Title: Cell type- and chromosome-specific chromatin landscapes and DNA replication programs of *Drosophila* testicular stem cells

Authors: Jennifer A. Urban[*,2], Daniel Ringwalt[2], John M. Urban[1,3], Wingel Xue[2], Ryan Gleason[2], Keji Zhao[4], Xin Chen[1,2,*]

[1]: Howard Hughes Medical Institute, Baltimore, MD 21218, USA

[2]: Department of Biology, The Johns Hopkins University, 3400 North Charles Street, Baltimore, Baltimore, MD 21218, USA

[3]: Biosphere Sciences and Engineering, Carnegie Institution for Science, Baltimore, MD 21218, USA

[4]: Systems Biology Center, National Heart, Lung and Blood Institute, NIH, 10 Center Dr., MSC 1674, Building 10, Room 7B05, Bethesda, MD 20892, USA

[*] Correspondence to: Jennifer A. Urban, Ph.D, Department of Biology, 3400 North Charles Street, The Johns Hopkins University, Baltimore, MD 21218-2685, USA., Email: jenurbanphd@gmail.com

Xin Chen, Ph.D., Howard Hughes Medical Institute, Department of Biology, 3400 North Charles Street, The Johns Hopkins University, Baltimore, MD 21218-2685, USA. Tel: 410-516-4576; Fax: 410-516-5213; Email: xchen32@jhu.edu

Running title: Chromatin and replication programs in stem cells

Key words: transcriptome, chromatin, replication, stem cells, *Drosophila*, testis

ABSTRACT

Stem cells have the unique ability to self-renew and differentiate into specialized cell types. Epigenetic mechanisms, including histones and their post-translational modifications, play a crucial role in regulating programs integral to a cell's identity, like gene expression and DNA replication. However, the transcriptional, chromatin, and replication timing profiles of adult stem cells *in vivo* remain poorly understood. Containing germline stem cells (GSCs) and somatic cyst stem cells (CySCs), the *Drosophila* testis provides an excellent *in vivo* model for studying adult stem cells. However, the small number of stem cells and cellular heterogeneity of this tissue have limited comprehensive genomic studies.

In this study, we developed cell type-specific genomic techniques to analyze the transcriptome, histone modification patterns, and replication timing of GSC-like and CySC-like cells. Single cell RNA-sequencing validated previous findings on GSC-CySC intercellular communication and revealed high expression of chromatin regulators in GSCs. To characterize chromatin landscapes, we developed a cell-type-specific chromatin profiling assay to map H3K4me3-, H3K27me3-, and H3K9me3-enriched regions, corresponding to euchromatic, facultative heterochromatic, and constitutive heterochromatic domains, respectively. Finally, we determined cell type-specific replication timing profiles, integrating our *in vivo* datasets with published data using cultured cell lines. Our results reveal that GSCs display a distinct replication program compared to somatic lineages, that aligns with chromatin state differences. Collectively, our integrated transcriptomic, chromatin, and replication datasets provide a comprehensive framework for understanding genome regulation differences between these *in vivo* stem cell populations, demonstrating the power of multi-omics in uncovering cell type-specific regulatory features.

INTRODUCTION

Stem cells are defined by their ability to self-renew and differentiate into specialized cell types. During development and tissue homeostasis, asymmetric cell division of stem cells produces two genetically identical daughter cells that adopt distinct cell fates. It is hypothesized that the asymmetric

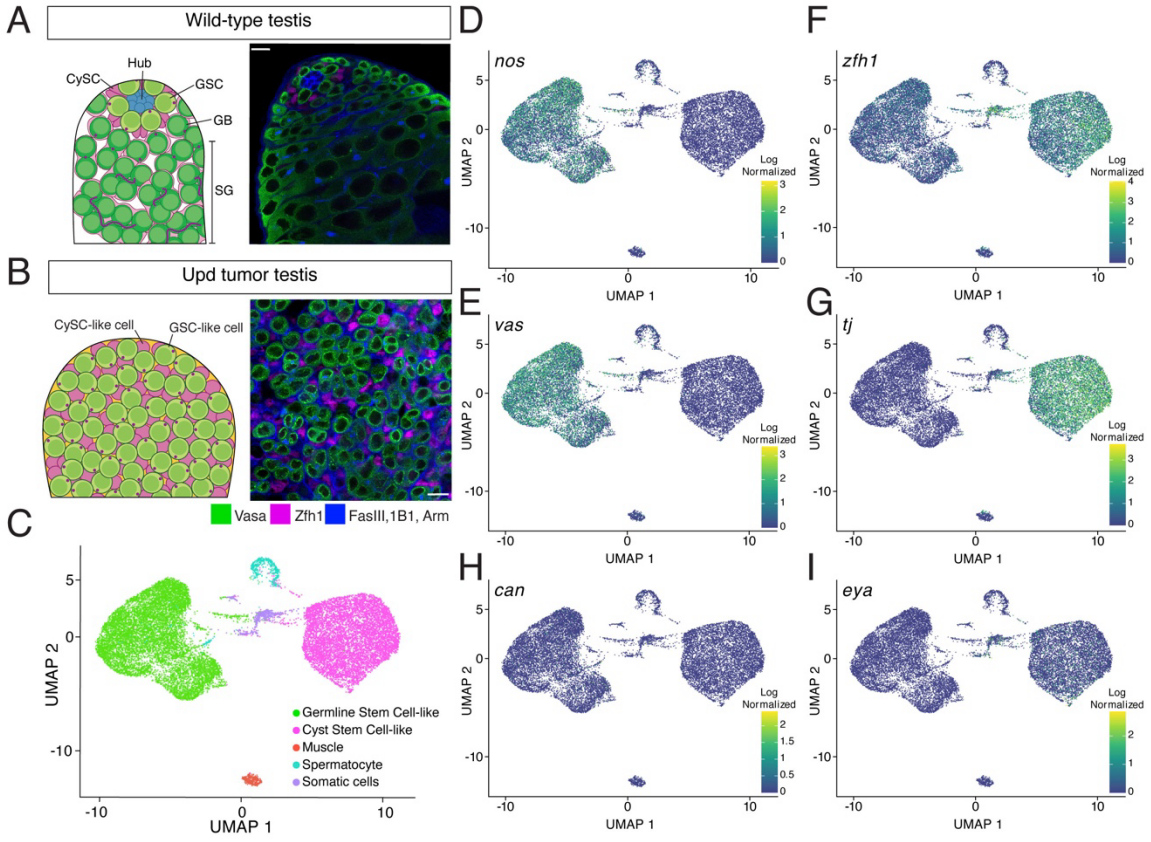
52 distribution and inheritance of cell fate determinants play a critical role in specifying these identities.
53 Epigenetic mechanisms, such as histones and their post-translational modifications, regulate chromatin
54 compaction, transcription, and DNA replication timing, to influence cell fate decisions. In stem cells,
55 epigenetic regulation is thought to be a key contributor to their differentiation potential. However, its
56 precise roles in *in vivo* stem cell systems remain to be fully elucidated.

57 The *Drosophila* male germline is a well-established model for studying stem cell biology with well-
58 defined cell lineages and differentiation programs. Germline stem cells (GSCs) reside near the stem cell
59 niche and undergo asymmetric cell divisions to generate a self-renewing GSC and a differentiating
60 gonialblast (GB, Figure 1A) (Fuller and Spradling 2007; Matunis et al. 2012). Concurrently, cyst stem
61 cells (CySCs) divide to produce a CySC and a cyst cell (CC) (Cheng et al. 2011; Gönczy and Dinardo
62 1996). The GB, encapsulated by two CCs, undergoes four mitotic divisions with incomplete cytokinesis,
63 forming a 16-cell spermatogonial (SG) cyst that subsequently enters meiosis (Figure 1A) (Matunis et al.
64 2012).

