Chromosome-specific maturation of the epigenome in the Drosophila male germline

James Anderson[1], Steven Henikoff[1,2*], Kami Ahmad[1*]

[1]Basic Sciences Division, Fred Hutchinson Cancer Center; Seattle WA 98109 USA

[2]Howard Hughes Medical Institute; Chevy Chase MD USA

Abstract

As cells differentiate in the Drosophila male germline, the *Y* chromosome becomes broadly upregulated in spermatocytes prior to meiosis, whereas the *X* and *4*th chromosomes become gradually inactive. Here we characterize the chromatin marks associated with chromosome-specific regulation in the male germline. We profile two distinct features of active chromatin, dimethylation of histone H3 at lysine 4 (H3K4me2) and elongating RNA Polymerase II. Our profiling of these active marks reveals a unique distribution of H3K4me2 at late-germline genes. Furthermore, we demonstrate that enrichment of RNA Polymerase II Serine-2 phosphate (RNAPolIIser2phos) on the X-linked genes in spermatocytes is no different from RNAPolIIser2phos enrichment at autosomal genes. We find that histone marks typically associated with chromosomal silencing in other tissues and developmental contexts are not enriched on the *X* or *4th* chromosomes in Drosophila spermatocytes. Instead, reduced expression can be attributed to the lack of *X* chromosome dosage compensation in the male germline. Interestingly, our profiling also uncovers prominent ubiquitinylation of histone H2A and lysine-16 acetylation of histone H4 across the activating *Y* chromosome in spermatocytes. Our results reveal a surprising degree of diversity in chromosome-specific features of the male germline.

Significance Statement

The *Drosophila* male germline uses distinct mechanisms of gene regulation that do not occur in somatic tissues, including widespread activation of the *Y* chromosome and reduced transcription from the *X* and *4th* chromosomes. It has been unclear whether reduced *X* chromosome expression results from direct silencing processes, or from the lack of dosage compensation in this tissue. To resolve this, we profiled features of active and silent chromatin in Drosophila spermatocytes, revealing chromosome-specific patterns of histone modification. Our results demonstrate that the *X* chromosome in the male germline does not acquire modifications typically associated silencing, but also lacks those of dosage compensation. We also find that a unique combination of histone marks is associated with high-level *Y* chromosome expression.

Introduction

In the male germline of *Drosophila melanogaster*, the *X*, *Y*, and *4th* chromosomes each have distinct chromosomal features which uniquely impact their expression over the course of differentiation. The *Y* chromosome is heterochromatic and gene-poor, but becomes highly active in primary spermatocytes, forming long, diffuse loops of actively transcribing genes within the interior of the nucleus [1], and many of these genes are essential for male fertility [2] [3]. Previous studies have suggested that expression from the *X* chromosome becomes progressively reduced as cells differentiate in the male germline, as does expression from the small autosomal "dot" *4th* chromosome [4] [5]. Recently there has been great interest in explaining the chromatin features underlying reduced chromosomal expression in *Drosophila*.

Here, we apply CUT&Tag chromatin profiling [6] on the adult testis of *D. melanogaster* to characterize the chromatin features of the spermatogenic transcription program. We reveal trends of chromatin regulation that are unique to the spermatogenic transcription program. We describe a germline-specific pattern of active histone modification around germline-specific promoters that extends into gene bodies. We also describe the changing chromosome-specific enrichment of elongating RNA Polymerase II in somatic tissues *vs* spermatocytes. Regarding repressive chromatin, we identify a broad gain of heterochromatic histone modifications across most chromosomes in spermatocytes. However, the *X* chromosome does not accumulate repressive chromatin features, supporting a model where this chromosome is not silenced in the Drosophila male germline. Surprisingly, spermatocytes acquire high levels of ubiquitinylated histone H2A (uH2A) across the *Y* chromosome, *4th* chromosome, and at many transposable elements, but many of these are transcriptionally active in this stage, pointing to a specialized role for this otherwise repressive histone modification in the male germline.

Results

Profiling active promoters in the Drosophila testes

The testis forms a spiraled tube, with germline stem cells (GSCs) residing at the apical tip (Figure 1A) [7]. Following asymmetric division of the GSC, one daughter cell migrates distally in the tissue and proceeds through four rounds of mitosis as spermatogonia. These cells undergo one final S phase of the cell cycle and then grow for roughly 3 days as primary spermatocytes before meiosis. During this spermatocyte stage cells upregulate more than one thousand genes for meiosis and for differentiation into mature spermatozoa [7].

To identify active promoters in *Drosophila* spermatogenesis, we performed CUT&Tag profiling for dimethylation of histone H3 at lysine 4 (H3K4me2), which marks active promoters and enhancers [8], on the testes of one-day-old adult males. To distinguish promoters active in the germline from those active in somatic cell types, we also profiled the H3K4me2 mark in wing imaginal discs from male larvae. We performed

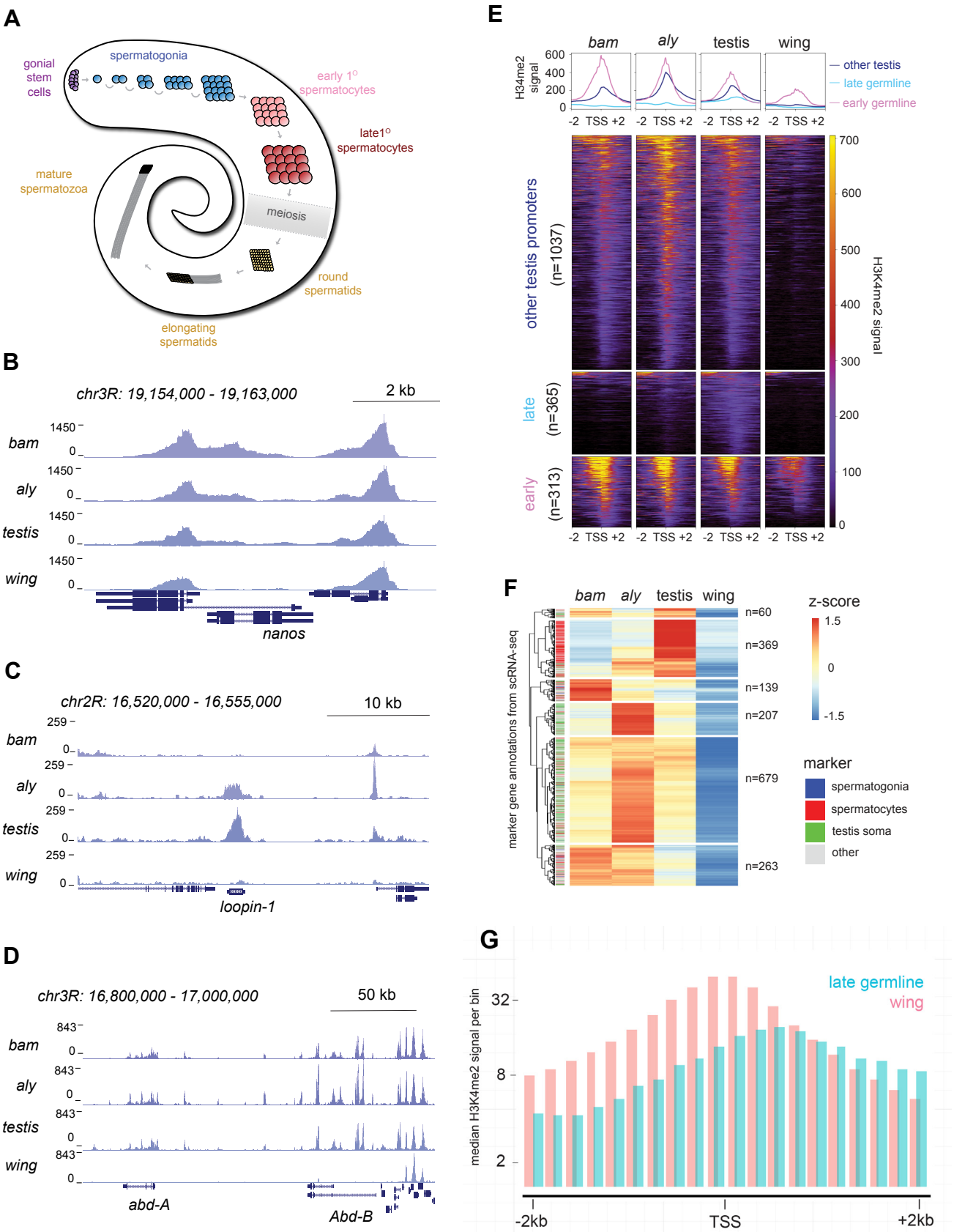


Figure 1. CUT&Tag for the H3K4me2 histone modification classifies active promoters by tissue- and stage-specificity. (A) The testis of *D. melanogaster* is a specialized, differentiating tissue. Somatic cells of the testis (not shown) include cyst cells, sheath cells, pigment cells, and terminal epithelium. Cell types of the male germline, from early to later in developmental time, comprise germline stem cells, spermatogonia, early primary spermatocytes, late primary spermatocytes, round spermatids, elongating spermatozoa, and mature spermatozoa.

(B) Tracks for the H3K4me2 histone modification in *bam* testis, *aly* testis, wild-type testis, and wing imaginal disc centered on the *nanos* gene, a marker of germline stem cells.

each profiling with 3-5 replicates, and between 100,000 - 10 million reads were mapped to a repeat-masked version of the *Drosophila* dm6 genome assembly.

To annotate testis-specific promoters from profiling data, we summed the enrichment of H3K4me2 around all annotated promoters in testis and wing disc (Supplementary Table 1. We identified 1,128 promoters (representing 941 genes) with differentially enriched promoter signals in testis samples, and 1,217 promoters (1,069 genes) in wing discs (Supplementary Figure S1, Supplementary Table 1). The list of promoters specifically active in the testis includes genes that are highly expressed in male germline cells (e.g. *loop-in-1*, Figure 1C) and genes that are expressed in the somatic cells of the testis (eg. *abdominal-A* and *Abdominal-B*, Figure 1D). In addition, some testis-enriched signals mark alternative promoters for commonly expressed genes [9]. Conversely, promoters with H3K4me2 signal enriched in wing discs mark somatically active promoters, including those for genes that are specifically active in wing tissue (eg. *nubbin*, Supplementary Table 1).

To further refine this set of active testis-specific promoters, we profiled the H3K4me2 modification in testes from two stage-arrest mutants. The germline of *bag-of-marbles* (*bam*) stage-arrest mutant males arrests in the spermatogonial stage, so their testes are full of early germline cells [10]. The germline of *always early* (*aly*) mutants arrests in the early primary spermatocyte stage, thus enriching for these early primary spermatocytes [11]. Profiling testes from these mutants can thereby identify when testis-specific promoters become active and enrich for promoters that may have activity limited to these stages. For example, the promoter for the early germline marker *nanos* (*nos*) is marked with H3K4me2 in *bam* testis, with less signal in in *aly* and wild-type testes (Figure 1B, Supplementary Table 1). Profiling also reveals the timing of activation for late germline genes; the promoter of the highly expressed germline-specific gene *loopin-1* lacks H3K4me2 signal in *bam* testes but is heavily marked in *aly* testes and in wild-type testis (Figure 1C). Finally, many promoters for post-meiotic spermatogenesis are unique to wild-type testes (e. g. *Mst36Fa* Supplementary Table 1), as the later stages of spermatogenesis are absent from *bam* and *aly* testes.

We combined the lists of active promoters in wild-type, *bam*, and *aly* testes identifying 1,715 promoters representing 1,429 genes (Supplementary Figure S2). Overall, the set of 365 promoters uniquely detectable in wild-type testes have moderate levels of H3K4me2, consistent with the relatively low signals at developmentally regulated genes (Figure 1E). The 313 promoters that are uniquely significant in *bam* testes also show H3K4me2 signal in wild-type and *aly* testes, and more moderately in wing discs (Figure 1E). Since wild-type and *aly* testes also include early germline cells, some signal at these promoters in all testes samples is expected. Finally, since early spermatocytes are present in both wild-type and in *aly* mutant testes, promoters uniquely identified in *aly* testes show moderate signal in wild-type testis as well.

To define when in spermatogenesis these promoters become active, we normalized the H3K4me2 signal at promoters as z-scores, and hierarchically clustered these sites. Six groups are apparent, with 139 promoters most active in *bam* mutant testes, 806 promoters most active in *aly* mutants (represented by two clusters in the heatmap, with 207 and 679 promoters respectively), and 365 most active in wild-type testes (Figure 1F).

We compared promoter activity to cell-type-specific gene expression determined by single-cell RNA-seq (scRNA-seq) of the *Drosophila* testis [5]. Aligning H3K4me2 promoter scores with marker gene expression reveals a strong correspondence in germline cell types between these two approaches, supporting the idea that CUT&Tag profiling of H3K4me2 on stage-arrest mutants is a reliable approach for mapping the timing of promoter activation during differentiation in the male germline.

