otsuka K, Yoshimura S, Kitamura Y, et al. 2024. ATF7IP2/MCAF2 directs H3K9 methylation and meiotic gene regulation in the male germline. *Genes Dev* **38:** 115–130. doi:10.1101/gad.351569.124

Bao J, Yan W. 2012. Male germline control of transposable elements. *Biol Reprod* **86:** 162. doi:10.1095/biolreprod.111.095463

Bao W, Kojima KK, Kohany O. 2015. Repbase update, a database of repetitive elements in eukaryotic genomes. *Mob DNA* **6:** 11. doi:10.1186/s13100-015-0041-9

Baudat F, Buard J, Grey C, Fledel-Alon A, Ober C, Przeworski M, Coop G, de Massy B. 2010. PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice. *Science* **327:** 836–840. doi:10.1126/science.1183439

Bruno M, Mahgoub M, Macfarlan TS. 2019. The arms race between KRAB-zinc finger proteins and endogenous retroelements and its impact on mammals. *Annu Rev Genet* **53:** 393–416. doi:10.1146/annurev-genet-112618-043717

Chen Y, Lyu R, Rong B, Zheng Y, Lin Z, Dai R, Zhang X, Xie N, Wang S, Tang F, et al. 2020. Refined spatial temporal epigenomic profiling reveals intrinsic connection between PRDM9-mediated H3K4me3 and the fate of double-stranded breaks. *Cell Res* **30:** 256–268. doi:10.1038/s41422-020-0281-1

Chuong EB, Elde NC, Feschotte C. 2017. Regulatory activities of transposable elements: from conflicts to benefits. *Nat Rev Genet* **18:** 71–86. doi:10.1038/nrg.2016.139

Crichton JH, Dunican DS, Maclennan M, Meehan RR, Adams IR. 2014. Defending the genome from the enemy within: mechanisms of retrotransposon suppression in the mouse germline. *Cell Mol Life Sci* **71:** 1581–1605. doi:10.1007/s00018-013-1468-0

Davis MP, Carrieri C, Saini HK, van Dongen S, Leonardi T, Bussotti G, Monahan JM, Auchynnikava T, Bitetti A, Rappsilber J, et al. 2017. Transposon-driven transcription is a conserved feature of vertebrate spermatogenesis and transcript evolution. *EMBO Rep* **18:** 1231–1247. doi:10.15252/embr.201744059

Del Villano BC, Lerner RA. 1976. Relationship between the oncornavirus gene product gp70 and a major protein secretion of the mouse genital tract. *Nature* **259:** 497–499. doi:10.1038/259497a0

Di Giacomo M, Comazzetto S, Saini H, De Fazio S, Carrieri C, Morgan M, Vasiliauskaite L, Benes V, Enright AJ, O'Carroll D. 2013. Multiple epigenetic mechanisms and the piRNA pathway enforce LINE1 silencing during adult spermatogenesis. *Mol Cell* **50:** 601–608. doi:10.1016/j.molcel.2013.04.026

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29:** 15–21. doi:10.1093/bioinformatics/bts635

Ecco G, Imbeault M, Trono D. 2017. KRAB zinc finger proteins. *Development* **144:** 2719–2729. doi:10.1242/dev.132605

Elbarbary RA, Lucas BA, Maquat LE. 2016. Retrotransposons as regulators of gene expression. *Science* **351:** aac7247. doi:10.1126/science.aac7247

Emerson JJ, Kaessmann H, Betrán E, Long M. 2004. Extensive gene traffic on the mammalian X chromosome. *Science* **303:** 537–540. doi:10.1126/science.1090042

Erwin JA, Marchetto MC, Gage FH. 2014. Mobile DNA elements in the generation of diversity and complexity in the brain. *Nat Rev Neurosci* **15:** 497–506. doi:10.1038/nrn3730

Friedli M, Trono D. 2015. The developmental control of transposable elements and the evolution of higher species. *Annu Rev Cell Dev Biol* **31:** 429–451. doi:10.1146/annurev-cellbio-100814-125514

Fu Q, Wang PJ. 2014. Mammalian piRNAs: biogenesis, function, and mysteries. *Spermatogenesis* **4:** e27889. doi:10.4161/spmg.27889

Garcia-Perez JL, Widmann TJ, Adams IR. 2016. The impact of transposable elements on mammalian development. *Development* **143:** 4101–4114. doi:10.1242/dev.132639