65 Germline and cyst cell lineages interact through cell-cell adhesion molecules and signaling pathways,
66 ensuring proper germline cyst differentiation to become mature sperm (Matunis et al. 2012). Notably, the
67 germline differentiation process is partially reversible. For example, partially differentiated SG cells can
68 dedifferentiate, reverting to a GSC-like state (Brawley and Matunis 2004; Sheng et al. 2009). Even
69 though their redifferentiation potency remains unclear and needs further investigation, this ability for
70 differentiated cells to dedifferentiate makes the male germline a powerful system for studying factors that
71 regulate changes in cellular identity.

72 In the *Drosophila* male germline, traditional genetic and cell biological approaches have provided key
73 insights into lineage-specific differences between GSCs and CySCs, as well as how these two lineages
74 interact to ensure normal stem cell functions. However, due to the cell heterogeneity and limited number
75 of stem cells in wild-type testis tissue, comprehensive cell-type-specific genomic datasets defining their
76 global transcriptional, chromatin, and replication landscapes have remained elusive. To address this, we

77 developed cell-type-specific genomic strategies for testicular stem cells, enabling a detailed analysis of
78 their transcriptome, chromatin landscapes, and replication timing dynamics. This approach uncovered
79 previously unrecognized chromosome-specific differences in chromatin regulation between GSCs and
80 CySCs, shedding new light on the molecular mechanisms regulating stem cell identity and activity.



81

82 **Figure 1: Single-Cell RNA-seq of *Upd*-tumor Testes Classifies Two Stem Cell Populations**

83 A: Schematic of a wild-type testis alongside a single z-slice confocal image demonstrating the
84 differentiation process of germline (green) and cyst (magenta) cells. At the apical tip, GSCs and CySCs
85 surround the stem cell niche (FasIII, blue cell cluster). Individual spermatogonial (SG) cysts are identified
86 by the spectrosome/fusome and the cell adhesion protein armadillo (blue).

87 B: A schematic and single z-slice confocal image of the *upd*-tumor. The tumor lacks a stem cell niche and
88 is composed of GSC-like (green) and CySC-like (magenta) cells. In blue, is the spectrosome, a spectrin-
89 containing structure that appears as a dot in GSCs.

90 C: The *upd*-tumor is composed of approximately five cell clusters, with the two largest groups
91 representing GSC-like and CySC-like cells.
92 D-I. Uniform Manifold Approximation Projection (UMAP) graphs show the cell- and stage-specific log
93 normalized expression of known marker transcripts that accurately define the GSC-like and CySC-like
94 clusters.
95
96 **RESULTS**
97 **Using a GSC and CySC tumor model for genomic studies**
98 Our understanding of the transcriptional, chromatin, and replication landscapes of *Drosophila*
99 GSCs is limited due to their small number in wild-type testes (Figure 1A, S1A). GSCs were
100 underrepresented in previous bulk genomic analyses of wild-type or early-stage germ cell-enriched testes
101 (Gan et al. 2010; Vedelek et al. 2018). Furthermore, even single-cell or single-nucleus transcriptomic
102 methods using whole wild-type testes yielded sparse GSC data (Witt et al. 2019; Yu et al. 2023; Raz et al.
103 2023; Shi et al. 2020).
104 To overcome this limitation, we used testes enriched with GSC-like and somatic CySC-like cells
105 by over-expressing the Janus kinase (JAK)-signal transducer and activator of transcription (STAT)
106 pathway ligand, Unpaired (Upd) (Figure 1B, S1B). Hyper-activation of the JAK-STAT pathway in the
107 testis results in over-proliferation of both stem cell types (Kiger et al. 2001; Tulina and Matunis 2001;
108 Leatherman and Dinardo 2010), resulting in a less heterogeneous tissue composition for genomic
109 profiling of stem cells from both germline and somatic lineages.
110 In our experimental design, we used two different cell type-specific drivers to overexpress Upd in
111 the fly testes. Driving expression of Upd with either a germline driver (*nanos* or *nos*-Gal4) or somatic
112 driver (*traffic jam* or *tj*-Gal4) leads to testis tumors with over-proliferative GSC-like and CySC-like cells.
113 For validation, a GFP-tagged histone H3 (H3-GFP) was driven along with Upd in either GSC-like cells
114 (*nos>upd*, *H3-GFP*) or CySC-like cells (*tj>upd*, *H3-GFP*). This tumor model reduces cellular

115 heterogeneity, making it an ideal system for studying GSC- or CySC-specific transcriptional, chromatin,
116 and replication landscapes.
117
118 **scRNA-seq profiling of GSC-like and CySC-like cells**
119 We performed single cell RNA-sequencing (scRNA-seq) on *nos*- and *tj*-driven Upd-
120 overexpressing tumor samples (*upd*-tumors), with two replicates each. Across these four samples, we
121 consistently identified distinct germline and somatic lineages, demonstrated through single-cell
122 integration, principal component analysis (PCA), and clustering (Figure S1C, S1D).
123 After individually clustering each sample, we integrated them to reduce batch effect variance, and
124 subset the data to GSC and CySC clusters. The majority of transcriptional variance was orthogonal to
125 tumor genotype variable (PC 1 *vs. nos>upd, H3-GFP*/*tj>upd, H3-GFP*: $R^2 = 0.06$, Figure S1D, S1E).
126 Consequently, most cells are clustered into two groups: GSC-like cells (expressing *vasa* and *nanos*) or
127 CySC-like cells (expressing *tj* and *zfh1*) (Figure 1C-G, Supplemental Table 1). Beyond these two main
128 clusters, we identified three additional clusters: primary spermatocytes (*mst87f*-expressing), muscle cells
129 (*act57B*-expressing), and an unidentified *wb*-expressing somatic lineage (Figure 1C, S1F, Supplemental
130 Table 1) (Kuhn et al. 1991; Kelly et al. 2002). For this study, we focus on the GSC-like and CySC-like
131 clusters.
132 The GSC-like and CySC-like clusters followed expected gene expression patterns for known cell
133 type- and stage-specific markers. For example, GSC-like cells, but not CySC-like cells, express *vasa* and
134 *nanos* (Figure 1D, 1E, Supplemental Table 1). On the other hand, CySC-like cells, but not GSC-like cells,
135 express *zfh1* and *tj* (Figure 1F, 1G, Supplemental Table 1). As a negative control, *eyeless* (*ey*), a gene
136 required for eye development, was not detected in any clusters (Figure S1G) (Halder et al. 1995).
137 The expression of stage-specific transcripts further confirms that the *upd*-tumor is enriched with
138 stem cell-like germline and cyst cells. As expected, *stat92E* is detected in both clusters, indicating JAK-
139 STAT pathway activation due to *upd* overexpression (Figure S1G, Supplemental Table 1) (Amoyel and