The late-germline class of promoters is distinct for two reasons. First, the gain of H3K4me2 at these late-germline genes is much more subtle than the dramatic H3K4me2 enrichment at the “early-germline” or “all other testis-specific” promoter groups (Figure 1E). This is apparent from the dim signal for H3K4me2 in the late-germline group of the heatmap (Figure 1E). Second, accumulation of H3K4me2 at late germline-specific genes peaks ~800 bp downstream of the TSS (Figure 1G). This is surprising given that H3K4me2 enrichment is typically centered around the TSSs of tissue-specific genes, such as wing-specific genes (Figure 1G).

Finally, comparison of H3K4me2 enrichment in testis and wing discs at all annotated promoters separated according to their chromosomal location reveals there are 886 promoters on the *X* chromosome that are significantly enriched in wing *vs* 460 promoters significantly enriched in testis (FDR < 0.05) (Supplementary Figure S3), consistent with previous

(C) Tracks for the H3K4me2 histone modification in *bam* testis, *aly* testis, wild-type testis, and wing imaginal disc centered on the *loopin-1* gene, a marker of spermatocytes.
(D) Tracks for the H3K4me2 histone modification in *bam* testis, *aly* testis, wild-type testis, and wing imaginal disc centered on the *abd-A* and *Abd-B* genes which are expressed in somatic cells of the testis and repressed in wing discs.
(E) Classification of testis-specific promoters into early-germline promoters (n=313), late-germline promoters (n=365), and common testis-specific promoters (n=1037). H3K4me2 signal from CUT&Tag on *bam* testis, *aly* testis, wild-type testis, and wing is shown centered on these promoters, represented as profile plots and heatmaps.
(F) Hierarchical clustering of H3K4me2 z-normalized signal at 1715 testis-specific promoters. Cell type annotations from scRNA-seq are on the left.
(G) Binned display H3K4me2 signal at 1217 promoters active in wing discs (pink bars) and at 365 late-germline-specific promoters (blue bars). H3K4me2 signal was divided into 200 bp bins ±2kb around each promoter. The height of each bin indicates median H3K4me2 signal. Genes shorter than the distance of each respective bin were excluded when calculating the median.

reports that the X-linked genes are less active in the male germline [5]. This trend on the *X* chromosome contrasts with the trend of promoters on the *Y* chromosome: every *Y*-linked promoter with significant tissue-specific enrichment is specific in testis and is not expressed in wing. This is consistent with the of broad activation of the *Y* chromosome in spermatocytes.

FACS-CUT&Tag profiling of RNA Polymerase II in primary spermatocytes

Given that spermatocytes are the cell type where the most striking transcriptional changes occur in the male germline, we used FACS to isolate spermatocytes from the testis for chromatin profiling. The *hephaestus* (*heph*) gene is expressed specifically in spermatocytes and in no other cell type of the testis and encodes an RNA-binding factor that binds *Y* chromosome transcripts [1]. A *heph-GFP* transgene visually marks spermatocytes (Figure 2A) [1]. We thus per-

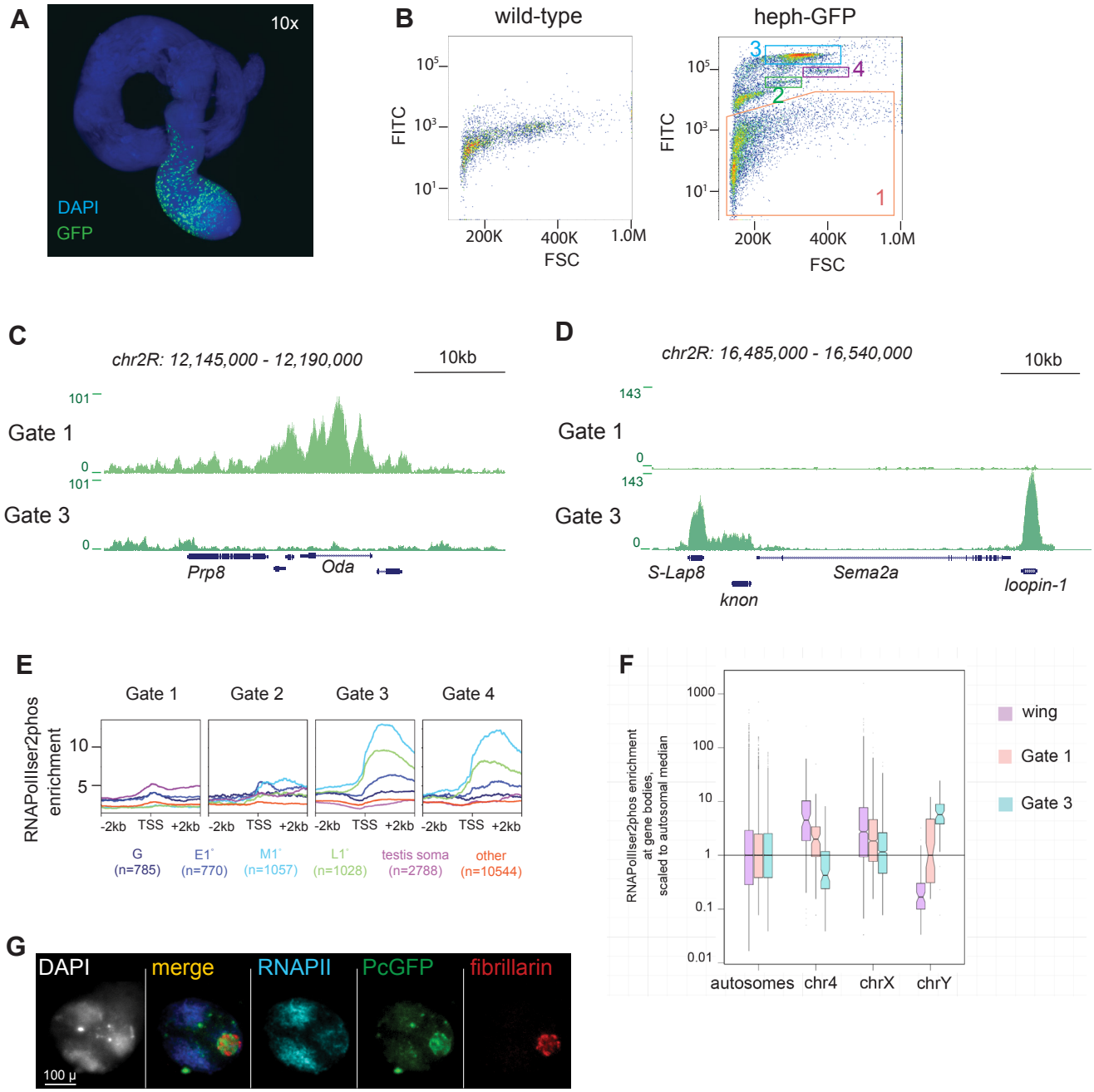


Figure 2. FACS-CUT&Tag for RNA Polymerase II in the male germline. (A) A *heph-GFP* testis stained with DAPI (blue) at 10x magnification. Spermatocytes show intense GFP signal (green).

(B) (Left) FSC-vs-GFP plot of recorded events from a FACS experiment on 10 dissociated w^{1118} testes, as a negative control. (Right) FSC-vs-GFP plot of all recorded events from a FACS experiment on 40 dissociated *heph-GFP* testes.

(C) Genome browser tracks of RNAPolIIser2phos at the spermatogonial marker gene *prp8* and the testis soma marker gene *Oda* from FACS-CUT&Tag on Gate 1 and Gate 3 cells from heph-GFP testis.

(D) Genome browser tracks of RNAPolIIser2phos at the spermatocyte marker genes *S-Lap8* and *loopin-1* from FACS-CUT&Tag on Gate 1 and Gate 3 cells from *heph-GFP* testis.

(E) Profile plots centered on TSSs showing average signal of RNAPolIIser2phos profiled by FACS-CUT&Tag. The color-coded cate-

formed fluorescence-activated cell sorting (FACS) on dissociated testes from *heph-GFP* flies, recording the forward scatter (FSC) and GFP signal of each event (Figure 2B). FACS profiles from *heph-GFP* samples display a large proportion of events with high GFP signal. These signals are absent in FACS profiles of wild-type testes, confirming that these events represent GFP-labeled spermatocytes. These GFP-labeled spermatocytes also appear as clusters of cells with distinct sizes (as inferred from FSC), consistent with the progressive growth of spermatocytes as they approach meiosis [7].

We collected four distinct FACS populations from dissociated *heph-GFP* samples based on GFP and FSC signals. In a typical FACS experiment with 40 dissociated *heph-GFP* testes as input, we collected ~6,000 events in that lack GFP signal (Gate 1), ~500 events with moderate GFP signal and moderate size (Gate 2), ~5,000 events with very high GFP and moderate size (Gate 3), and ~500 events with high GFP and large size (Gate 4). Since Gate 1 resembles the FACS profile of wild-type testis and lacks GFP signal, this Gate must contain a mixture of spermatogonia, post-meiotic spermatids, and somatic cells of the testis. The GFP-positive Gates 2, 3, and 4 represent progressively larger spermatocytes. The cells of Gate 4 are dimmer than Gate 3 and are the largest, thus representing the latest stage of spermatocytes when heph-GFP signal decreases [1]). Because Gate 3 contains the most FACS events from the spermatocyte gates, we focused further analysis on these cells.

For each FACS-CUT&Tag experiment, we FACS-isolated cells and then performed profiling for each chromatin epitope (Supplementary Table 5). Genome browser tracks demonstrate the high technical quality of these FACS-CUT&Tag profiles for RNA Polymerase II Serine-2 phosphate (RNAPolIIser2phos), the elongating form of RNA Polymerase II. For example, the genes *prp8* and *Oda* are marker genes for spermatogonia and cyst cells respectively, so these genes should be enriched with elongating RNA Polymerase II only in Gate 1 and not in Gate 3. Genome browser tracks confirm that this is the case (Figure 2C). Additionally, the genes *S-Lap8* and *loopin-1* are marker genes for spermatocytes, so these genes should be enriched with elongating RNA Polymerase II only in Gate 3 and not in Gate 1. Genome browser tracks also confirm that this is the case (Figure 2D).

To more comprehensively describe genes which are expressed in each sorted FACS population, we profiled RNAPolIIser2phos by FACS-CUT&Tag and summarized enrichment at all annotated promoters, separated by marker gene annotation (Figure 2E). As expected, Gate 1 shows measurable RNAPolIIser2phos signal for somatic cell type marker genes and lacks signal at marker genes for spermatocyte stages. In contrast, in Gate 2 samples, moderate RNAPolIIser2phos signal accumulates at mid-spermatocyte marker genes, while in Gates 3 and 4 there is dramatic enrichment at markers for all spermatocyte stages (Figure 2E).

To assess chromosomal transcriptional activity in the male germline, we summed the enrichment of RNAPolIIser2phos across all gene bodies in wing disc, testes Gate 1, and spermatocyte Gate 3, separating genes according to their chromosomal location and scaling every gene-body-sum to the median value of all gene-body-sums in the “autosomes” category. This approach reveals that the median enrichment of RNAPolIIser2phos on X-linked genes in spermatocytes is similar to RNAPolIIser2phos enrichment on autosomal genes in spermatocytes. A different trend is apparent on the *Y* chromosome, where in Gate 1 this chromosome is more enriched with RNAPolIIser2phos than in wing disc cells, and even further enriched in Gate 3 cells. These patterns are consistent with the dramatic upregulation of male fertility factors on the *Y* chromosome.

To further characterize the pattern of RNAPolIIser2phos of in the male germline, we assessed the cytological enrichment of RNAPolIIser2phos by immunostaining primary spermatocytes from *Polycomb-GFP* (*Pc-GFP*) flies [12]. Pc marks the nucleolus in spermatocytes and thus provides a useful cytological landmark for the *X* chromosome, where some of the rDNA genes are located. Of the three DAPI-bright chromosomal compartments in each nucleus, cytology reveals reduced signal for elongating RNA Polymerase II on the *X* chromosome abutting the nucleolus (Figure 2G), as has been previously described [10].

Profiling silencing chromatin marks in spermatocytes

In mammalian species, *X* chromosome inactivation in male germlines is associated with repressive histone modifications [13]. We therefore profiled three repressive histone modifications that drive silencing in other tissues: ubiquitinylation of histone H2A at lysine-119 (uH2A), associated with repression by Polycomb Repressive Complex 1 (PRC1) [14]; dimethylation of histone H3 at lysine-9 (H3K9me2) [15]; and H3K27me3, associated with repression by Polycomb Repressive Complex 2 (PRC2) [16]. We performed FACS-CUT&Tag profiling of uH2A, H3K9me2, and H3K27me3 in sorted spermatocytes (Figure 3). Profiles of repressive marks reveal chromatin features that are especially prominent on the *Su(Ste)* repeat clusters, the gene *Ppr-Y*, and the large cluster of genes including the *Mst77Y* gene (Figure 3A).