Gill ME, Rohmer A, Erkek-Ozhan S, Liang CY, Chun S, Ozonov EA, Peters A. 2023. De novo transcriptome assembly of mouse male germ cells reveals novel genes, stage-specific bidirectional promoter activity, and noncoding RNA expression. *Genome Res* **33:** 2060–2078. doi:10.1101/gr.278060.123

Groner AC, Meylan S, Ciuffi A, Zangger N, Ambrosini G, Dénervaud N, Bucher P, Trono D. 2010. KRAB-zinc finger proteins and KAP1 can mediate long-range transcriptional repression through heterochromatin spreading. *PLoS Genet* **6:** e1000869. doi:10.1371/journal.pgen.1000869

Hasegawa K, Sin HS, Maezawa S, Broering TJ, Kartashov AV, Alavattam KG, Ichijima Y, Zhang F, Bacon WC, Greis KD, et al. 2015. SCML2 establishes the male germline epigenome through regulation of histone H2A ubiquitination. *Dev Cell* **32:** 574–588. doi:10.1016/j.devcel.2015.01.014

Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. 2010. Simple combinations of lineage-

determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. *Mol Cell* **38:** 576–589. doi:10.1016/j.molcel.2010.05.004
Helleboid PY, Heusel M, Duc J, Piot C, Thorball CW, Coluccio A, Pontis J, Imbeault M, Turelli P, Aebersold R, et al. 2019. The interactome of KRAB zinc finger proteins reveals the evolutionary history of their functional diversification. *EMBO J* **38:** e101220. doi:10.15252/embj.2018101220
Hermann BP, Cheng K, Singh A, Roa-De La Cruz L, Mutoji KN, Chen IC, Gildersleeve H, Lehle JD, Mayo M, Westernströer B, et al. 2018. The mammalian spermatogenesis single-cell transcriptome, from spermatogonial stem cells to spermatids. *Cell Rep* **25:** 1650–1667.e8. doi:10.1016/j.celrep.2018.10.026
Hirota T, Blakeley P, Sangrithi MN, Mahadevaiah SK, Encheva V, Snijders AP, ElInati E, Ojarikre OA, de Rooij DG, Niakan KK, et al. 2018. SETDB1 links the meiotic DNA damage response to sex chromosome silencing in mice. *Dev Cell* **47:** 645–659.e6. doi:10.1016/j.devcel.2018.10.004
Huang DW, Sherman BT, Lempicki RA. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4:** 44–57. doi:10.1038/nprot.2008.211
Huang Y, Kim JK, Do DV, Lee C, Penfold CA, Zylicz JJ, Marioni JC, Hackett JA, Surani MA. 2017. Stella modulates transcriptional and endogenous retrovirus programs during maternal-to-zygotic transition. *eLife* **6:** e22345. doi:10.7554/eLife.22345
Imai Y, Baudat F, Taillepierre M, Stanzione M, Toth A, de Massy B. 2017. The PRDM9 KRAB domain is required for meiosis and involved in protein interactions. *Chromosoma* **126:** 681–695. doi:10.1007/s00412-017-0631-z
Imbeault M, Helleboid PY, Trono D. 2017. KRAB zinc-finger proteins contribute to the evolution of gene regulatory networks. *Nature* **543:** 550–554. doi:10.1038/nature21683
Iouranova A, Grun D, Rossy T, Duc J, Coudray A, Imbeault M, de Tribolet-Hardy J, Turelli P, Persat A, Trono D. 2022. KRAB zinc finger protein ZNF676 controls the transcriptional influence of LTR12-related endogenous retrovirus sequences. *Mob DNA* **13:** 4. doi:10.1186/s13100-021-00260-0
Khil PP, Smirnova NA, Romanienko PJ, Camerini-Otero RD. 2004. The mouse X chromosome is enriched for sex-biased genes not subject to selection by meiotic sex chromosome inactivation. *Nat Genet* **36:** 642–646. doi:10.1038/ng1368
Kinsella RJ, Kahari A, Haider S, Zamora J, Proctor G, Spudich G, Almeida-King J, Staines D, Derwent P, Kerhornou A, et al. 2011. Ensembl BioMarts: a hub for data retrieval across taxonomic space. *Database (Oxford)* **2011:** bar030. doi:10.1093/database/bar030