140 Bach 2012). Additionally, as expected for GSCs, neither *bag-of-marbles* (*bam*), which is a marker of
141 spermatogonia, nor *cannonball* (*can*), a spermatocyte-specific transcript, are strongly expressed in the
142 GSC-like cell cluster (Figure 1H, S1G, Supplemental Table 1) (McKearin and Spradling 1990; Hiller et
143 al. 2001). Consistently, *eyes-absent* (*eya*), a late-stage cyst cell marker, is not expressed in the CySC-like
144 cell cluster (Figure 1I, Supplemental Table 1) (Fabrizio et al. 2003).
145 Thus, the two major clusters from the *upd*-tumors correspond to GSCs and CySCs, expressing
146 characteristic cell type- and stage-specific transcripts. The size and purity of these clusters provide an
147 opportunity to robustly profile the transcriptomes of GSC- and CySC-like cells.
148
149 **CySC-like cells highly express intercellular communication genes while GSC-like cells enrich for**
150 **chromatin regulators**
151 To identify cell-type-enriched transcripts, we performed differential gene expression analysis. A
152 transcript was considered as having significantly higher expression in GSC-like cells if it had a fold
153 change >1.5 (GSC/CySC) and an S-value < 10^{-4} (estimated false sign rate) (Figure 2A, Supplemental
154 Table 2). We identified 827 genes enriched in GSC-like cells and 1,794 in CySC-like cells. These GSC-
155 like and CySC-like enriched gene sets contained expected markers, supporting cluster identity and the
156 robustness of our analysis (Figure 2B, Supplemental Table 2). First, the long-noncoding RNAs, *roX1* and
157 *roX2* (*RNA on the X*), components of the male dosage compensation complex found in somatic but not
158 germ cells (Shevelyov et al. 2022), rank among the top 100 genes with significantly higher expression in
159 CySC-like cells (Figure 2B, Supplemental Table 2). Second, *ago3*, encoding a germline-specific
160 Argonaute required for piRNA-mediated transposon repression (Williams and Rubin 2002), is among the
161 top 200 genes with significantly higher expression in GSC-like cells (Figure 2B, Supplemental Table 2).
162

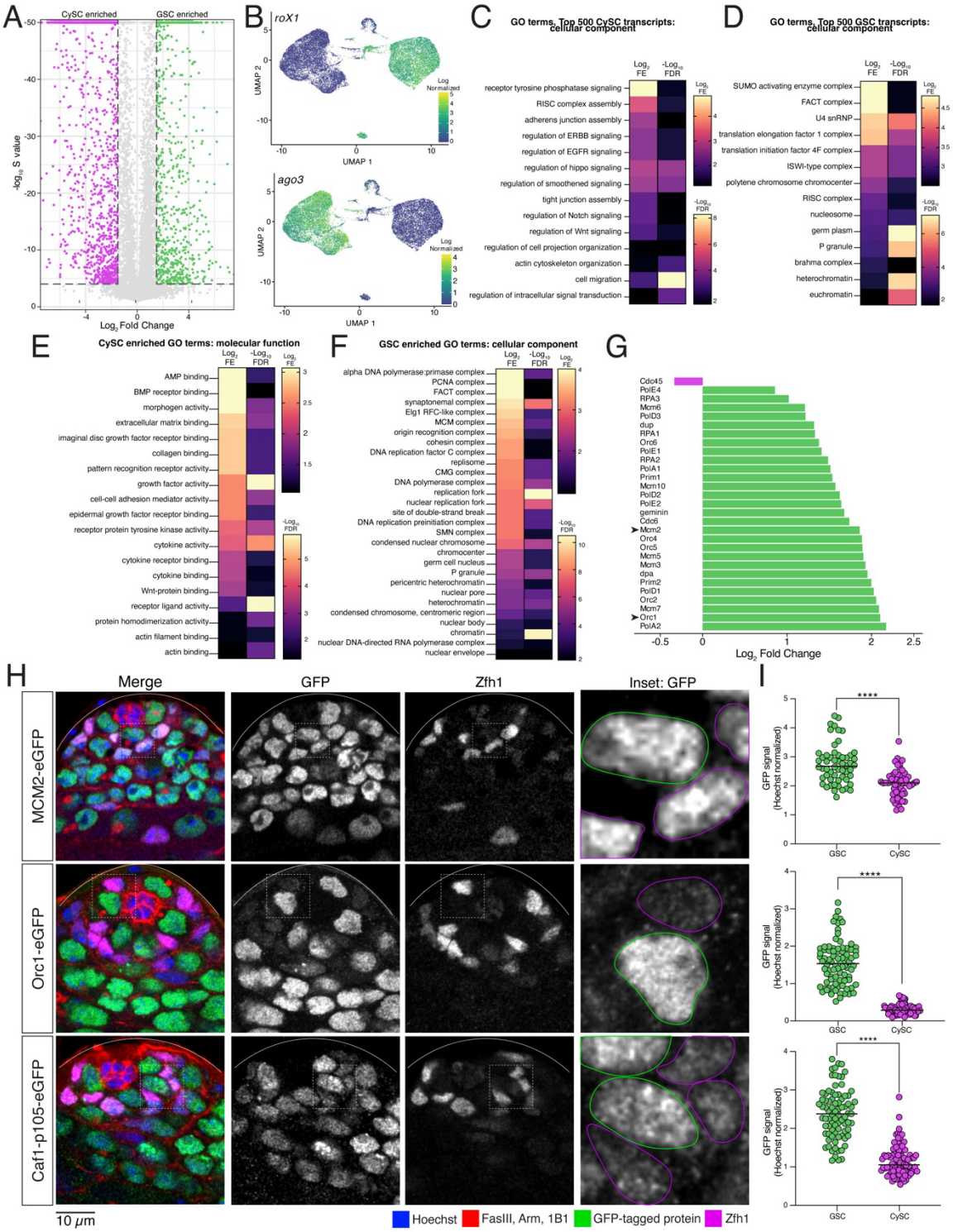


Figure 2: Intercellular communication in CySC-like cells and chromatin-based regulation in GSC-like cells

166 A: Differential gene expression analysis between the transcriptomes of the GSC-like and CySC-like
167 datasets. Differentially expressed transcripts are defined as having an absolute fold change difference
168 >1.5 (GSC/CySC) and an estimated false sign rate (S-value) $<10^{-4}$.
169 B: UMAP graphs of two known differentially expressed transcripts demonstrate the accuracy of cluster
170 identification. *roX1:* Log_2 FC= -6.01, S-value= 0; *ago3:* Log_2FC= 3.91, S-value= 0.
171 C-D: Gene Ontology (GO) analysis of the top 500 highest expressed genes in the GSC-like (at least
172 258.44274 CPM) and CySC-like (at least 246.321917 CPM) cell clusters. Shown are heatmaps of the
173 Log_2 Fold Enrichment (FE) and negative Log_{10} False Discovery Rate (FDR) results from the GO analysis.
174 E-F: Gene Ontology results for the significantly differentially expressed genes in GSC-like and CySC-
175 like cells. Shown are heatmaps of the Log_2 Fold Enrichment (FE) and negative Log_{10} False Discovery
176 Rate (FDR) results from the GO analysis.
177 G: Differential transcript expression (Log_2 GSC/CySC) for select genes involved in DNA replication,
178 shows the higher expression of these genes in GSC-like cells (green bars). One gene, Cdc45, is
179 differentially expressed in the CySC-like cells (magenta bar).
180 H: Confocal images of testis tips from fly strains with the indicated endogenously tagged GFP fusion
181 protein. Expression of each protein is restricted within the mitotically active spermatogonial region. The
182 inset displays a representative z-stack image containing a GSC (outlined in green) and CySC (outlined in
183 magenta) that would be used for quantification.
184 I: The amount of GFP fluorescence within the volume of the nucleus of a GSC or CySC was normalized
185 to the amount of Hoechst staining to produce the GFP signal value for that cell. The spectrosome, a germ
186 cell-specific organelle, and CySC-specific transcription factor Zfh1 were used to distinguish between
187 GSCs and CySCs (Leatherman and Di Nardo 2008; Lin et al. 1994). Two-tailed Mann-Whitney test,
188 ****= $p<0.0001$.
189