To get a broader picture of the distribution of uH2A, H3K9me2, and H3K27me3 genome-wide, we plotted these profiles as Circos plots (Figures 3B-3D). From inner ring to outer ring, these Circos plots show the profiles of the IgG negative control, wing discs, testes Gate 1, and spermatocytes Gate 3. The outermost ring indicates the $\log_2$ fold-change (log-FC) enrichment in Gate 3 / Gate 1, thus giving a sense of how enriched that mark is in spermatocytes compared to all other testis cell types. The heterochromatic histone modifi-

gories in these profile plots correspond with the marker gene annotations from scRNA-seq [5].

(F) Enrichment scores of RNAPolIIser2phos across all gene bodies in wing discs, testes samples, and in spermatocytes, separated by chromosomal location. Scores were scaled to the median score on autosomes.

(G) Immunostaining of germline cells for RNAPolIIser2phos to characterize the cytological location of elongating RNA polymerase.

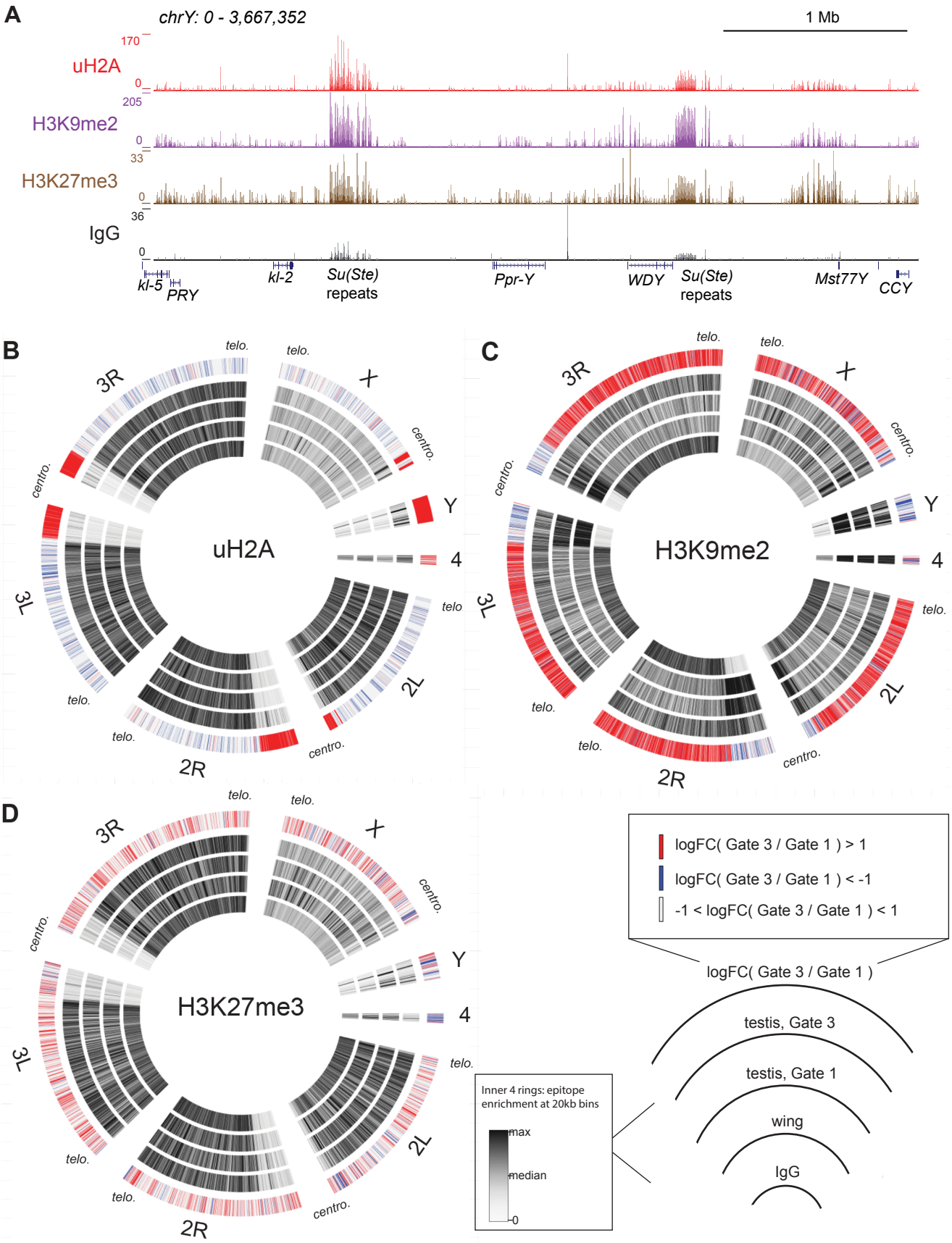


Figure 3. CUT&Tag of repressive histone marks in spermatocytes. (A) Genome browser tracks at the entire *Y* chromosome, showing FACS-CUT&Tag signal from Gate 3 profiles of the uH2A (red), H3K9me2 (purple), and H3K27me3 (brown) histone modifications, and IgG negative control (black).

(B-D) Circos plots of uH2A, H3K9me2, and H3K27me3 signal respectively. In each Circos plot, the four innermost rings indicate enrichment of the epitope of interest at all consecutive 20 kb genomic bins in: IgG-negative control (innermost ring), *heph-GFP* wing disc (next ring up), FACS Gate 1 from *heph-GFP* testis (next ring up), and FACS Gate 3 from *heph-GFP* testis (next ring up). The

black-white scaling is performed independently per ring, such that white indicates enrichment of 0 in any bin, black indicates the maximally enriched bin for that epitope, and gray scaling corresponds with intermediate enrichment. The outermost ring of each Circos plots indicates $\log_2$ fold-change of epitope enrichment from FACS Gate 3 / Gate 1 at all consecutive 20 kb genomic bins, color-coded according to logFC values (red, logFC>1; blue, logFC<-1; gray, logFC >-1 and <1). Thus, red indicate bins where the epitope is enriched in spermatocytes relative to all other cell types of the testis, and blue indicates bins where the epitope is depleted in spermatocytes relative to all other cell types of the testis.

cation H3K9me2 is increased in spermatocytes on all chromosomes except the *Y* chromosome, consistent with the transcriptional activation of this chromosome in this cell type (Figure 3C). In contrast, the Polycomb repressive modification H3K27me3 is largely invariant in spermatocytes (Figure 3D). However, Circos plots revealed a dramatic accumulation of the uH2A histone modification on the *4th* chromosome and on the *Y* chromosomes in spermatocytes (Figure 3B). The chromosome-specific trends of uH2A, H3K9me2, and H3K27me3 in spermatocytes is summarized as boxplots (Supplementary Figures S4) and show that uH2A is uniquely increased on the *Y* and *4th* chromosomes in spermatocytes. However, none of these three repressive modifications become specifically enriched on the *X* chromosome in spermatocytes. Thus, there is no evidence that the *X* chromosome is silenced in the male germline.

Since profiling revealed *Y* chromosome-specific accumulation of uH2A in spermatocytes, we assessed the cytological pattern of this modification by immunostaining (Figure 4). This cytology shows the accumulation of a uH2A-modified region in the nucleus beginning in mid-primary spermatocytes, which then appears to expand later into a much more prominent uH2A body in late-primary spermatocytes. To confirm this timing, we immunostained germline nuclei from *bam* and *aly* stage-arrest mutants. The early-germline nuclei from *bam* testes lack uH2A staining, while germline nuclei from *aly* testes have only a small uH2A body. Thus, the uH2A body accumulates during the early spermatocyte stage and expands as spermatocytes develop.

The uH2A body consistently abuts the nucleolus, consistent with it being uH2A modification of the *Y* chromosome. To confirm this chromosomal identity, we immunostained male larvae lacking a *Y* chromosome; in spermatocytes from these males only small puncta of uH2A staining are visible. This supports the idea that the uH2A body is the *Y* chromosome, which is transcriptionally active in this stage.

Loss of H3K9me2 and gain of uH2A histone modifications is not limited to the *Y* chromosome, but also occurs in pericentromeric regions (Figure 3B,C). We also investigated trends of enrichment for uH2A, H3K9me2, H3K27me3, and RNAPIIS2p at transposons (Supplementary Figure S5A-E, Supplementary Tables 8-12). Indeed, a large number of transposon families show increased transcription specifically in FACS-isolated spermatocytes (Supplementary Figure S5E). This transcriptional activity can account for the loss of H3K9me2 in pericentromeric regions. To further characterize the relationship between transcription and the uH2A modification, we plotted signal for uH2A and RNAPolIIser2phos in sorted spermatocytes at gene promoters and across transposon consensus sequences (Figure 5; Supplementary Figure S6). Strikingly, there is a strong correspondence between the uH2A modification and RNAPolIIser2phos at *Y*-linked promoters and at transposons (Figure 5A,B). In contrast, there is little uH2A at autosomal and *X*-linked promoters, regardless of transcriptional activity (Supplementary Figure S6B). Furthermore, the enrichment of uH2A at *Y*-linked promoters and at transposons is unique to spermatocytes (Supplementary Figures S6A,C).

The X chromosome lacks dosage compensation in spermatocytes

In somatic cells of males, gene expression on the single *X* chromosome is approximately doubled by the action of the Male Sex-Lethal (MSL) complex, which coats the male *X* chromosome and catalyzes acetylation of histone H4 at lysine 16 (H4K16Ac) [17]. Previous cytological and transcriptomics evidence has demonstrated that both MSL and H4K16Ac are depleted from the later male germline, suggesting that dosage compensation does not occur in the male germline [18] [19]. To assess H4K16Ac enrichment genome-wide in spermatocytes, we profiled this modification in male wing disc cells, in testes, and in FACS-isolated spermatocytes. The H4K16Ac modification is widespread across the genome, consistent with its association with transcriptional activity, and in wing disc (somatic) cells is noticeably enriched across the male *X* chromosome (Figure 6; Supplementary Figure S7A,B). In contrast, the *X* chromosome is depleted for H4K16ac relative to the bulk of the genome in spermatocytes (Figure 6), indicating the lack of dosage compensation of this chromosome in this stage. A small number of loci at the base of X chromosome gain H4K16Ac in spermatocytes, which includes the microRNAs *mir10404-1 – mir10404-5* and the genes *CG40813* and *CG41562* (Supplementary Figure S8). Profiling of RNAPolIIser2phos confirms that this cluster of genes is also transcribed in spermatocytes (Supplementary Figure S8). These profiling results suggest that the decreased expression of X-linked genes in spermatocytes is simply attributable to the loss of dosage compensation in the male germline.

Strikingly, there is substantial gain of H4K16ac across the *Y* chromosome in spermatocytes (Figure 6; Supplementary Figure S7), indicating that this modification accompanies transcriptional activation of the *Y* chromosome. Just as one sex chromosome loses H4K16Ac, the other gains it.

Discussion

Although gene expression in the testis of Drosophila has been detailed, [20] [5] [21], the chromatin features have been less thoroughly characterized, in part because of the complexity and small size of the tissue. We have addressed this by performing CUT&Tag chromatin profiling for both active and repressive chromatin marks, as well as elongating RNA polymerase II. We profiled whole testes from wildtype and from stage-arrest mutants to parse the chromatin fea-