Koito A, Ishizaka Y. 2013. Retroviruses, retroelements and their restrictions. *Front Microbiol* **4:** 197. doi:10.3389/978-2-88919-401-8
Ku HY, Lin H. 2014. PIWI proteins and their interactors in piRNA biogenesis, germline development and gene expression. *Natl Sci Rev* **1:** 205–218. doi:10.1093/nsr/nwu014
Kumar J, Kaur G, Ren R, Lu Y, Lin K, Li J, Huang Y, Patel A, Barton MC, Macfarlan T, et al. 2020. KRAB domain of ZFP568 disrupts TRIM28-mediated abnormal interactions in cancer cells. *NAR Cancer* **2:** zcaa007. doi:10.1093/narcan/zcaa007
Kumar S, Suleski M, Craig JM, Kasprowicz AE, Sanderford M, Li M, Stecher G, Hedges SB. 2022. TimeTree 5: an expanded resource for species divergence times. *Mol Biol Evol* **39:** msac174. doi:10.1093/molbev/msac174
Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9:** 357–359. doi:10.1038/nmeth.1923
Li X, Ito M, Zhou F, Youngson N, Zuo X, Leder P, Ferguson-Smith AC. 2008. A maternal-zygotic effect gene, *Zfp57*, maintains both maternal and paternal imprints. *Dev Cell* **15:** 547–557. doi:10.1016/j.devcel.2008.08.014
Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25:** 2078–2079. doi:10.1093/bioinformatics/btp352
Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30:** 923–930. doi:10.1093/bioinformatics/btt656
Liu Y, Zhang Y, Yin J, Gao Y, Li Y, Bai D, He W, Li X, Zhang P, Li R, et al. 2019. Distinct H3K9me3 and DNA methylation modifications during mouse spermatogenesis. *J Biol Chem* **294:** 18714–18725. doi:10.1074/jbc.RA119.010496
Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15:** 550. doi:10.1186/s13059-014-0550-8
Maezawa S, Hasegawa K, Yukawa M, Kubo N, Sakashita A, Alavattam KG, Sin HS, Kartashov AV, Sasaki H, Barski A, et al. 2018a. Polycomb protein SCML2 facilitates H3K27me3 to establish bivalent domains in the male germline. *Proc Natl Acad Sci* **115:** 4957–4962. doi:10.1073/pnas.1804512115
Maezawa S, Yukawa M, Alavattam KG, Barski A, Namekawa SH. 2018b. Dynamic reorganization of open chromatin underlies diverse transcriptomes during spermatogenesis. *Nucleic Acids Res* **46:** 593–608. doi:10.1093/nar/gkx1052
Maezawa S, Sakashita A, Yukawa M, Chen X, Takahashi K, Alavattam KG, Nakata I, Weirauch MT, Barski A, Namekawa SH. 2020. Super-enhancer switching drives a burst in gene expression at the mitosis-to-meiosis transition. *Nat Struct Mol Biol* **27:** 978–988. doi:10.1038/s41594-020-0488-3
Mouse Genome Sequencing Consortium. 2002. Initial sequencing and comparative analysis of the mouse genome. *Nature* **420:** 520–562. doi:10.1038/nature01262
Mueller JL, Mahadevaiah SK, Park PJ, Warburton PE, Page DC, Turner JM. 2008. The mouse X chromosome is enriched for multicopy testis genes showing postmeiotic expression. *Nat Genet* **40:** 794–799. doi:10.1038/ng.126
Mueller JL, Skaletsky H, Brown LG, Zaghlul S, Rock S, Graves T, Auger K, Warren WC, Wilson RK, Page DC. 2013. Independent specialization of the human and mouse X chromosomes for the male germ line. *Nat Genet* **45:** 1083–1087. doi:10.1038/ng.2705
Myers S, Bowden R, Tumian A, Bontrop RE, Freeman C, MacFie TS, McVean G, Donnelly P. 2010. Drive against hotspot motifs in primates implicates the *PRDM9* gene in meiotic recombination. *Science* **327:** 876–879. doi:10.1126/science.1182363