190 We performed Gene Ontology (GO) analysis to compare gene functions of expressed genes in
191 GSC-like and CySC-like cells. We first analyzed the top 500 expressed transcripts based on their CPM
192 (counts-per-million) values (top ~2.8%). CySCs are enriched with transcripts encoding cell-cell adhesion
193 molecules and signaling pathway components (Figure 2C, Supplemental Table 3). For example,
194 Cadherin-N (*cadN*) transcript level in the CySC-like cells (256.04 CPM) was significantly higher than in
195 GSC-like cells (0.17 CPM) (Supplemental Table 1). Components from several signaling pathways,
196 including the ErbB, Epidermal Growth Factor (EGF), Hippo, Wnt, and Notch pathways, were identified
197 as expressed highly in CySC-like cells (Figure 2C, Supplemental Table 3), consistent with their roles in
198 regulating CySC fate specification, division, and germline cyst enclosure (Schulz et al. 2002; Parrott et al.
199 2012; Kiger et al. 2000; Hudson et al. 2013; Ng et al. 2019). We also performed GO analysis on the 1,794
200 significantly higher expressed transcripts in CySCs. These transcripts are also enriched for genes involved
201 in intercellular communication, including paracrine signaling pathway components and adhesion proteins
202 (Figure 2E, S2A, Supplemental Table 3). For example, non-cell-autonomous Bone Morphogenic Protein
203 (BMP) signaling in somatic gonadal cells is essential for GSC maintenance and differentiation
204 (Shivdasani and Ingham 2003; Leatherman and Dinardo 2010; Kawase et al. 2004). Components of this
205 pathway, including the receptor Thickveins (Tkv) and the ligands Decapentaplegic (Dpp) and Glass
206 bottom boat (Gbb), are significantly higher expressed in GSC-like cells in our analysis (Supplemental
207 Table 2). Together, these profiles align with the known supportive roles of somatic gonadal cells in
208 spermatogenesis.
209 In contrast to CySC-like cells, the GO analysis of transcripts in GSC-like cells indicates the top
210 500 genes expressed in GSC-like cells are primarily chromatin regulators, such as the Facilitates
211 Chromatin Transcription (FACT), Imitation Switch (ISWI), and Brahma complexes (Figure 2D,
212 Supplemental Table 3) (Formosa and Winston 2020; Tyagi et al. 2016; Hu et al. 2021). For example,
213 FACT complex subunits were expressed approximately five times higher in GSCs than in CySCs
214 (Supplemental Table 1). We next examined the 827 significantly higher expressed genes in the GSC-like

215 cell cluster. The GSC-like transcriptome is differentially enriched for protein complexes involved in
216 chromatin regulation, such as DNA replication, heterochromatin maintenance, and chromatin accessibility
217 (Figures 2F, 2G, S2A, S2C-D, Table 1, Supplemental Table 3). Further, several chromatin-associated
218 complexes contain multiple components, all significantly differentially expressed in GSC-like cells (Table
219 1). For example, genes encoding both FACT complex subunits, five Minichromosome Maintenance 2-7
220 (MCM2-7) complex components, and four Origin Recognition Complex (ORC) components, are all
221 enriched in GSC-like cells compared to CySC-like cells (Table 1).

Complex Name	Component Names	Function
Origin Recognition Complex	**Orc1**, **Orc2**, Orc3, **Orc4**, **Orc5**, Orc6	Binds to origins of DNA replication
Minichromosome Maintenance Complex	**Mcm2**, **Mcm3**, **dpa**, **Mcm5**, Mcm6, **Mcm7, Mcm10**	DNA helicase unwinds dsDNA
Replication Factors	Cdc45, dup, **geminin**, **cdc6**	Factors regulating initiation of the replisome
Replication Protein A	RPA1, RPA2, RPA3	Binds to single stranded DNA
DNA polymerase epsilon	**Chrac-14**, PolE1, **PolE2**, PolE4	Synthesizes leading strand DNA
DNA polymerase alpha	**PolA1**, **PolA2**, **Prim1**, **Prim2**	Synthesizes the RNA primer that initiates DNA synthesis
DNA polymerase delta	**PolD1**, **PolD2**, PolD3	Synthesizes lagging strand DNA
Other DNA polymerases	PolH, PolZ1, PolZ2, PolI, PolG1, PolG2, PolQ	Varied polymerases for DNA Repair and mitochrondrial DNA synthesis
Replication Factor C/ Egl1 Complex	CG8142, **elg1**, **RfC3**, **RfC4**, **RfC38**	Regulate chromatin association of PCNA
DNA Ligases	**DNAlig1**, DNAlig3, DNAlig4	Joins DNA strands
DNA Helicases	Dna2, **RecQ4**, RecQ5, Pif1	DNA helicases functioning in DNA repair
DNA N-glycosylases	Tdg, Ogg1, Nthl1	Enzymes that remove damaged bases from DNA
DNA polymerase Rev1	**Rev1**	DNA polymerase involved in translesion synthesis
DNA Repair	**Brca2**	Involved in DNA Repair
Chromatin assembly factor 1	Caf-180, **Caf-105**, Caf-55	Assembles nucleosomes on newly synthesized DNA
Chromatin Accessibility Complex	**Acf**, **Chrac-14**, Chrac-16, Iswi	Regulates chromatin accessibility through nucleosome sliding
Histone Variants	His4r, **Cid**, **His2Av**	Histone variants
Histone Chaperones	Hira, Nap1	Histone chaperones
Nucleosome Remodeling Factor	Nurf-38, Caf1-55, E(bx), Iswi	Nucleosome sliding complex regulates accessibility for transcription

Remodeling and Spacing Factor Complex	CG8677, Iswi	Remodeling and spacing factor complex
SNF2-like Chromatin Remodelers and NuRD	Mi-2, dom, chd3	Disrupts DNA-histone interactions
FACT	**dre4**, **Ssrp**	Interacts with H2A-H2B dimers and facilitates nucleosome formation

222 **Table 1: Composition and function of select DNA replication and chromatin regulation complexes.**
223 Components in bold are differentially enriched in GSC-like cells.
224
225 To rule out potential bias caused by differences in cell cycles, we tested whether the enrichment
226 of DNA replication transcripts in GSC-like cells resulted from an over-representation of GSC-like cells in
227 G1- or S-phase on the dataset compared with CySC-like cells in these two cell cycle stages (Figure S2E).
228 Normalizing transcript representation across cell cycle phases such that both cell types have equivalent
229 percentages of cells in each cell cycle stage (33% G1, 33% S, 33% G2/M), did not change the enrichment
230 pattern of genes in GSC-like and CySC-like clusters (Figure S2F). Specifically, all assayed *mcm* and *orc*
231 genes remained GSC-enriched. Thus, the enrichment of DNA replication transcripts in GSC-like cells
232 reflects a true biological difference rather than an artifact of differences in their cell cycle phase
233 distribution.
234 We next tested whether the enrichment of DNA replication factors detected in GSC-like cells
235 from the *upd*-tumor also occurs in GSCs from non-tumor wild-type testes. Using genome editing, we
236 generated knock-in fly lines with *gfp*-tagged endogenous *mcm2*, *orc1*, and *caf1-p105*, genes considered
237 differentially enriched in GSC-like cells (Figure 2G, S2D, Table 1). We then performed immunostaining
238 on non-tumor testes to examine the expression patterns of these GFP-fusion proteins and quantified GFP
239 signals in wild-type GSCs and CySCs (Figure2H, 2I, S2G). In all instances, normalized GFP signal was
240 significantly higher in GSCs compared to CySCs (Figure 2I), demonstrating that the protein levels in
241 wild-type testes recapitulate the transcript enrichment identified in *upd*-tumor samples. Together, these
242 scRNA-seq data support the usage of *upd*-tumor testes in uncovering key characteristics of wild-type stem
243 cells through genomic strategies.