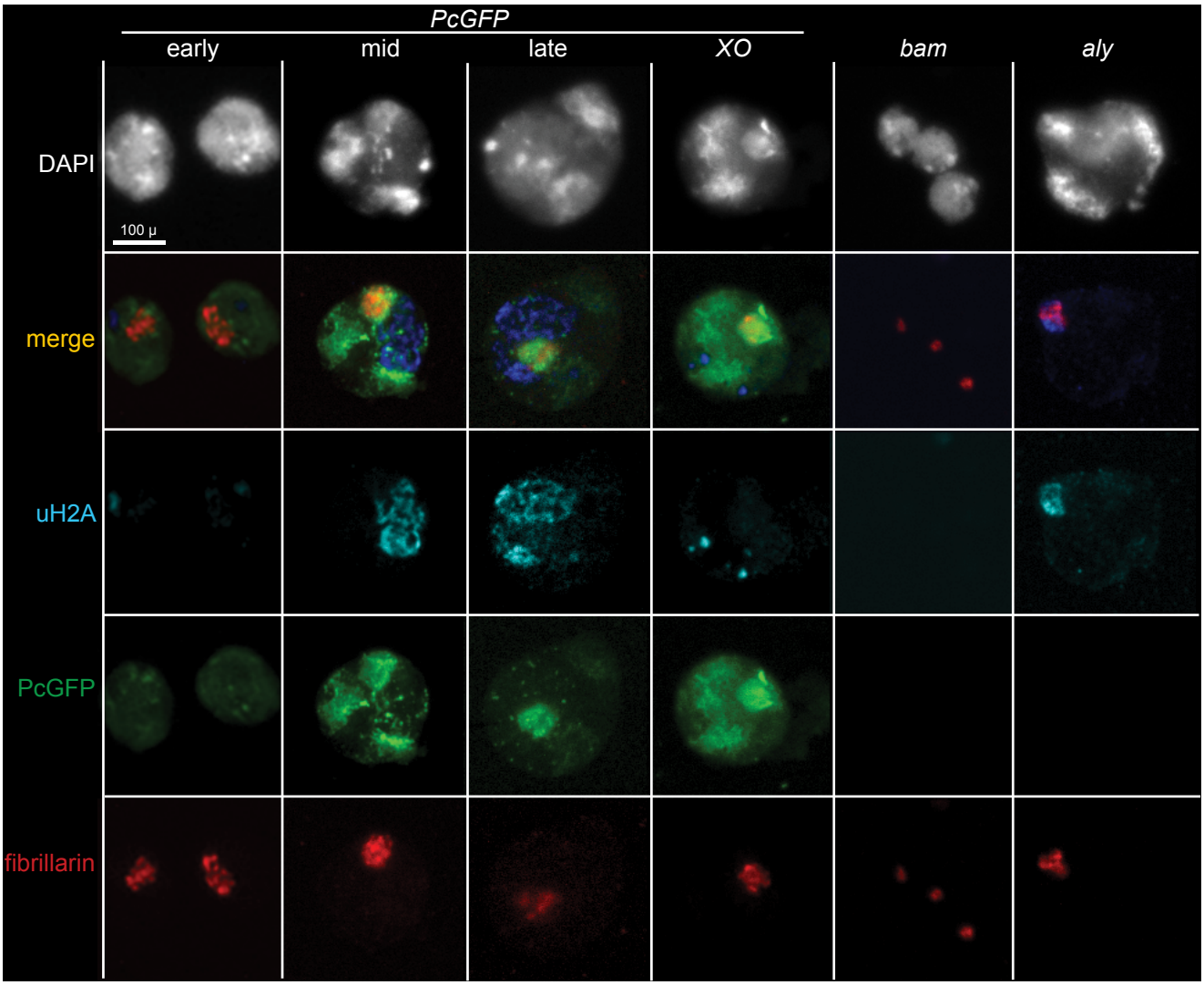


Figure 4. The spermatocyte uH2A body is the *Y* chromosome. Pc-GFP germline cells immunostained for the uH2A histone modification and fibrillarin.

tures of the early and late germline. Additionally, we profiled the elongating form of RNA polymerase II in isolated spermatocytes. These profiles provide a broad overview of the epigenomic status in the differentiating germline.

The late germline promoters are unique in having a smeary

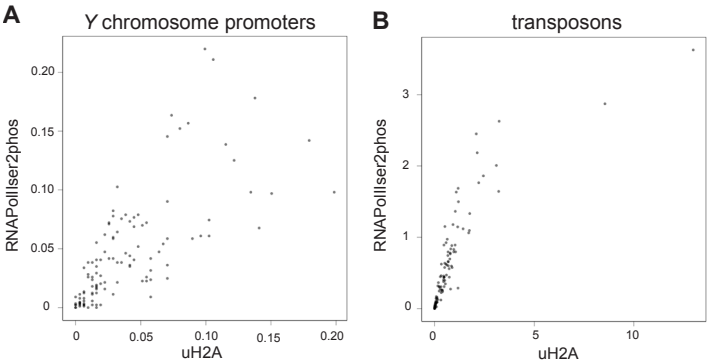


Figure 5. The uH2A histone modification in spermatocytes is enriched at transcribing transposons and *Y*-linked promoters. (A) uH2A signal of uH2A vs RNAPolIIser2phos at 1 kb bins centered around the TSSs of genes on chromosome *Y*.
(B) uH2A signal vs RNAPolIIser2phos at 128 transposon consensus sequences from [34] [35].

distribution of H3K4me2 downstream of the TSS, within gene bodies. This contrasts with the typical distribution of H3K4me2 at other tissue-specific genes, which tends to be centered on the TSS [8]. Our observation of smeared distribution of the H3K4me2 modification in Drosophila spermatocytes is reminiscent of the broad H3K4me3 domains at cell identity genes in mammals; these both might result from increased polymerase pausing at developmental genes [22] [23].

We profiled several repressive marks in the Drosophila male germline. Although the H3K27me3 histone modification marks Polycomb silencing of tissue-specific genes [16], H3K27me3 landscapes do not differ significantly between spermatocytes and the rest of the testis. This is consistent with previous studies which have demonstrated cytologically that the H3K27me3 writer Enhancer of zeste (E(z)) protein is absent from spermatocytes [10]. The heterochromatic H3K9me2 histone modification in spermatocytes is lost from the *Y* chromosome, pericentric heterochromatin, and transposons, paralleling the activation of these regions in spermatocytes.

Is the X chromosome silenced in spermatocytes?
In mammals the *X* and *Y* chromosomes become progres-

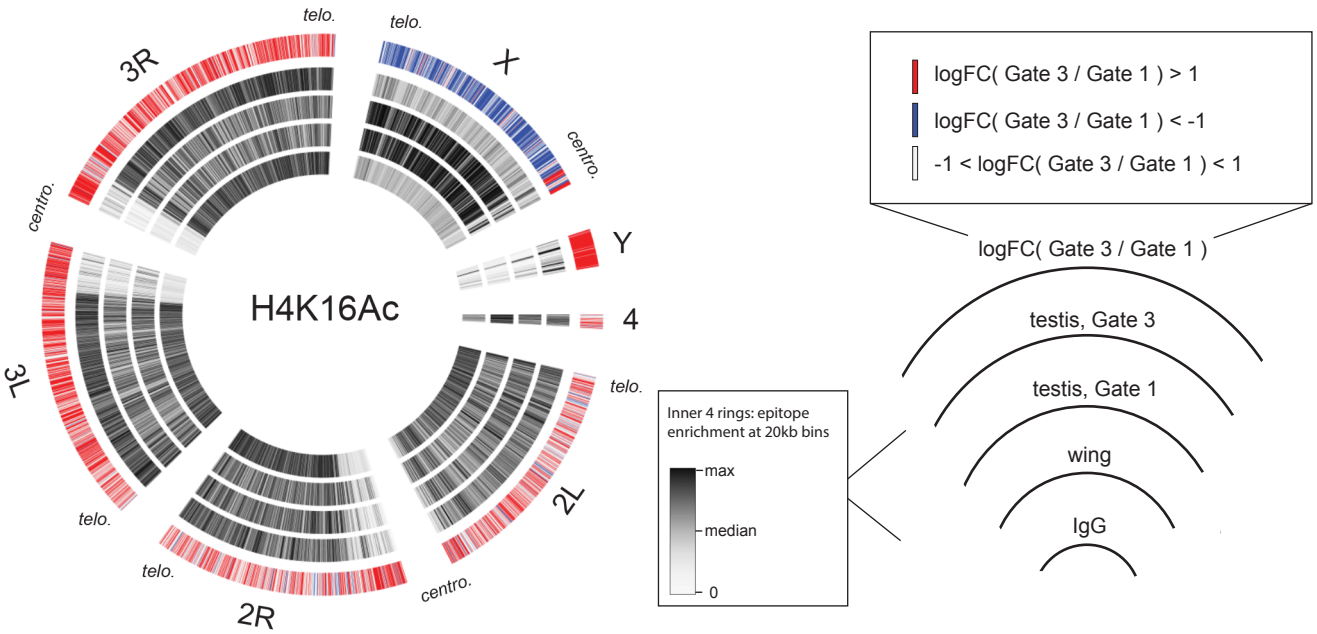


Figure 6. The *X* chromosome lacks dosage compensation in spermatocytes. Circos plot of H4K16Ac. The four innermost rings indicate enrichment of the epitope of interest at all consecutive 20kb genomic bins in: IgG-negative control (innermost ring), *heph-GFP* wing disc (next ring up), FACS Gate 1 from *heph-GFP* testis (next ring up), and FACS Gate 3 from *heph-GFP* testis (next ring up). The black-white scaling is performed independently per ring, such that white indicates enrichment of 0 in any bin, black indicates the maximally enriched bin for that epitope, and gray scaling corresponds with intermediate enrichment. The outermost ring of each circos plots indicates $\log_2$ fold-change of H4K16Ac enrichment from FACS Gate 3 / Gate 1 at all consecutive 20kb genomic bins, color-coded according to logFC values (red, logFC>1; blue, logFC<-1; gray, logFC >-1 and <1). Thus, red indicate bins where the epitope is enriched in spermatocytes relative to all other cell types of the testis, and blue indicates bins where H4K16Ac is depleted in spermatocytes relative to all other cell types of the testis.

sively silenced as the male germline differentiates [24]. Two of the repressive chromatin features associated with mammalian MSCI are H3K9me2 and uH2A, which accumulate on the unsynapsed sex chromosomes and mediate repression. It has been debated whether Drosophila spermatocytes also silence their unpaired sex chromosomes [25]. We report that neither H3K9me2 nor uH2A histone marks become enriched on the *X* chromosome in the Drosophila male germline. Additionally, the elongating form of RNA polymerase II is not reduced at *X*-linked genes compared to autosomal genes in spermatocytes. Thus, there is no evidence for transcriptional silencing of the sex chromosomes in Drosophila. In mammals, meiotic pairing of homologs prevents chromosomal silencing can occur at any unpaired chromosome [26], but only the sex chromosomes are typically unpaired. However, meiosis in Drosophila is unusual in that all chromosomes do not synapse [27]. The evolutionary loss of synapsis in Drosophila males must then have been accompanied by the loss of a silencing response to unpaired chromosomes. The epigenomic status of the germline *X* chromosome does differ from that in somatic cells however, in that canonical dosage compensation is absent in the male germline of Drosophila [19] [18]. Our profiling results further support the idea that dosage compensation is absent from the Drosophila male germline.

Activation of the Y chromosome

The uH2A histone modification is typically associated with Polycomb-mediated gene silencing [14]. We observe that this modification forms a striking chromatin body in spermatocytes, but this body includes the *Y* chromosome that is highly transcribed in the male germline. The uH2A mark also accumulates on transposons that are transcribed in the male germline. While silencing ubiquitylation of histone H2A is catalyzed by the Sce/RING1B enzyme subunit of the Polycomb Repressive Complex 1 (PRC1) in somatic cells [28], the Polycomb subunit of PRC1 does not co-localize with the uH2A body in spermatocytes, suggesting that the germline modification is not catalyzed by Sce. Indeed, the Drosophila genome encodes many other ubiquitinylating enzymes, some of which target the H2A histone [29] [30]. The activated *Y* chromosome also becomes enriched with the H4K16Ac histone modification in spermatocytes, raising the possibility that the acetylation of H4K16 and ubiquitylation of H2A are involved in stage-specific activation of Y-linked genes in spermatocytes. Indeed, previous reports have demonstrated the association of uH2A with certain active genes in other cell types [31].

Methods

Fly strains

All crosses were performed at 25˚C. All mutations and chromosomal rearrangements used here are described in Flybase (http://www.flybase.org) and are listed in Supplementary Table 2. The w^{1118} strain was used as a wildtype control. The *heph-GFP* males have the genotype *y w / Y ; P{PTT-GC}heph*^{CC00664}*/TM3, Ser Sb*. The *bam* mutant refers to males with genotype *w / Y ; e bam*^{D86} */ Df(3R)FDD-0089346*. The *aly* mutant refers to males with genotype *P{ry11}ry2, mwh aly ry*^{506} *e / Df(3L)BSC428*.

Imaging whole testes

Testes from one-day old adult males were dissected and fixed for 10 minutes in 4% formaldehyde/PBST (Phosphate-buffered saline (PBS) with 0.1% Triton-X100), and then incubated twice in 0.3% sodium deoxycholate/PBST for 20' each. Samples were incubated with primary antiserum diluted in PBST supplemented with 10% goat serum for 20', and then incubated with primary antibody at 4˚ overnight, and finally with fluorescently labeled secondary antibodies (1:200 dilution, Jackson ImmunoResearch). All tissues were stained with 0.5 µg/mL DAPI/PBS, mounted in 80% glycerol on slides, and imaged by epifluorescence on an EVOS FL Auto 2 inverted microscope (Thermo Fisher Scientific) with a 10X objective. All antibodies used are listed in Supplementary Table 3.

Imaging spermatocytes

Testes from third-instar male larvae were prepared as described by Bonaccorsi et al. (Bonaccorsi et al 2000). Briefly, a testis was dissected in a drop of PBS on a Histobond glass slide (VWR), then squashed gently with a RainX-coated coverslip (ITW Global Brands), then flash-frozen in liquid nitrogen. After popping off the coverslip, the sample was fixed with 4% formaldehyde/PBST, and then incubated with 0.3% sodium deoxycholate/PBST for 20'. Samples were then incubated with primary antiserum/PBST supplemented with 0.1% bovine serum albumin (BSA) at 4˚ overnight, and finally with fluorescently labeled secondary antibodies (1:200 dilution, Jackson ImmunoResearch). Slides were stained with 0.5 µg/mL DAPI/PBS, mounted in 80% glycerol, and imaged by epifluorescence on an EVOS FL Auto 2 inverted microscope (Thermo Fisher Scientific) with a 40X objective. Pseudo-colored images were adjusted and composited in Adobe Photoshop and Adobe Illustrator. All antibodies used are listed in Supplementary Table 3.