Namekawa SH, Park PJ, Zhang LF, Shima JE, McCarrey JR, Griswold MD, Lee JT. 2006. Postmeiotic sex chromatin in the male germline of mice. *Curr Biol* **16:** 660–667. doi:10.1016/j.cub.2006.01.066
Patel A, Horton JR, Wilson GG, Zhang X, Cheng X. 2016. Structural basis for human PRDM9 action at recombination hot spots. *Genes Dev* **30:** 257–265. doi:10.1101/gad.274928.115
Patel A, Yang P, Tinkham M, Pradhan M, Sun MA, Wang Y, Hoang D, Wolf G, Horton JR, Zhang X, et al. 2018. DNA conformation induces adaptable binding by tandem zinc finger proteins. *Cell* **173:** 221–233.e12. doi:10.1016/j.cell.2018.02.058
Patel L, Kang R, Rosenberg SC, Qiu Y, Raviram R, Chee S, Hu R, Ren B, Cole F, Corbett KD. 2019. Dynamic reorganization of the genome shapes the recombination landscape in meiotic prophase. *Nat Struct Mol Biol* **26:** 164–174. doi:10.1038/s41594-019-0187-0
Peaston AE, Evsikov AV, Graber JH, de Vries WN, Holbrook AE, Solter D, Knowles BB. 2004. Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. *Dev Cell* **7:** 597–606. doi:10.1016/j.devcel.2004.09.004
Pi W, Yang Z, Wang J, Ruan L, Yu X, Ling J, Krantz S, Isales C, Conway SJ, Lin S, et al. 2004. The LTR enhancer of ERV-9 human endogenous retrovirus is active in oocytes and progenitor cells in transgenic zebrafish and humans. *Proc Natl Acad Sci* **101:** 805–810. doi:10.1073/pnas.0307698100
Pontis J, Planet E, Offner S, Turelli P, Duc J, Coudray A, Theunissen TW, Jaenisch R, Trono D. 2019. Hominoid-Specific transposable elements and KZFPs facilitate human embryonic genome activation and control transcription in naive human ESCs. *Cell Stem Cell* **24:** 724–735.e5. doi:10.1016/j.stem.2019.03.012
Pontis J, Pulver C, Playfoot CJ, Planet E, Grun D, Offner S, Duc J, Manfrin A, Lutolf MP, Trono D. 2022. Primate-specific transposable elements shape transcriptional networks during human development. *Nat Commun* **13:** 7178. doi:10.1038/s41467-022-34800-w
Potrzebowski L, Vinckenbosch N, Marques AC, Chalmel F, Jégou B, Kaessmann H. 2008. Chromosomal gene movements reflect the recent origin and biology of therian sex chromosomes. *PLoS Biol* **6:** e80. doi:10.1371/journal.pbio.0060080
Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26:** 841–842. doi:10.1093/bioinformatics/btq033
Ramirez F, Ryan DP, Gruning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dundar F, Manke T. 2016. Deeptools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* **44:** W160–W165. doi:10.1093/nar/gkw257
R Core Team. 2018. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna. https://www.R-project.org/.
Rebollo R, Romanish MT, Mager DL. 2012. Transposable elements: an abundant and natural source of regulatory sequences for host genes. *Annu Rev Genet* **46:** 21–42. doi:10.1146/annurev-genet-110711-155621
Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011. Integrative genomics viewer. *Nat Biotechnol* **29:** 24–26. doi:10.1038/nbt.1754
Sakashita A, Maezawa S, Takahashi K, Alavattam KG, Yukawa M, Hu YC, Kojima S, Parrish NF, Barski A, Pavlicev M, et al. 2020. Endogenous retroviruses drive species-specific germline transcriptomes in mammals. *Nat Struct Mol Biol* **27:** 967–977. doi:10.1038/s41594-020-0487-4
Schultz DC, Ayyanathan K, Negorev D, Maul GG, Rauscher FJ III. 2002. SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific

methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes Dev* **16:** 919–932. doi:10.1101/gad.973302

Seah MKY, Wang Y, Goy PA, Loh HM, Peh WJ, Low DHP, Han BY, Wong E, Leong EL, Wolf G, et al. 2019. The KRAB-zinc-finger protein ZFP708 mediates epigenetic repression at RMER19B retrotransposons. *Development* **146:** dev170266. doi:10.1242/dev.170266

Senft AD, Macfarlan TS. 2021. Transposable elements shape the evolution of mammalian development. *Nat Rev Genet* **22:** 691–711. doi:10.1038/s41576-021-00385-1

Sherman BT, Hao M, Qiu J, Jiao X, Baseler MW, Lane HC, Imamichi T, Chang W. 2022. DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res* **50:** W216–W221. doi:10.1093/nar/gkac194

Sin HS, Ichijima Y, Koh E, Namiki M, Namekawa SH. 2012. Human postmeiotic sex chromatin and its impact on sex chromosome evolution. *Genome Res* **22:** 827–836. doi:10.1101/gr.135046.111

Sin HS, Kartashov AV, Hasegawa K, Barski A, Namekawa SH. 2015. Poised chromatin and bivalent domains facilitate the mitosis-to-meiosis transition in the male germline. *BMC Biol* **13:** 53. doi:10.1186/s12915-015-0159-8

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

Storer J, Hubley R, Rosen J, Wheeler TJ, Smit AF. 2021. The Dfam community resource of transposable element families, sequence models, and genome annotations. *Mob DNA* **12:** 2. doi:10.1186/s13100-020-00230-y

Sundaram V, Cheng Y, Ma Z, Li D, Xing X, Edge P, Snyder MP, Wang T. 2014. Widespread contribution of transposable elements to the innovation of gene regulatory networks. *Genome Res* **24:** 1963–1976. doi:10.1101/gr.168872.113

Tadepally HD, Burger G, Aubry M. 2008. Evolution of C2H2-zinc finger genes and subfamilies in mammals: species-specific duplication and loss of clusters, genes and effector domains. *BMC Evol Biol* **8:** 176. doi:10.1186/1471-2148-8-176

Takahashi N, Coluccio A, Thorball CW, Planet E, Shi H, Offner S, Turelli P, Imbeault M, Ferguson-Smith AC, Trono D. 2019. ZNF445 is a primary regulator of genomic imprinting. *Genes Dev* **33:** 49–54. doi:10.1101/gad.320069.118

Thompson PJ, Macfarlan TS, Lorincz MC. 2016. Long terminal repeats: from parasitic elements to building blocks of the transcriptional regulatory repertoire. *Mol Cell* **62:** 766–776. doi:10.1016/j.molcel.2016.03.029

Trono D. 2015. Transposable elements, polydactyl proteins, and the genesis of human-specific transcription networks. *Cold Spring Harb Symp Quant Biol* **80:** 281–288. doi:10.1101/sqb.2015.80.027573

Turner JM. 2015. Meiotic silencing in mammals. *Annu Rev Genet* **49:** 395–412. doi:10.1146/annurev-genet-112414-055145

Wang Y, Wang H, Zhang Y, Du Z, Si W, Fan S, Qin D, Wang M, Duan Y, Li L, et al. 2019. Reprogramming of meiotic chromatin architecture during spermatogenesis. *Mol Cell* **73:** 547–561.e6. doi:10.1016/j.molcel.2018.11.019

Watanabe T, Cheng EC, Zhong M, Lin H. 2015. Retrotransposons and pseudogenes regulate mRNAs and lncRNAs via the piRNA pathway in the germline. *Genome Res* **25:** 368–380. doi:10.1101/gr.180802.114

Wolf G, de Iaco A, Sun MA, Bruno M, Tinkham M, Hoang D, Mitra A, Ralls S, Trono D, Macfarlan TS. 2020. KRAB-zinc finger protein gene expansion in response to active retrotransposons in the murine lineage. *eLife* **9:** e56337. doi:10.7554/eLife.56337

Yang P, Wang Y, Macfarlan TS. 2017. The role of KRAB-ZFPs in transposable element repression and mammalian evolution. *Trends Genet* **33:** 871–881. doi:10.1016/j.tig.2017.08.006

Zamudio N, Bourc'his D. 2010. Transposable elements in the mammalian germline: a comfortable niche or a deadly trap? *Heredity (Edinb)* **105:** 92–104. doi:10.1038/hdy.2010.53

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, et al. 2008. Model-based Analysis of ChIP-Seq (MACS). *Genome Biol* **9:** R137. doi:10.1186/gb-2008-9-9-r137

Zhou S, Sakashita A, Yuan S, Namekawa SH. 2023. Retrotransposons in the mammalian male germline. *Sex Dev* **16:** 404–422. doi:10.1159/000520683

Zhu LJ, Gazin C, Lawson ND, Pagès H, Lin SM, Lapointe DS, Green MR. 2010. ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. *BMC Bioinformatics* **11:** 237. doi:10.1186/1471-2105-11-237

Zhu Z, Li C, Yang S, Tian R, Wang J, Yuan Q, Dong H, He Z, Wang S, Li Z. 2016. Dynamics of the transcriptome during human spermatogenesis: predicting the potential key genes regulating male gametes generation. *Sci Rep* **6:** 19069. doi:10.1038/srep19069

Received August 13, 2024; accepted in revised form February 6, 2025.