244 **Cell type-specific chromatin landscapes profiled by H3K4me3, H3K27me3, and H3K9me3**

245 The high expression of chromatin regulators in GSC-like cells, revealed by our scRNA-seq
246 analysis, suggests their potential roles in shaping distinct chromatin landscapes to regulate germline
247 functions. This prompted us to investigate cell-type-specific chromatin patterns in GSC-like and CySC-
248 like cells, respectively. To achieve this, we optimized a low-input cell-type-specific chromatin profiling
249 method using *upd*-tumor testes, which employs micrococcal nuclease (MNase) digestion of targeted
250 chromatin (Figure 3A). In this method, we co-expressed the Upd ligand and histone H3-GFP fusion
251 protein in either GSC-like cells (*nos>upd*, *H3-GFP*) or CySC-like cells (*tj>upd*, *H3-GFP*), as described
252 earlier. MNase digestion specifically targeted chromatin containing H3-GFP using an anti-GFP antibody
253 (Figure 3A). After separating the soluble (GFP-containing) from insoluble (GFP-negative) chromatin, 5%
254 of the soluble chromatin was sequenced as 'input', while the remaining chromatin was prepared for low
255 input ChIP-sequencing (ChIP-seq) using antibodies against H3K4me3, H3K27me3, or H3K9me3 (Figure
256 3A). We chose these histone modifications because they represent distinct chromatin domains: H3K4me3
257 (euchromatin), H3K27me3 (facultative heterochromatin), and H3K9me3 (constitutive heterochromatin)
258 (Kouzarides 2007). We quantified histone mark enrichment using a multiple regression model that
259 includes the input log-enrichment as intercept, a ChIP Log_2 Fold Enrichment (L_2FE) coefficient, and
260 covariates. The L_2FE coefficients robustly characterize the chromatin of the *Drosophila* chromosomes in
261 each sample (Figure 3B, S3A).

262 In agreement with previous findings (Kharchenko et al. 2011), H3K4me3 and H3K27me3 are
263 enriched along the euchromatic chromosome arms but depleted in the pericentromeric regions of
264 chromosomes 2, 3, and X (Figure 3B, S3A). In contrast, H3K9me3, which associates with constitutive
265 heterochromatin, shows enrichment at pericentromeric regions (Figure 3B, S3A). Together, these results
266 demonstrate that our cell-type-specific chromatin profiling assay accurately recapitulates chromatin
267 signatures consistent with the known distributions of H3K4me3, H3K27me3, and H3K9me3 (Kharchenko
268 et al. 2011).

269 To investigate relationships between chromatin patterns and cell-type-specific gene expression,
270 we integrated our cell-type-specific chromatin profiles with the scRNA-seq results. Focusing on the GSC-
271 like and CySC-like clusters, we defined a transcript as expressed ('on') if its CPM was above 5, a
272 threshold that reliably separates known cell-type-specifically expressed genes from non-expressed genes
273 (CPM< 5, 'off'). For example, using this cutoff, *ago3* is classified as 'off' in the CySC-like cluster (CPM
274 = 2.11), but is robustly expressed in the GSC-like cluster (CPM = 697.21) (Figure 2D, Supplemental
275 Table 1), indicating strong cell type-specific expression (Williams and Rubin 2002). We further classified
276 'on' genes into three categories based on their transcript levels: low, medium, and high (Figure 3C). The
277 GSC-like and CySC-like clusters contained similar numbers of 'on' and 'off' genes (GSCs: 6,842 'on'
278 and 10,716 'off'; CySCs: 6,484 'on' and 10,775 'off'), while the remaining three smaller clusters
279 displayed comparable numbers of 'on' and 'off' genes (Figure S3B).
280 Using these categories, we plotted mononucleosomal enrichment (as a density estimation of input
281 monosomal fragment midpoints) to understand how nucleosomal position and density influence gene
282 expression. The phased distribution of the H3 signal, indicative of nucleosome positioning, displays the
283 characteristic mononucleosomal pattern generated by the MNase (Figure 3D). This approach reveals the
284 strongly positioned +1 nucleosome near the transcription start sites (TSSs) of actively expressed genes in
285 both cell types, with positioning strength diminishing as gene expression levels decrease (Figure 3D,
286 Supplemental Table 4). This observation aligns with the known association between strongly positioned
287 nucleosomes and chromatin organization at active genes, underscoring the role of nucleosome positioning
288 in regulating gene expression and further validating our experimental approach (Chereji et al. 2018).

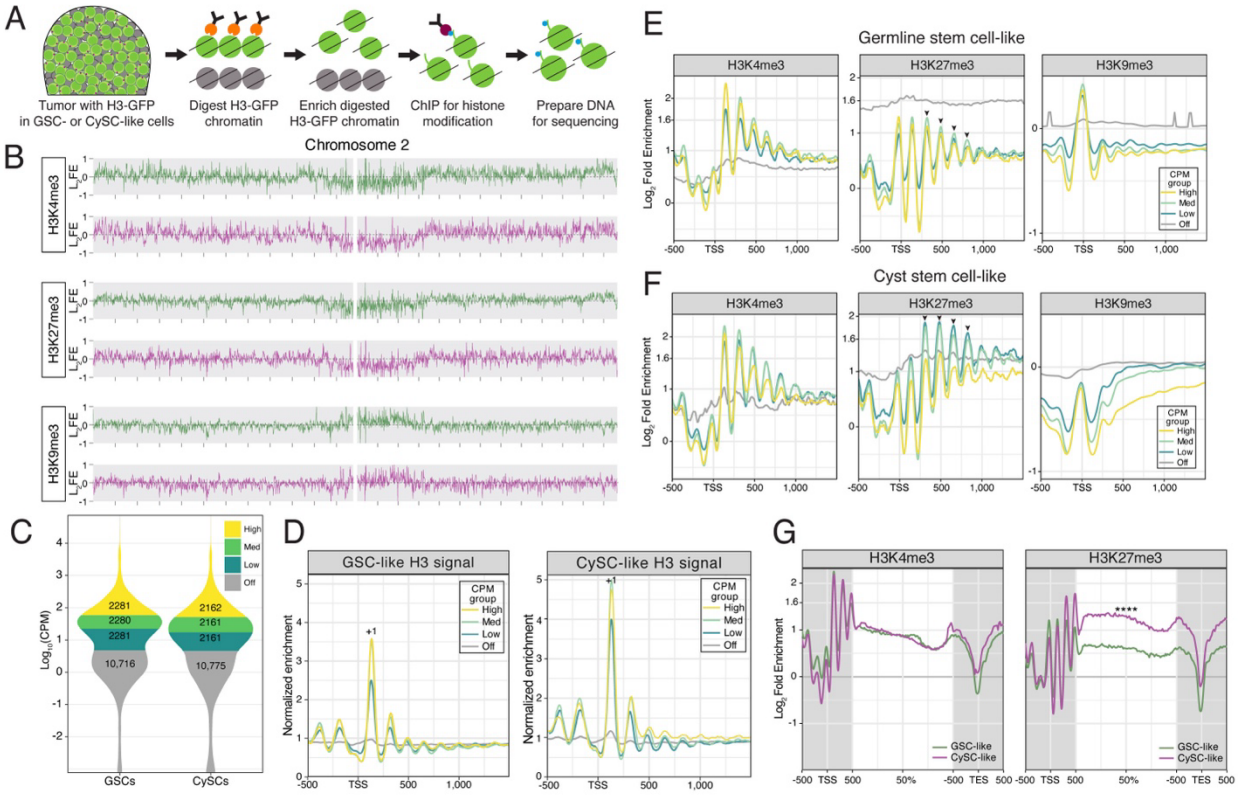


Figure 3: The cell-specific chromatin profiles of H3K4me3, H3K27me3, and H3K9me3

A: Schematic of the cell-specific chromatin immunocleavage (ChIC) coupled with chromatin immunoprecipitation (ChIP) protocol. Chromatin containing a histone H3-GFP fusion protein is targeted for digestion by antibody-guided MNase. After enriching for digested chromatin, a low-input ChIP protocol is performed for either the H3K4me3, H3K27me3, or H3K9me3 chromatin modification.