Whole-mount CUT&Tag

To adapt CUT&Tag [6] for whole tissues ("whole-mount CUT&Tag"), we dissected one-day-old adult testes, or imaginal wing discs from 3rd instar larvae, in PBS buffer with Roche cOmplete protease inhibitor. We used 6-10 testes and 6-10 wing imaginal discs for each chromatin profiling experiment. Experiments were performed in replicate (with 2-5 replicates per sample). We then permeabilized whole dissected tissues with 0.1% Triton/PBS for 30 minutes at room temperature, and then exchanged CUT&Tag solutions by manually transferring the tissues to sequential wells of a dissection plate, each bearing 150 µL of each respective CUT&Tag solution. Tissues were incubated with primary antibody in Wash+ buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 0.5 mM spermidine, 2 mM EDTA, 1% BSA, with Roche cOmplete protease inhibitor) overnight at 4˚, incubated with secondary antibody in Wash+ buffer for 1 hour at room temperature, and then incubated with protein-A-Tn5 loaded with adapters in 300Wash + buffer (20 mM HEPES pH 7.5, 300 mM NaCl, 0.5 mM spermidine with Roche cOmplete protease inhibitor) for 1 hour. After one wash with 300Wash+ buffer, samples were incubated in 300Wash+ buffer supplemented with 10 mM MgCl2 for 1 hour at 37˚ to tagment chromatin. To then dissociate the tissues, we transferred tissues to collagenase (2 mg/mL) (Sigma, C9407) in "HEPESCA" solution (50 mM HEPES buffer pH 7.5, 0.36 mM $CaCl_2$) and incubated at 37˚ for 1 hour. To further dissociate the tissue and break-up cell contents, we then added SDS to 0.16% and protease K to 0.3 mg/mL and incubated at 58˚ for 1 hour. DNA was purified by phenol:chloroform extraction and ethanol precipitation. Libraries were prepared as described (protocol: https://www.protocols.io/view/bench-top-cut-amp-tag-kqdg34qdpl25/v3), with 14 cycles of PCR with 10 second combined annealing and extension for enrichment of short DNA fragments. Libraries were sequenced for 50 cycles in paired-end mode (PE50) on the Illumina HiSeq 2500 platform at the Fred Hutchinson Cancer Center Genomics Shared Resource. Paired-end reads were mapped to release r6.30 of the *D. melanogaster* genome obtained from FlyBase using Bowtie2 (using parameters, e.g.: --end-to-end --very-sensitive --no-mixed --no-discordant -q --phred33 -I 10 -X 700), with repetitive sequences having been masked-out of the genome. This repeat-masked genome can be found at: http://hgdownload.cse.ucsc.edu/goldenPath/dm6/bigZips/dm6.fa.masked.gz

FACS-CUT&Tag of spermatocytes

To prepare samples for FACS, 40 testes were dissected from one-day old adult *heph-GFP* males. These 40 dissected testes were transferred into a 200 µL dissection well containing 2 mg/mL collagenase (Sigma, C9407) in 1x "HEPESCA" solution (50mM HEPES buffer pH 7.5, 0.36mM $CaCl_2$), and then this dissection well was left in a 37˚ room for one hour. This dissociated sample was then pipetted up-and-down with a P200 pipette tip to further encourage dissociation. Next, the dissociated material was passed through a 35 µM filter with 5 mL collection tube (Corning, 352235), and 1x PBS was passed through the filter afterwards to bring the total volume of collected filtrate to 1 mL. This collection of dissociated, filtered testes was kept on ice until performing FACS. A MA900 Multi-Application Cell Sorter (Sony) was used for sorting cell populations from dissociated *heph-GFP* testis with a 100 µM nozzle and flow pressure of 2. Identification of distinct cell populations was based on forward scatter (FSC)-vs-GFP plots. To enable visualization of all recorded events on the FSC-GFP plot, the default laser settings of the Sony MA900 were adjusted to FITC/GFP=32.0% and FSC=1. Each sorted population was collected in 1 mL of PBS in 5 mL tubes. Standard benchtop CUT&Tag was performed on these samples as described [6] https://www.protocols.io/view/bench-top-cut-amp-tag-kqdg34qdpl25/v3

Identifying testis-specific promoters

We used EPDnew for promoter locations in dm6 genome assembly [32] [33]. Analysis and display were done using deepTools v3.3.2, degust v. 4.1.1 (degust.erc.monash.edu), and R Studio (R Studio version 1.3.1056, R version 4.1.1). Promoter signals from whole-mount CUT&Tag for the H3K4me2 histone modification in *bam* testis, *aly* testis, w^{1118} testis, and wing imaginal disc were extracted as fragment counts $\pm$500 bp around annotated gene Transcriptional Start Sites (TSSs) using deepTools/multiBamSummary command to generate a count table. Testis-specific promoters were defined using degust v. 4.1.1 (degust.erc.monash.edu) to normalize counts with limma-voom and display comparisons as MA plots summarizing H3K4me2 enrichment at all promoters, in all pairwise comparisons, setting the threshold for

significantly-differing promoters at FDR<0.0001. Promoter names, fold-changes, and FDRs are listed in Supplementary Table 1. Any promoter called testis-specific in at least one comparison with wing was considered a testis-specific promoter. This process produced the 1715 testis-specific promoters (Supplementary Table).

To further sub-categorize these 1715 promoters, promoters called testis-specific only in the *bam*-vs wing MA plot are considered early-germline promoters. Promoters called testis-specific only in the wild-type-testis-vs-wing MA plot are considered late-germline promoters. Testis-specific promoters not classified as early-germline or late-germline are considered "other testis-specific" promoters.

Promoter heatmaps
For promoter heatmaps, profiling coverage from deepTools/bamCoverage was binned by 10 bp, ±2 kb around annotated TSSs, ordered by mean signal and plotted using deepTools/computeMatrix and plotHeatmap commands.

Promoter scores
We assigned promoter activity scores by binning H3K4me2 signal ±500 bp around each of our 1715 testis-specific promoters in *bam*, *aly*, and wild-type testis, as well as wing disc. We calculated this enrichment-per-promoter using deepTools/multiBamSummary command, with parameters:

BED-file –BED (see *bed, below) --outRawCounts

Regarding *bed: This .bed file lists all annotated promoters in the dm6 genome from the EPDnew database. This list of all fly promoters can be found at: https://epd.epfl.ch//index.php.

Activity scores were transformed into z-scores per promoter, and the heatmap was ordered by hierarchical clustering of these transformed promoter scores. The visualization and clustering of this heatmap was performed with the pheatmap package in R, using standard settings.

For the leftmost column indicating marker gene annotations, we used scRNA-seq annotations [5]. The "testis soma" marker gene category is a merged list of genes annotated for the somatic P, T, C1, C2, C3, or C4 cell types.

Gene body scores and scaled boxplots
We calculated gene body enrichment scores for the RNAPolII-Iser2phos FACS-CUT&Tag dataset by summing RNAPolII-ser2phos enrichment across each gene body, using deepTools/multiBamSummary command, with parameters:

BED-file –BED {see *bed, below} –outRawCounts

Regarding *bed: The .bed file used for this multiBamSummary lists the TSS to TES of all annotated genes in the dm6 genome, and was derived from the "Batch Download" tool on Flybase.org.

We then plotted the distributions of these gene-enrichment scores as boxplots, discarding genes with enrichment score = 0 in either wing, Gate 1, or Gate 3. The gene body enrichment score for every gene was scaled by dividing each score by the median of the scores in the "all other genes" category.

Scaled boxplots of histone modifications
We calculated enrichment at all consecutive 20 kb genomic bins for the uH2A, H3K9me2, H3K27me3, and H4K16Ac histone modifications profiled in *heph-GFP* dissociated and sorted testis by FACS-CUT&Tag.

To calculate enrichment across all genomic bins, we used deepTools/multiBamSummary command, with parameters:

bins -bs 20000 –outRawCounts

We then plotted the distributions of these bin-enrichment scores as boxplots representing wing, Gate 1, and Gate 3 from the FACS-isolated populations. We discarded any gene from consideration if its enrichment score = 0 in either wing, Gate 1, or Gate 3. The bin-enrichment score for every bin was scaled to the "autosomes" category, by dividing each score by the median of the scores in the "autosomes" category.

Circos plots
We visualized CUT&Tag data for uH2A, H3K9me3, H3K27me3 (Figures 4B-D), and H4K16Ac (Figure 7) as Circos plots, using the circlize package in R with default settings (circlize version 0.4.15). The inner ring of each Circos plot represents a genome coverage file for the epitope of interest from whole-mount CUT&Tag on wild-type testis, with all values log-transformed for ease of visualization. These genome coverage files were generated by deepTools/bamCoverage command, using parameters:

-bs 20000 --centerReads --effectiveGenomeSize 142573017 -of bedgraph

The outer ring of each Circos plot represents consecutive 20 kb bins in the genome, with the color of each bin (red, blue, or gray) corresponding with the results of a Gate3/Gate1 enrichment analysis, which was performed by calculating the enrichment of each epitope at all consecutive 20kb genomic bins using deepTools/multiBigwigSummary command, with parameters:

bins -bs 20000 –outRawCounts

and then calculating the $\log_2$ fold-change of Gate 3/Gate 1 signals at each genomic bin.

The innermost 4 rings of each Circos plot represent genomic coverage. The color-coding of bins was set independently per ring.

Transposon heatmaps
For all CUT&Tag datasets shown in Figure 6, we assembled a modified version of the repeat-masked dm6 genome, appending the consensus sequences for 128 transposons, listed at: https://github.com/bergmanlab/drosophila-transposons/blob/master/misc/D_mel_transposon_sequence_set.fa [34] [35]. This dm6+tp genome was made by concatenating the transposon sequences to the end of the repeat-masked dm6 genome.

To calculate enrichment of each epitope across each transposons, we used deepTool's multiBamSummary command, with parameters:

BED-file –BED {see *bed, below} –outRawCounts

Regarding *bed: The .bed file used for this multiBamSummary lists the beginning to end of all 128 transposons according to our new "dm6 masked plus transposons" genome. The resulting count table was transformed into counts-per-million

(CPM), and then these CPM values were transformed into z-scores independently per transposon. We removed transposons with all enrichment counts = 0, and heatmaps were made using the pheatmap package in R, with all the default parameters.

Genome Browser Tracks
Track screenshots were produced using the UCSC Genome browser (http://genome.ucsc.edu).

References

1. Fingerhut, J.M., J.V. Moran, and Y.M. Yamashita, *Satellite DNA-containing gigantic introns in a unique gene expression program during Drosophila spermatogenesis.* PLoS Genet, 2019. **15**(5): p. e1008028.

2. Brown, E.J., A.H. Nguyen, and D. Bachtrog, *The Drosophila Y Chromosome Affects Heterochromatin Integrity Genome-Wide.* Mol Biol Evol, 2020. **37**(10): p. 2808-2824.

3. Zhang, J., et al., *The Role of Y Chromosome Genes in Male Fertility in Drosophila melanogaster.* Genetics, 2020. **215**(3): p. 623-633.

4. Kemkemer, C., W. Hense, and J. Parsch, *Fine-scale analysis of X chromosome inactivation in the male germ line of Drosophila melanogaster.* Mol Biol Evol, 2011. **28**(5): p. 1561-3.

5. Mahadevaraju, S., et al., *Dynamic sex chromosome expression in Drosophila male germ cells.* Nat Commun, 2021. **12**(1): p. 892.

6. Kaya-Okur, H.S., et al., *CUT&Tag for efficient epigenomic profiling of small samples and single cells.* Nat Commun, 2019. **10**(1): p. 1930.

7. White-Cooper, H., *Molecular mechanisms of gene regulation during Drosophila spermatogenesis.* Reproduction, 2010. **139**(1): p. 11-21.

8. Bernstein, B.E., et al., *Genomic maps and comparative analysis of histone modifications in human and mouse.* Cell, 2005. **120**(2): p. 169-81.

9. Lu, D., et al., *Developmental regulation of cell type-specific transcription by novel promoter-proximal sequence elements.* Genes Dev, 2020. **34**(9-10): p. 663-677.

10. Chen, X., et al., *Sequential changes at differentiation gene promoters as they become active in a stem cell lineage.* Development, 2011. **138**(12): p. 2441-50.

11. Laktionov, P.P., et al., *Genome-wide analysis of gene regulation mechanisms during Drosophila spermatogenesis.* Epigenetics Chromatin, 2018. **11**(1): p. 14.

12. Dietzel, S., et al., *The nuclear distribution of Polycomb during Drosophila melanogaster development shown with a GFP fusion protein.* Chromosoma, 1999. **108**(2): p. 83-94.

13. Turner, J.M., *Meiotic Silencing in Mammals.* Annu Rev Genet, 2015. **49**: p. 395-412.

14. Baarends, W.M., et al., *Histone ubiquitination and chromatin remodeling in mouse spermatogenesis.* Dev Biol, 1999. **207**(2): p. 322-33.

15. Khalil, A.M., F.Z. Boyar, and D.J. Driscoll, *Dynamic histone modifications mark sex chromosome inactivation and reactivation during mammalian spermatogenesis.* Proceedings of the National Academy of Sciences, 2004. **101**(47): p. 16583-16587.

16. Aloia, L., B. Di Stefano, and L. Di Croce, *Polycomb complexes in stem cells and embryonic development.* Development, 2013. **140**(12): p. 2525-34.

17. Kelley, R.L., et al., *Sex lethal controls dosage compensation in Drosophila by a non-splicing mechanism.* Nature, 1997. **387**(6629): p. 195-9.

18. Meiklejohn, C.D., et al., *Sex chromosome-specific regulation in the Drosophila male germline but little evidence for chromosomal dosage compensation or meiotic inactivation.* PLoS Biol, 2011. **9**(8): p. e1001126.

19. Rastelli, L. and M.I. Kuroda, *An analysis of maleless and histone H4 acetylation in Drosophila melanogaster spermatogenesis.* Mech Dev, 1998. **71**(1-2): p. 107-17.

20. White-Cooper, H., et al., *Molecular mechanisms of gene regulation during Drosophila spermatogenesis.* Reproduction, 2010. **139**(1): p. 11-21.

21. Lawlor, M.A., W. Cao, and C.E. Ellison, *A transposon expression burst accompanies the activation of Y-chromosome fertility genes during Drosophila spermatogenesis.* Nat Commun, 2021. **12**(1): p. 6854.

22. Benayoun, B.A., et al., *H3K4me3 breadth is linked to cell identity and transcriptional consistency.* Cell, 2014. **158**(3): p. 673-88.

23. Park, S., et al., *Broad domains of histone H3 lysine 4 trimethylation in transcriptional regulation and disease.* Febs j, 2020. **287**(14): p. 2891-2902.

24. Turner, J.M., *Meiotic sex chromosome inactivation.* Development, 2007. **134**(10): p. 1823-31.

25. Vibranovski, M.D., *Meiotic sex chromosome inactivation in Drosophila.* J Genomics, 2014. **2**: p. 104-17.

26. Burgoyne, P.S., S.K. Mahadevaiah, and J.M. Turner, *The consequences of asynapsis for mammalian meiosis.* Nat Rev Genet, 2009. **10**(3): p. 207-16.

27. McKee, B.D. and M.A. Handel, *Sex chromosomes, recombination, and chromatin conformation.* Chromosoma, 1993. **102**(2): p. 71-80.

28. Blackledge, N.P., et al., *PRC1 Catalytic Activity Is Central to Polycomb System Function.* Mol Cell, 2020. **77**(4): p. 857-874.e9.

29. Gorfinkiel, N., et al., *The Drosophila Polycomb group gene Sex combs extra encodes the ortholog of mammalian Ring1 proteins.* Mech Dev, 2004. **121**(5): p. 449-62.

30. Tasaki, T., et al., *A family of mammalian E3 ubiquitin ligases that contain the UBR box motif and recognize N-de-*

grons. Mol Cell Biol, 2005. **25**(16): p. 7120-36.

31. Loubiere, V., et al., *Widespread activation of developmental gene expression characterized by PRC1-dependent chromatin looping.* Sci Adv, 2020. **6**(2): p. eaax4001.

32. Dreos, R., et al., *The Eukaryotic Promoter Database: expansion of EPDnew and new promoter analysis tools.* Nucleic Acids Res, 2015. **43**(Database issue): p. D92-6.

33. Meylan, P., et al., *EPD in 2020: enhanced data visualization and extension to ncRNA promoters.* Nucleic Acids Res, 2020. **48**(D1): p. D65-d69.

34. Sackton, T.B., et al., *Population genomic inferences from sparse high-throughput sequencing of two populations of Drosophila melanogaster.* Genome Biol Evol, 2009. **1**: p. 449-65.

35. Méndez-Lago, M., et al., *A large palindrome with interchromosomal gene duplications in the pericentromeric region of the D. melanogaster Y chromosome.* Mol Biol Evol, 2011. **28**(7): p. 1967-71.

Supplementary Table Legends

Supplementary Table 1. H3K4me2 at promoters, including promoter-annotations based on CUT&Tag profiling.
Log of all promoters referred to throughout the H3K4me2 analysis presented in Figure 1.

Supplementary Table 2. Fly genotypes.
All fly genotypes used throughout this study.

Supplementary Table 3. Antibodies.
All antibodies used throughout this study.

Supplementary Table 4. CUT&Tag samples.
Log of all testis or wing samples which were profiled by whole-mount CUT&Tag or by standard CUT&Tag throughout this study. The standard CUT&Tag protocol is described here: https://www.protocols.io/view/bench-top-cut-amp-tag-kqdg34qdpl25/v3.

Supplementary Table 5. All FACS-CUT&Tag samples.
Log of all FACS-CUT&Tag samples in this paper.

Supplementary Table 6. RNAPolIISer2phos enrichment at gene bodies.
Reads as CPM of RNAPolIIser2phos enrichment at all genes, as summarized by the boxplot in Figure 2G.

Supplementary Table 7. Transposon from Bergman Lab.
A list of 128 transposons provided by: https://github.com/bergmanlab/drosophila-transposons/blob/master/misc/D_mel_transposon_sequence_set.fa

The leftmost column, ID_full_length, provides the transposon IDs as they are listed on the source website. In the following three columns, we sub-divided these IDs into three parts to facilitate sorting/filtering the transposons.

Supplementary Table 8. uH2A, H3K9me2, H3K27me3, and RNAPolIIser2phos at transposons.
This is the list of 126 transposons shown in the heatmap of Figure 6A, sorted in the same order as the heatmap after having been hierarchically clustered. Counts are in CPM.

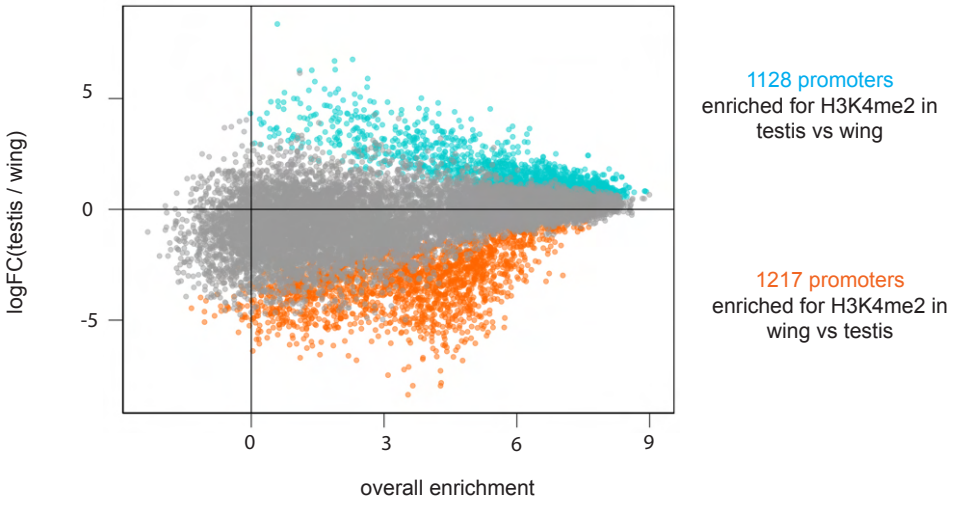
Supplementary Table 9. uH2A at transposons.
This is the list of 126 transposons shown in the heatmap of Figure 6B, sorted in the same order as the heatmap after having been hierarchically clustered. Counts are in CPM.

Supplementary Table 10. H3K9me2 at transposons.
This is the list of 126 transposons shown in the heatmap of Figure 6C, sorted in the same order as the heatmap after having been hierarchically clustered. Counts are in CPM.

Supplementary Table 11. H3K27me3 at transposons.
This is the list of 125 transposons shown in the heatmap of Figure 6D, sorted in the same order as the heatmap after having been hierarchically clustered. Counts are in CPM.

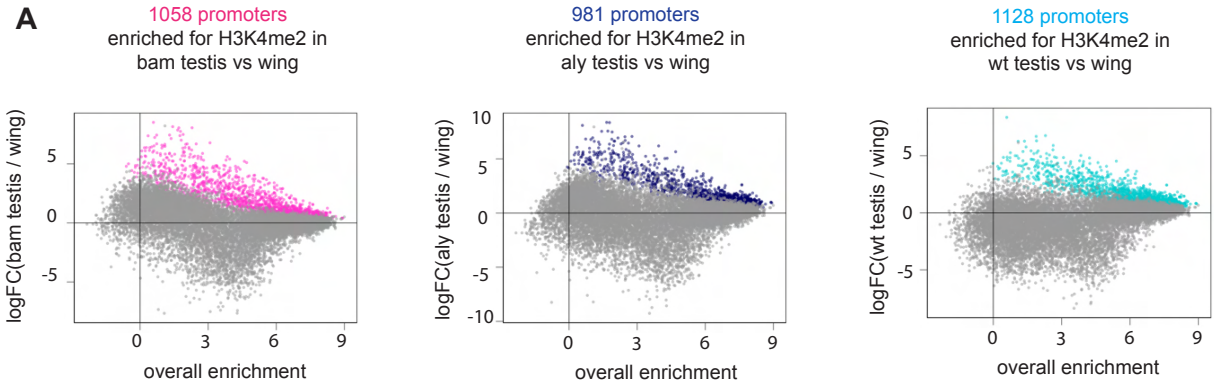
Supplementary Table 12. RNAPolIISer2phos at transposons.
This is the list of 123 transposons shown in the heatmap of Figure 6E, sorted in the same order as the heatmap after having been hierarchically clustered. Counts are in CPM.

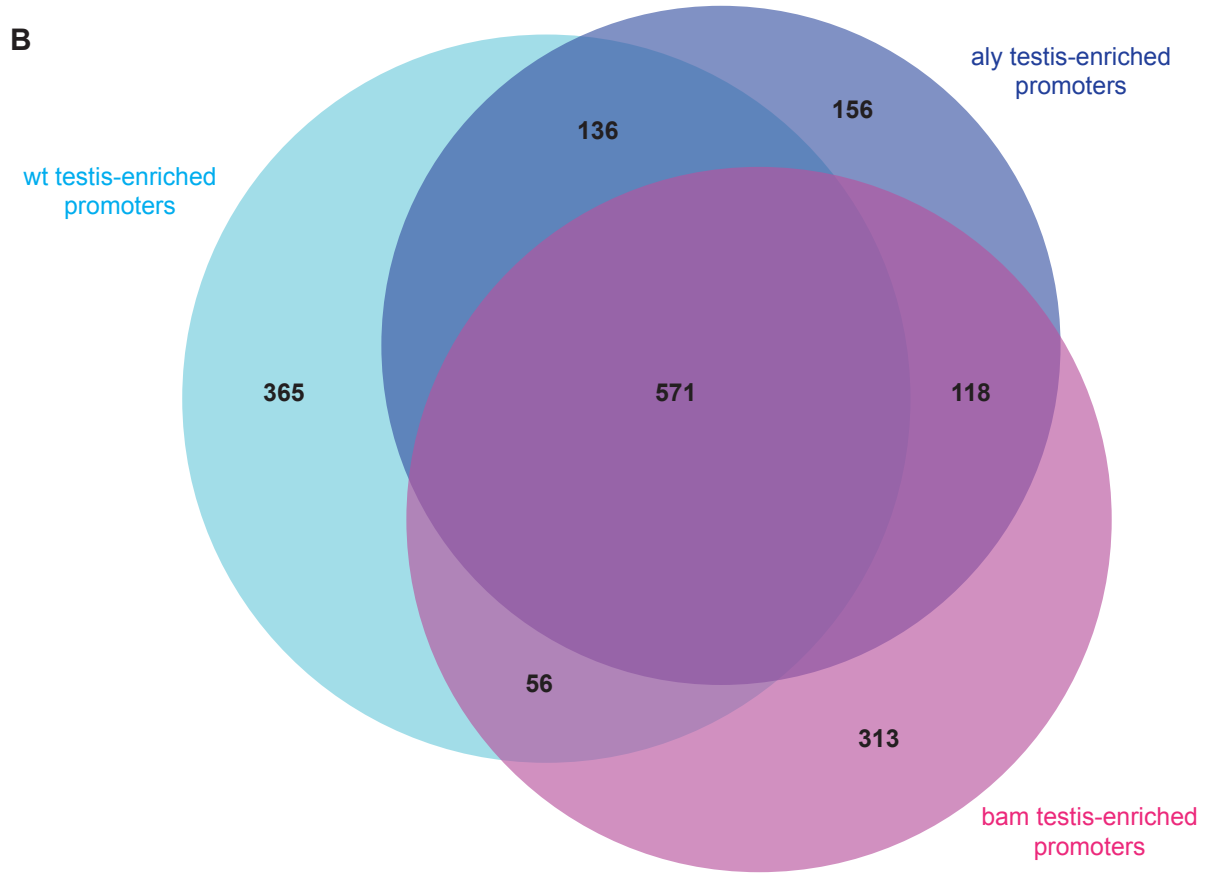
Supplementary Figure S1



Supplementary Figure S1. MA plots for the identification of testis- or wing-specific active promoters based on H3K4me2 CUT&Tag profiles. MA plot of H3K4me2 enrichment at all promoters in testis-vs-wing. In this MA plot, the y-axis indicates log2 fold-change of H3K4me2 enrichment at each promoter in testis / wing. Thus, the y-axis illustrates overall trends of enrichment for H3K4me2 in testis vs soma, with points above y = 0 indicating promoters which are enriched for H3K4me2 in testis vs wing, and points below y = 0 indicating promoters which are enriched for H3K4me2 in wing vs testis. The x-axis indicates overall enrichment of H3K4me2 in general, regardless of logFC. The H3K4me2 enrichment counts underlying this plot are given in Supplementary Table 1. We have color-coded the points to indicate whether a promoter was deemed significantly enriched in testis (cyan), significantly enriched in wing (orange), or not significantly enriched in one tissue vs another (gray) (FDR<0.0001).