B: The Log2 Fold Enrichment (L_2FE) of H3K4me3, H3K27me3, and H3K9me3 along the euchromatic arms and pericentromeric region of Chromosome 2. Enrichment tracks from GSC-like cells are in green, whereas the tracks associated with CySC-like cells are in magenta.

C: Using a cut-off of CPM>5 as "on", the transcriptomes of the GSC-like and CySC-like clusters obtained from the scRNA-seq were divided into three equal parts to define genes as either off, low, medium, or high expression.

302 D: Enrichment of mononucleosome fragments is plotted at transcription start sites (TSS) of genes
303 separated by their expression level as defined in C. Normalized enrichment of mononucleosome
304 fragments was determined by normalizing to the median mononucleosomal fragment count on autosomes.
305 The +1 nucleosome is indicated in each graph. Enrichment at the maximum L_2FE value was compared to
306 the "off" gene value using a T-test for each CPM group, indicating a strong significant difference in
307 fragment enrichment (Supplemental Table 4).
308 E&F: Cell-specific histone modification enrichment was calculated as the Log2 Fold Enrichment
309 Log_2(Average mark signal / H3 (input) signal), then plotted at the TSS of genes categorized by their
310 expression level. The arrowheads indicate the position of nucleosomes where a difference in H3K27me3
311 enrichment is observed between GSC-like and CySC-like chromatin profiles. As done earlier, enrichment
312 at the maximum L_2FE value for each CPM group was compared to the "off" gene value using a T-test,
313 indicating a strong significant difference in modification enrichment (Supplemental Table 4).
314 G: Comparing the average gene profiles of L_2FE along gene bodies between GSC-like and CySC-like
315 cells shows no difference in enrichment for H3K4me3, but a significant increase in H3K27me3
316 enrichment in CySC-like cells (Supplemental Table 4).
317
318 **H3K27me3 plays a more prominent role in CySC-like cells compared to GSC-like cells**
319 To investigate the relationship between histone modifications and gene expression, we analyzed
320 modification enrichments at TSSs relative to transcript levels using the input-normalized Log_2 Fold
321 Enrichment (L_2FE) of H3K4me3, H3K27me3, and H3K9me3 for GSC-like and CySC-like cells. We
322 applied a LOESS model, regressing the histone modification L_2FE at TSSs with the gene expression
323 levels ($Log_{10}CPM$), to estimate mean L_2FE values for genes categorized as 'off', 'low', 'medium', and
324 'high' based on their median $Log_{10}CPM$ values (Figure S3C, S3E). In both GSC-like and CySC-like cells,
325 highly expressed genes are enriched for H3K4me3 but show depletion of H3K27me3 and H3K9me3

326 (Figure S3C). In contrast, silenced genes in both GSC-like and CySC-like cells display higher H3K9me3
327 enrichment compared to expressed genes (Figure S3C).
328 When examining the average histone modification enrichment at TSSs, we detected a significant
329 H3K4me3 peak at the +1 nucleosome for all expressed genes in both cell types (Figure 3E-F,
330 Supplemental Table 4). The H3K4me3 enrichment positively correlates with gene expression (Figures
331 S3C, S3E), as anticipated. Conversely, H3K9me3 shows minimal enrichment at expressed genes (Figures
332 3E-F, S3C, S3E), but is highly enriched at transposable elements, consistent with its known role in
333 marking constitutively repressed genomic regions (Figure S3D, Supplemental Table 4).
334 Among the three profiled histone modifications, H3K27me3 displays distinguishable cell-type-
335 specific differences between GSC-like and CySC-like cells, indicating potentially distinct roles in these
336 two cell types (Figure S3C, S3E). In GSC-like cells, H3K27me3 shows moderate enrichment across all
337 expressed gene categories (Figure 3E, S3C, S3E), whereas silenced genes display stronger enrichment
338 despite lacking well-positioned nucleosomes (Figure 3E). This non-canonical H3K27me3 enrichment
339 resembles the pattern observed in female *Drosophila* GSCs, where H3K27me3 is detectable at both
340 inactive and active gene loci (Deluca et al. 2020).
341 In contrast, in CySC-like cells, H3K27me3 is most enriched at the lowest-expressed genes and
342 gradually decreases with increasing gene expression levels (Figure 3F, S3C, S3E). Notably, this
343 H3K27me3 pattern is more consistent with its well-established role in gene repression. Further, in CySC-
344 like cells, H3K27me3 enrichment increases at the +3 nucleosome position and extends into the gene body
345 for medium- and low-expression genes, a distinct feature not observed in GSC-like cells (Figure 3E, 3F).
346 Previous studies demonstrate that broad H3K4me3 domains mark genes critical for cell identity
347 and function in stem and progenitor cells. As cells differentiate and lose potency, these broad H3K4me3
348 domains become refined (Zhang et al. 2021). Further, broad H3K27me3 enrichment across gene bodies is
349 associated with gene repression (Young et al. 2011). To investigate whether broad histone modification
350 patterns are present in the two populations of stem cells in testis, we analyzed histone modification

351 enrichment across gene bodies in genes expressed in both GSC-like and CySC-like cells (Figure 3G). Our
352 analysis revealed similar H3K4me3 enrichment along gene bodies in both cell types (Figure 3G,
353 Supplemental Table 4). However, H3K27me3 was significantly enriched along gene bodies in CySC-like
354 cells (Figure 3G, Supplemental Table 4). This differential enrichment indicates a cell-type-specific
355 regulatory mechanism for which H3K27me3 plays a more prominent role in CySC-like cells than in
356 GSC-like cells.

357 Previous functional studies in *Drosophila* testes demonstrate that the histone methyltransferase
358 Enhancer of Zeste [E(z)], which generates H3K27me3, plays a critical non-cell autonomous role in the
359 CySC lineage by preventing ectopic expression of a key somatic transcription factor in germ cells (Eun et
360 al. 2014). Conversely, inactivation of E(z) in the germline results in only mild defects, which arise in
361 later-stage germ cells rather than GSCs (Eun et al. 2017). This indicates H3K27me3 may not play a large
362 role in the early germline. Consistent with these findings, our scRNA-seq data show moderate *E(z)*
363 expression in both GSC-like and CySC-like cells (GSC-like CPM = 74.33, CySC-like CPM = 20.89,
364 Figure S3F, Supplemental Table S1). However, the transcripts for the demethylases responsible for
365 erasing the H3K27me3 modification, *Jarid2* and *Utx,* are more highly expressed in the GSC-like cell
366 cluster (*Jarid:* GSC-like CPM = 138.24, CySC-like CPM = 68.11; *Utx:* GSC-like CPM = 64.17, CySC-
367 like CPM = 31.06) (Figure S3G, S3H, Supplemental Table S1). Together, these findings highlight a non-
368 cell-autonomous role for E(z) and H3K27me3 in the CySC lineage, where they regulate the identity and
369 activity of the GSC lineage. Moreover, these data suggest a non-canonical H3K27me3 chromatin pattern
370 in male GSC-like cells, resembling the pattern detected in female GSCs (Deluca et al. 2020).
371

372 **Distinct replication timing profiles in GSC-like and CySC-like cells**

373 Chromatin organization plays an integral role in regulating transcription and DNA replication.
374 Given the differences in chromatin landscapes between GSC-like and CySC-like cells, we aimed to
375 understand how replication timing differs between these cell types using Replication-sequencing (Repli-

376 seq) (Figure 4A, S4A-C) (Marchal et al. 2018). First, we dissected *upd*-tumors expressing H3-GFP in
377 either GSC-like or CySC-like cells and incubated them with the thymidine analog Bromodeoxyuridine
378 (BrdU) to label replicating DNA. Next, GSC-like or CySC-like nuclei were extracted and sorted *via* flow
379 cytometry into four S-phase fractions corresponding to early, early-mid, late-mid, and late S-phase, based
380 on their DNA content (Figure 4A, S4A-C). BrdU-labeled DNA was then immunoprecipitated, followed
381 by library preparation and sequencing (Figure 4A).