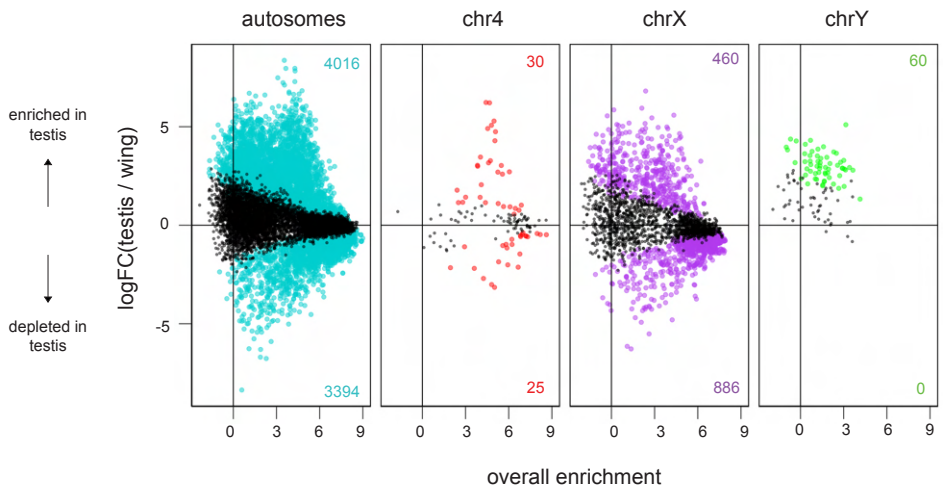
Supplementary Figure S2





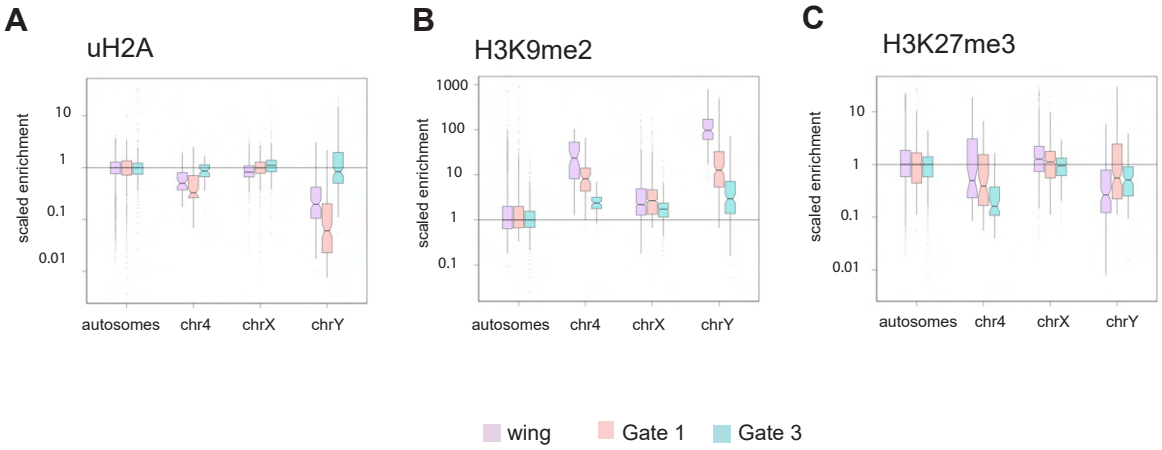
Supplementary Figure S2. MA plots for the identification of early-germline-, late-germline-, or "other testis-specific" promoters based on H3K4me2 CUT&Tag profiles. (A) See Supplementary Figure S1A for an explanation of what an MA plot represents. The H4K4me2 enrichment counts underlying this plot are given in Supplementary Table 1. We have color-coded the points such that gray indicates promoters which were deemed not significantly enriched in one tissue vs another, and any coloring (purple, navy, or cyan) represents promoters which were deemed significantly enriched (FDR < 0.0001) in bam testis-, aly testis-, or wild-type testis-vs-wing.
(B) The MA plots in Supplementary Figure S2A identified 1058, 981, and 1128 promoters respectively which are enriched in testis-vs-wing according to bam testis- (left), aly testis- (middle), and wild-type testis (right)-vs-wing comparisons. As expected, there's a large overlap between the promoters which were called as testis-specific via these three separate approaches. For example, there are 571 promoters which were called as testis-specific according to all three MA plots in Supplementary Figure S2A. In Figure 1E (and henceforth), we use the term "late-germline promoters" to refer to the 365 promoters which were identified as being testis-specific only in the wild-type-vs-wing MA plot, not in the other MA plots. Also in Figure 1E (and henceforth), we use the term "early-germline promoters" to refer to the 313 promoters which were identified as being testis-specific only in the bam testis-vs-wing MA plot, not in the other MA plots. For all remaining 1037 promoters which were called as testis-specific promoters, we simply use the term "other testis-specific promoters."

Supplementary Figure S3



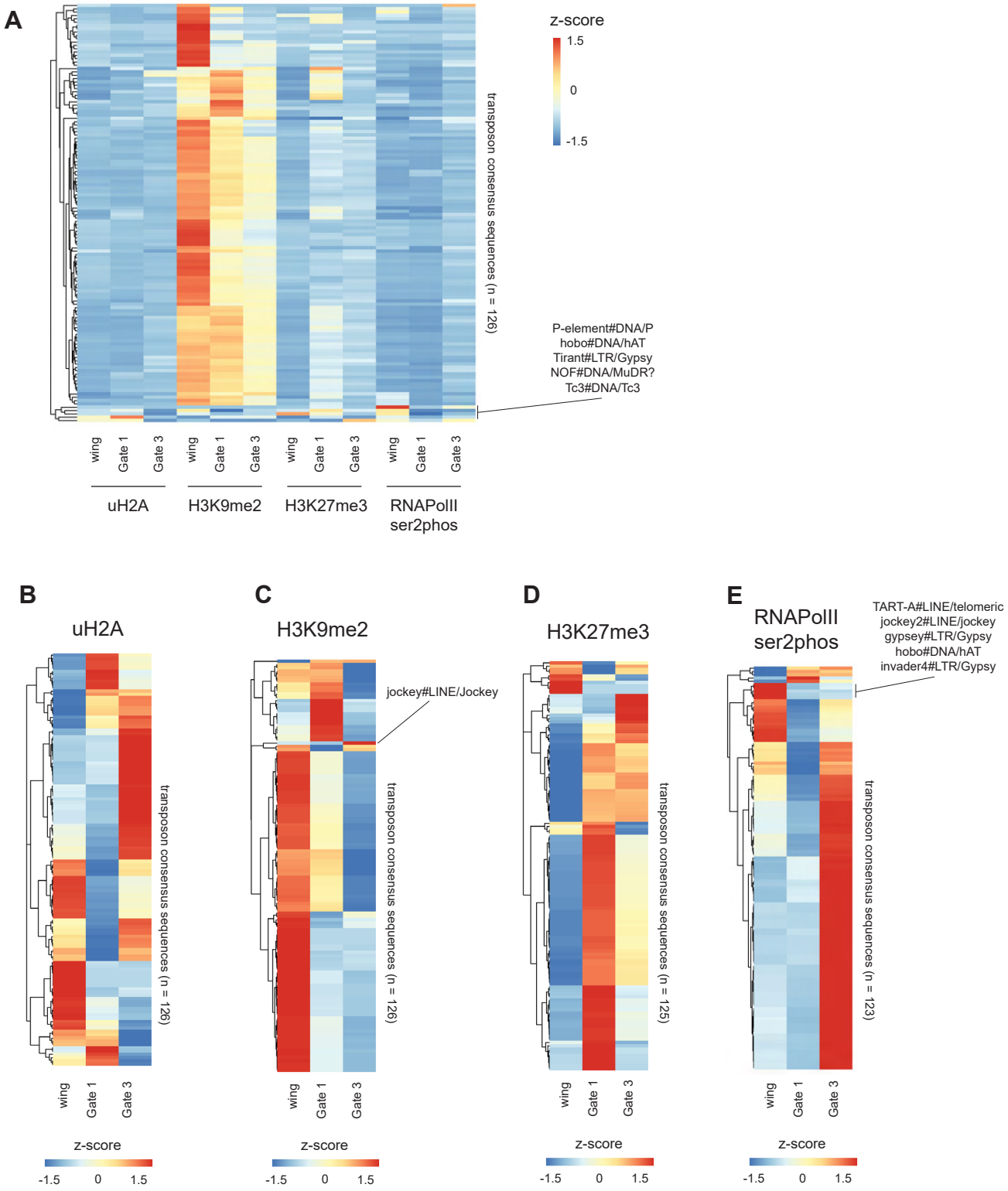
Supplementary Figure S3. MA plots show H3K4me2 in testis vs wing. This is the same plot displayed in Supplementary Figure1A. The only difference is that, here, we have separated all promoters according to their chromosomal location. The H3K4me2 enrichment counts underlying this plot are given in Supplementary Table 1.
We have color-coded the points such that small black points indicate promoters which were deemed not significantly enriched either tissue vs the other (FDR > 0.05), and any colored point (cyan, red, purple, or green for: autosomes, chr4, chrX, or chrY respectively) indicate promoters which were deemed significantly enriched in either testis or wing (FDR < 0.05). The numbers on the right corners of each plot indicate how many promoters were deemed significantly enriched in testis (top right number) or wing (lower right number).

Supplementary Figure S4



Supplementary Figure S4. Boxplots show the differential enrichment of uH2A, H3K9me2, and H3K27me3 across different chromosomes in testis Gate 1, testis Gate 3, and wing. (A) Boxplots showing the distribution of enrichment scores for uH2A enrichment from FACS-CUT&Tag at all consecutive 20kb genomic bins, with all bins separated according to chromosomal location and scaled to the median score of the "autosomes" category.
(B) Boxplots showing the distribution of enrichment scores for H3K9me2 enrichment from FACS-CUT&Tag at all consecutive 20kb genomic bins, with all bins separated according to chromosomal location and scaled to the median score of the "autosomes" category.
(C) Boxplots showing the distribution of enrichment scores for H3K27me3 enrichment from FACS-CUT&Tag at all consecutive 20kb genomic bins, with all bins separated according to chromosomal location and scaled to the median score of the "autosomes" category.

Supplementary Table S5



Supplementary Figure S5. Different epitopes have distinct patterns of enrichment at transposons in wing and spermatocytes. (A) To address whether uH2A, H3K9me2, H3K27me3, or RNAPolIIser2phos may be associated with repression of transposons in spermatocytes, we aligned FACS-CUT&Tag sequencing to a custom repeat-masked dm6 genome with the consensus sequences for 128 families of transposable elements to assess the enrichment of chromatin features on transposons. For each epitope at each transposon family, we calculated counts-per-million (CPM) signal in wing disc, Gate 1, and Gate 3 samples, transformed these counts into z-scores per transposon, and summarized the trends of epitope enrichment as heatmaps. When evaluating uH2A, H3K9me2, H3K27me3, and RNAPolIIser2phos signals across all transposons in a single heatmap (Supplementary Figure S5A), the most striking trend is that the heterochromatic marker H3K9me2 is much more enriched at nearly all transposons in general than any other epitope. Of the 126 transposons shown in Supplementary Figure 5A, only one cluster of 5 transposons do not display this trend of high H3K9me2 enrichment (see the bottom five rows of Supplementary Figure S6A, with bracket pointing toward the transposon names in this cluster). Interestingly, Supplementary Figure 5A also illustrates that one particular transposon, P-element#DNA/P, is highly transcribed in wing (as indicated by the high enrichment of RNAPolIISer2phos) while being relatively depleted of uH2A, H3K9me2, and H3K27me3.