382 We generated replication timing profiles, identifying early and late replicating genomic regions
383 on each chromosome. In these profiles, enrichment of Repli-seq reads in the early fraction yields a
384 positive score, just as late fraction enrichment yields a negative score. Consistent with known replication
385 timing patterns, pericentromeric heterochromatin is largely late-replicating in both GSC-like and CySC-
386 like cells (Figure 4B, 4C). On the other hand, in both cell types, replication timing oscillates between
387 early and late across the chromosomal arms of chromosomes 2, 3, and X (Figure 4B, 4C). The 2nd and 3rd
388 chromosomes comprise 70-80% of the *Drosophila* genome and display distinct chromatin landscapes
389 between the chromosome arms and pericentromeric regions (Figure 3B, S3A). To examine replication
390 progression relative to these chromatin domains, we plotted replication timing values for these two major
391 autosomes, separating the arms from the pericentromeric regions. Minimal replication timing differences
392 were detected between GSC-like and CySC-like cells along the major autosomal arms (Figure 4C).
393 Notably, while pericentromeric regions for both major autosomes were consistently late-replicating in
394 CySC-like cells, they displayed early replication features in GSC-like cells (Figures 4B, 4C,
395 Supplemental Table 5). Surprisingly, the 4th chromosome replicates early in GSC-like cells, despite its
396 known heterochromatic nature (Figure 4B, 4C) (Haynes et al. 2007). Overall, GSC-like cells exhibited a
397 trend where traditionally heterochromatic regions replicate earlier than expected, suggesting previously
398 unrecognized roles for these regions in regulating GSC functions.

399 We next compared replication timing values between GSC-like and CySC-like cells to identify
400 genomic regions with significantly different timing regimes. Using 3kb sliding windows, we tested the

401 null hypothesis that replication timing is equal between the two stem cell types. Nested peak calling was
402 then performed to identify regions at least 20kb in size with significant timing differences between the
403 two cell types. Regions were then categorized as 'GSC earlier' or 'CySC earlier' depending on whether
404 they replicate earlier in GSC-like or CySC-like cells. Overall, 25,12 Mb of the genome replicates earlier
405 in GSC-like cells compared to CySC-like cells, while 14.588 Mb replicates earlier in CySCs-like cells
406 (Figure 4D, 4E, Supplemental Table 6). The total genomic size of earlier replicating regions is similar
407 between GSC-like and CySC-like cells across the major autosomal arms and the X chromosome (Figure
408 4D). A notable exception is the right arm of chromosome 3, where the total amount of kb replicating
409 earlier in GSC-like cells is much greater than CySC-like cells (Figure 4D, Supplemental Table 6).
410 In heterochromatic regions, GSC-like cells exhibited a greater total genomic size of earlier-
411 replicating DNA compared with CySC-like cells (Figure 4E). A notable exception is the Y chromosome,
412 where most differentially timed regions replicate earlier in CySC-like cells than in GSC-like cells (Figure
413 4E). To examine whether differences in replication timing align with differences in histone modification
414 enrichment, we analyzed L_2FE values from our cell type-specific histone modification data. As expected,
415 regions that replicate earlier in GSC-like cells show significantly greater H3K4me3 enrichment in this cell
416 type (Figure 4F, Supplemental Table 5). Similarly, earlier replicating regions in CySC-like cells are
417 enriched for H3K4me3 in these cells (Figure 4F, Supplemental Table 5). These results indicate that cell
418 type-specific replication timing differences align with cell type-specific chromatin features. For the first
419 time, our results enable precise identification of specific chromosomal regions that differ in replication
420 timing and chromatin landscapes between germline and somatic stem cell populations in the *in vivo*
421 context.

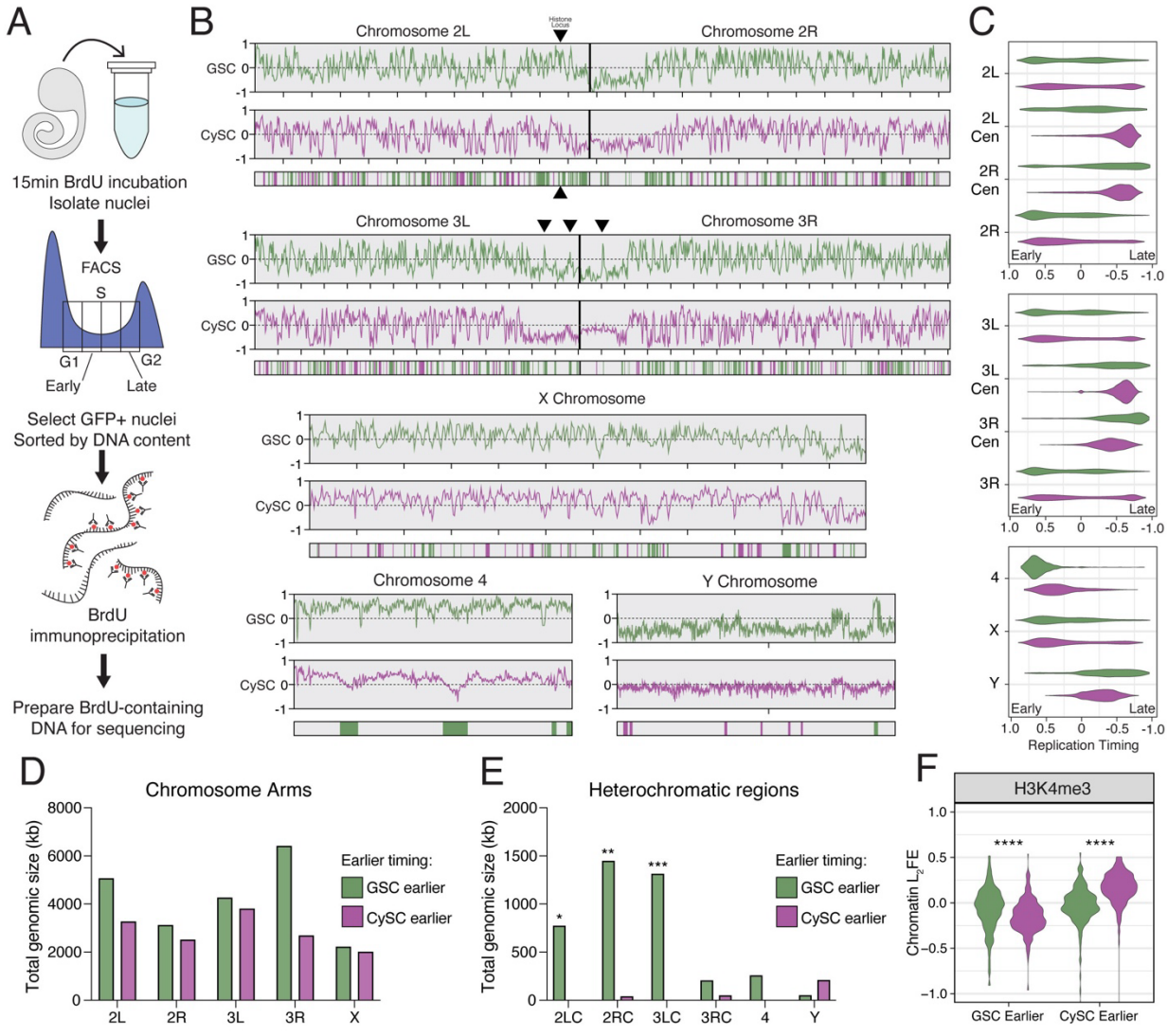


Figure 4: The replication timing profiles of GSC-like and CySC-like cells are distinct