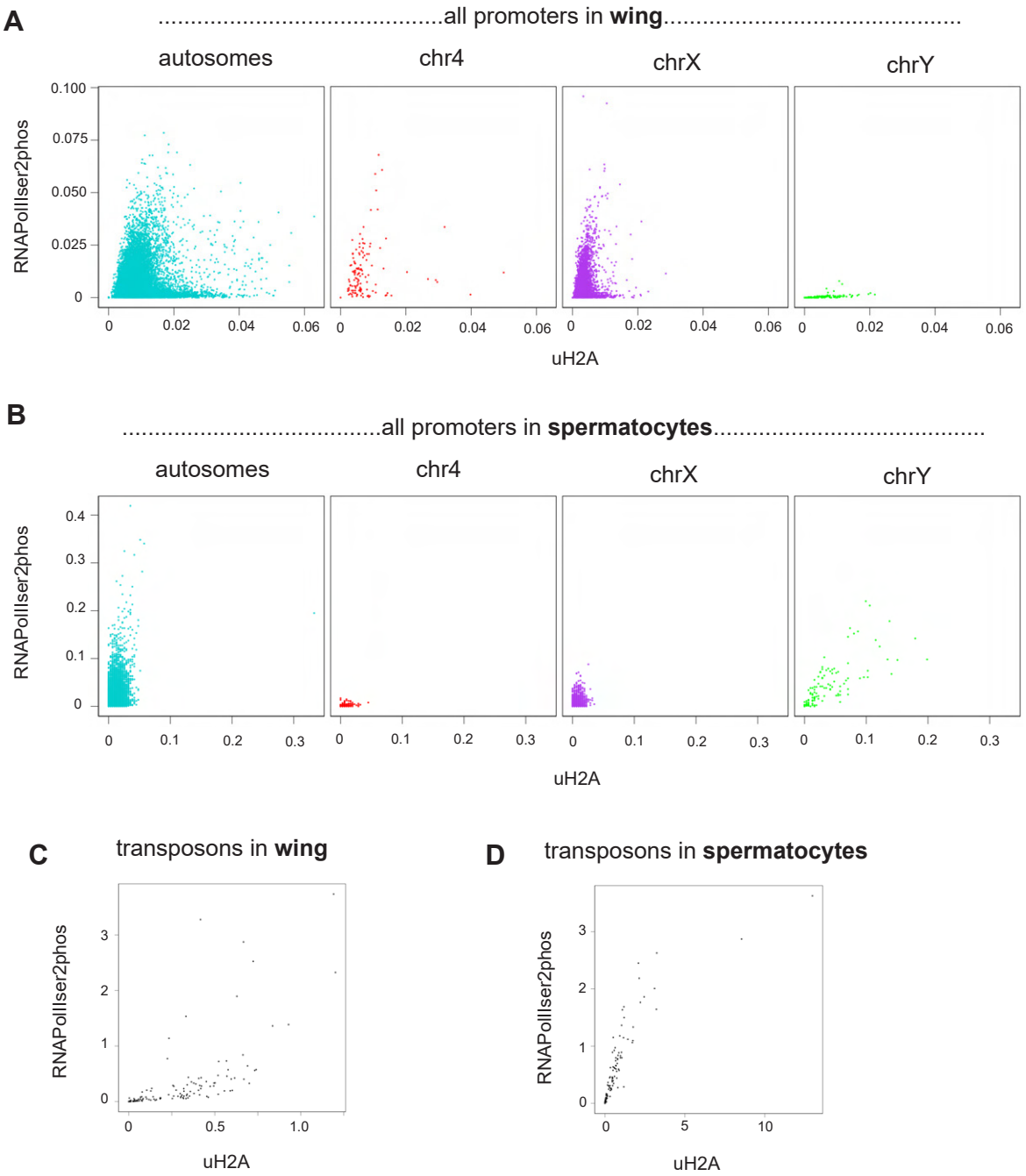
(B-E) To assess more subtle changes in chromatin features, we generated heatmaps for each epitope individually (Supplementary Figures S5B-5E; Supplementary Tables 8-12). These heatmaps reveal distinct trends of enrichment for uH2A, H3K9me2, H3K27me3, and RNAPolIIser2phos in wing and spermatocytes. First, uH2A (Supplementary Figure S5B) is broadly enriched at the majority of all transposons in spermatocytes (i.e. Gate 3) compared to other testis cell types (Gate 1), suggesting that germline cells accumulate uH2A at a wide variety of transposons. In total, 85 out of 126 transposons have higher enrichment of uH2A in Gate 3 than in Gate 1. This parallels the gain of uH2A across heterochromatic regions (Figure 4B).

H3K9me2 (Supplementary Figure S5C) becomes relatively depleted in spermatocytes (Gate 3) compared to other testis cell types (Gate 1) or compared to wing. In particular, 122 out of 126 transposons are depleted of H3K9me2 in Gate 3 compared to wing disc. The one particularly dramatic exception to this trend is the element jockey#LINE/Jockey (indicated with a line in Figure S5C), which accumulates a high amount of H3K9me2 in Gate 3 compared to either Gate 1 or wing.

Transposons in wing disc are generally depleted of H3K27me3 compared to either Gate 1 or Gate 3 (Supplementary Figure S5D). More specifically, only 10 out of 123 transposons (indicated in the uppermost 10 rows of Supplementary Figure 5D) are highly enriched with H3K27me3 in wing. Comparing H3K27me3 enrichment in Gate 1 vs Gate 3, we see that 76 out of 125 transposons (indicated in the lower 76 rows of Figure S5D) are depleted of H3K27me3 in spermatocytes (Gate 3) compared to all other cell types of the testis (Gate 1).

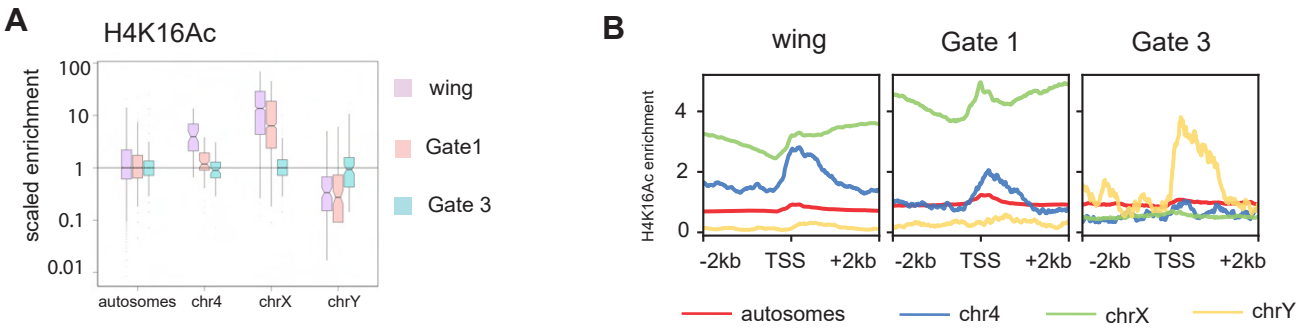
Interestingly, RNAPolIIser2phos (Supplementary Figure S5E) becomes enriched at 95 out of 123 transposons in spermatocytes, indicating that these transposons are transcribed specifically in the male germline. A small fraction of transposons become transcriptionally inactive in spermatocytes, including: TART-A#LINE/telomeric, jockey2#LINE/jockey, gypsey#LTR/Gypsy, hobo#DNA/hAT, and invader4#LTR/Gypsy.

Supplementary Figure S6



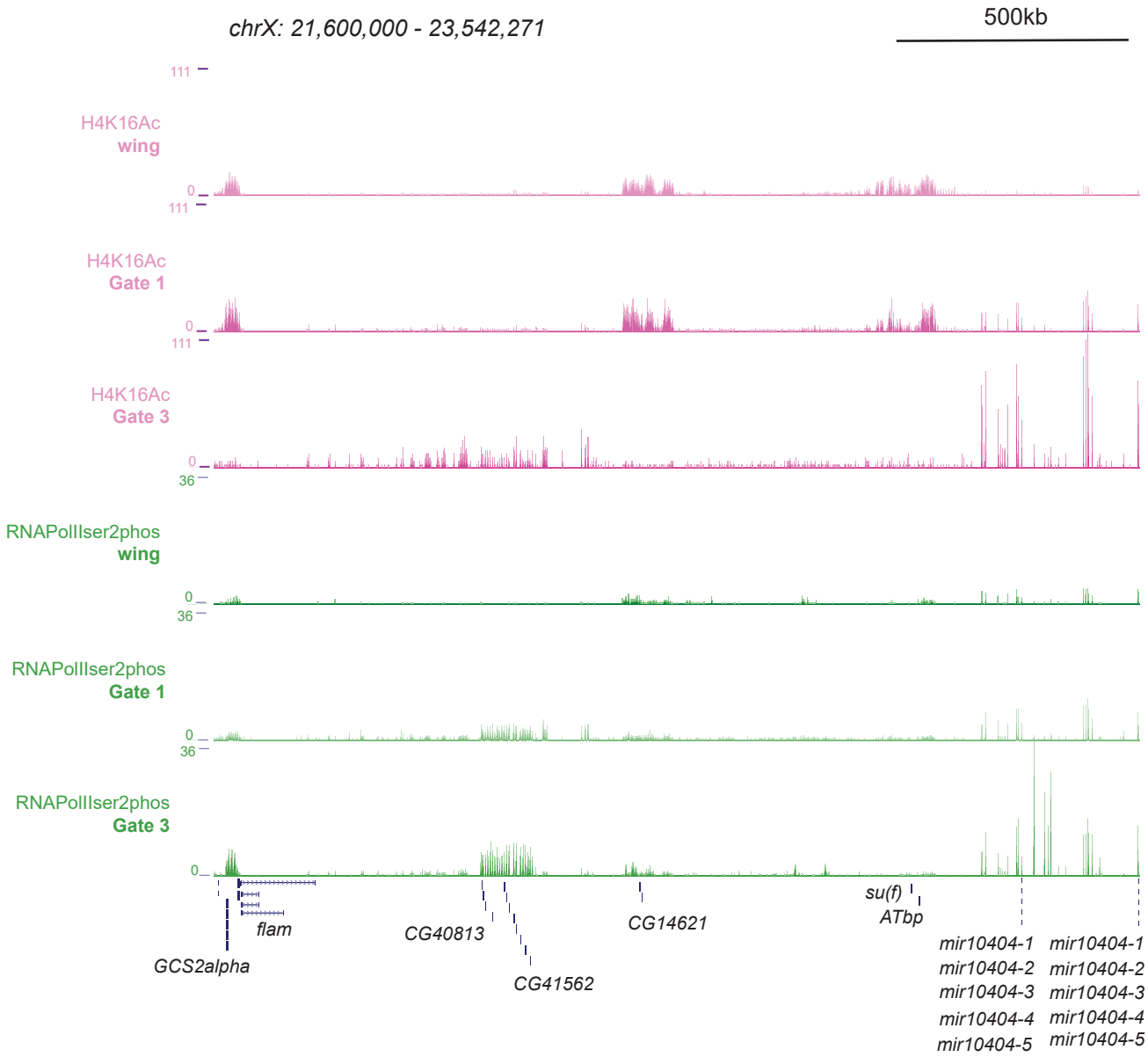
Supplementary Figure S6. Correspondence of uH2A and RNAPolIIser2phos in spermatocytes and wing disc, at promoters or transposons. (A) Scatterplots showing enrichment of uH2A vs RNAPolIIser2phos in wing disc at all 1kb promoter bins genome-wide, separated according to chromosomal location. Axis values are counts per million (CPM).
(B) Scatterplots showing enrichment of uH2A vs RNAPolIIser2phos in sorted spermatocytes (Gate 3) at all 1kb promoter bins genome-wide, separated according to chromosomal location. Axis values are counts per million (CPM).
(C) Scatterplots showing enrichment of uH2A vs RNAPolIIser2phos in wing disc at all transposon consensus sequences. Axis values are counts per million (CPM).
(D) Scatterplots showing enrichment of uH2A vs RNAPolIIser2phos in sorted spermatocytes (Gate 3) at all transposon consensus sequences. Axis values are counts per million (CPM).

Supplementary Figure S7



Supplementary Figure S7. H4K16Ac becomes depleted from chromosome X in sorted spermatocytes. (A) Boxplots showing the distribution of enrichment scores for H4K16Ac enrichment from FACS-CUT&Tag at all consecutive 20kb genomic bins, with all bins separated according to chromosomal location and scaled to the median score of the “autosomes” category.
(B) Profile plots show the enrichment of H4K16Ac around all promoters (the start of the promoter is labeled as “TSS” in the plot). We separated promoters into four categories to determine whether there were chromosomal trends in the changing dynamics of H4K16Ac in Gate 3, the four categories being: autosomes (red), chromosome 4 (blue), chromosome X (green), and chromosome Y (yellow).

Supplementary Figure S8



Supplementary Figure S8. The distal region of the X chromosome has several loci which gain H4K16Ac in sorted spermatocytes. Genome browser tracks showing H4K16Ac and RNAPolIIser2phos profiles in wing, Gate 1, and Gate 3 at the distal region of the X chromosome.