A: A schematic of the Repli-seq protocol. First, dissected *upd*-tumors are incubated with BrdU, followed by nuclei isolation. GFP-positive nuclei are then sorted by DNA content, separating into one of four S-phase fractions: early, early-mid, late-mid, or late replicating. Next, the BrdU-containing DNA is immunoprecipitated and prepared for sequencing to identify genomic locations undergoing replication within the specific stages of S-phase.

B: Replication timing is scored from -1 to 1, where positive values represent early replicating and negative values represent late replicating. Replication timing line plots for GSC-like (green) and CySC-like (magenta) are separated for each chromosome. Beneath the line plots are the genomic windows

432 where replication timing significantly differed between GSC-like and CySC-like cells, where the color
433 indicates the cell type with earlier replication.
434 C: The replication timing values for each chromosome are plotted as violin plots and separated by
435 heterochromatic or euchromatic regions. A T-test of the logit-transformed violins indicated there is a
436 significant difference in the replication timing values for all comparisons between GSC-like and CySC-
437 like distributions, except the X chromosome. Cen = pericentromeric heterochromatin.
438 D&E: The total genomic size in kb for each chromosome for which we rejected the null hypothesis of no
439 difference in timing among GSC-like or CySC-like cells, grouped by the earlier-replicating cell type. In
440 general, heterochromatic regions are earlier replicating in GSC-like cells.
441 F: The Log_2 Fold Enrichment values for H3K4me3 are significantly elevated in regions that are cell-
442 specifically earlier replicating.
443
444 **Replication timing of GSC-like cells is distinct from that of CySC-like and cultured *Drosophila* cells**
445 As most studies on replication timing have focused on cultured cell lines, there are only a few
446 studies from cells *in vivo* in a normal developmental context (Armstrong et al. 2018; Pratto et al. 2021).
447 To gain cell-type specific insights and compare replication timing between *in vivo* and *in vitro* cells, we
448 compared our datasets with published replication timing data from two well-studied *Drosophila* cell lines,
449 Kc167 and S2 (Lubelsky et al. 2014). Clustering analysis on Euclidean distance demonstrated that the
450 replication timing profile of GSC-like cells is distinct from the other three cell types, while the two cell
451 lines cluster together (Figure 5A).
452 To facilitate cross-cell-type comparisons, we standardized replication timing along the reference
453 genome to have unit variance, yielding RT z-scores. We next categorized the genome into 12
454 chromosomal domains with their mean RT z-scores: two major autosomal arms and their pericentromeric
455 regions, the 4th, X, and Y chromosomes, and a gene consensus for the rDNA locus (Figure 5C). This
456 allowed us to characterize negatively and positively correlated cell type profiles (Figure 5B).

457 Interestingly, among the four cell types, GSC-like and Kc167 cells show the strongest correlation in their
458 chromosomal region RT z-score profiles (Figure 5B). This is intriguing because Kc167 cells are female
459 (XX karyotype) (Cherbas and Gong 2014). Similarly, CySC-like and S2 cells display positively
460 correlated RT z-scores, likely due to their shared somatic identity and male (XY) karyotype.

461 We additionally summarized the chromosomal region RT z-scores by their arithmetic mean
462 (Figure 5C). These summarized z-scores, relative to the combined dataset mean, indicate how each
463 genomic region in each cell type deviates from the dataset's average (Figure 5C). While chromosomal
464 arms display no overall trends, this method identified earlier replication timing in GSC-like cells for the
465 pericentromeric region of Chromosome 2L (Early-Mid), the 4th chromosome (Early), and the rDNA locus
466 (Early) (Figure 5C). Although replication timing for the X chromosome was found to be Early-Mid and
467 the Y chromosome to be Late-Mid in both GSC-like and CySC-like cells (Supplemental Table 5), the z-
468 scores revealed more genomic regions on both sex chromosomes that replicate earlier in CySC-like cells
469 (Figure 5C).

470 Pairwise comparisons of replication timing profiles across the four cell types identified genomic
471 windows with distinct replication timing. As shown above, we quantified the total size (kb) of earlier-
472 replicating regions and mapped their genomic locations by cell type. The chromosome arms of the 2nd and
473 3rd chromosomes show minimal differences in the amount or location of earlier-replicating regions across
474 cell types (Figure 5D, 5E, S5A, S5B). Thus, while small differences in replication timing may be present
475 along the chromosome arms, there are no apparent large-scale cell-specific differences.

476 We next focused on regions where the mean RT z-score suggests earlier replication in GSC-like
477 cells, such as the pericentromeric region of chromosome 2L and the 4th. In both cases, the total genomic
478 size of earlier replicating regions is greater in GSC-like cells than in other cell types (Figure 5F, 5G, S5C,
479 S5D, Supplemental Table 6). The pericentromeric region does not exhibit specific locations with
480 significantly different replication timing; rather, the entire region displays earlier replication in GSC-like

481 cells (Figure 5F). In contrast, the 4th chromosome contains three distinct regions with comparatively
482 earlier replication timing in GSC-like cells (Figure 5G, S5D).
483 The Y chromosome generally replicates late in S-phase across all three male cell types (Figure
484 5H). However, CySC-like cells show a relatively earlier replication timing profile, as reflected by their
485 mean RT z-score (Figure 5C). Indeed, the total genomic size of earlier-replicating regions is larger in both
486 CySC-like and S2 cells compared to GSC-like cells, while CySC-like and S2 cells show similar values
487 (Figure 5H, Supplemental Table 6). Unlike the 4th chromosome, replication timing for the Y chromosome
488 does not reveal any consistent earlier-replicating regions. However, a few specific loci show dramatically
489 different replication timing when comparing GSC-like to CySC-like cells, which are subject for
490 subsequent analyses (Figure 5H).
491 Finally, when comparing male cells (GSC-like, CySC-like, or S2) to female Kc167 cells, the X
492 chromosome consistently replicates earlier in the male cell, whereas the average replication timing of the
493 X chromosome was remarkably similar among male cell types (all with a summary RT value of 0.48 ±
494 0.01) (Figure 5I, S5E, Supplemental Table 5). This predominance of early-replication of the X
495 chromosome in male cells likely reflects the accessible chromatin environment that facilitates male-
496 specific X chromosome dosage compensation (Urban et al. 2017). Interestingly, early X chromosome
497 replication is preserved in GSC-like cells, despite the absence of dosage compensation in this cell type
498 (Figure 2D) (Meiklejohn et al. 2011). Our scRNA-seq further supports this note, showing no evidence of
499 X-linked gene dosage compensation in GSC-like cells or primary spermatocytes, based on the X-to-
500 autosome (X:A) transcript abundance ratios (Figure S5F). Overall, our analysis identifies key genomic
501 regions with cell type-specific replication timing differences, including heterochromatic pericentromeric
502 regions and the 4th, X, and Y chromosomes, suggesting distinct functional requirements for these domains
503 across different cell types.
504

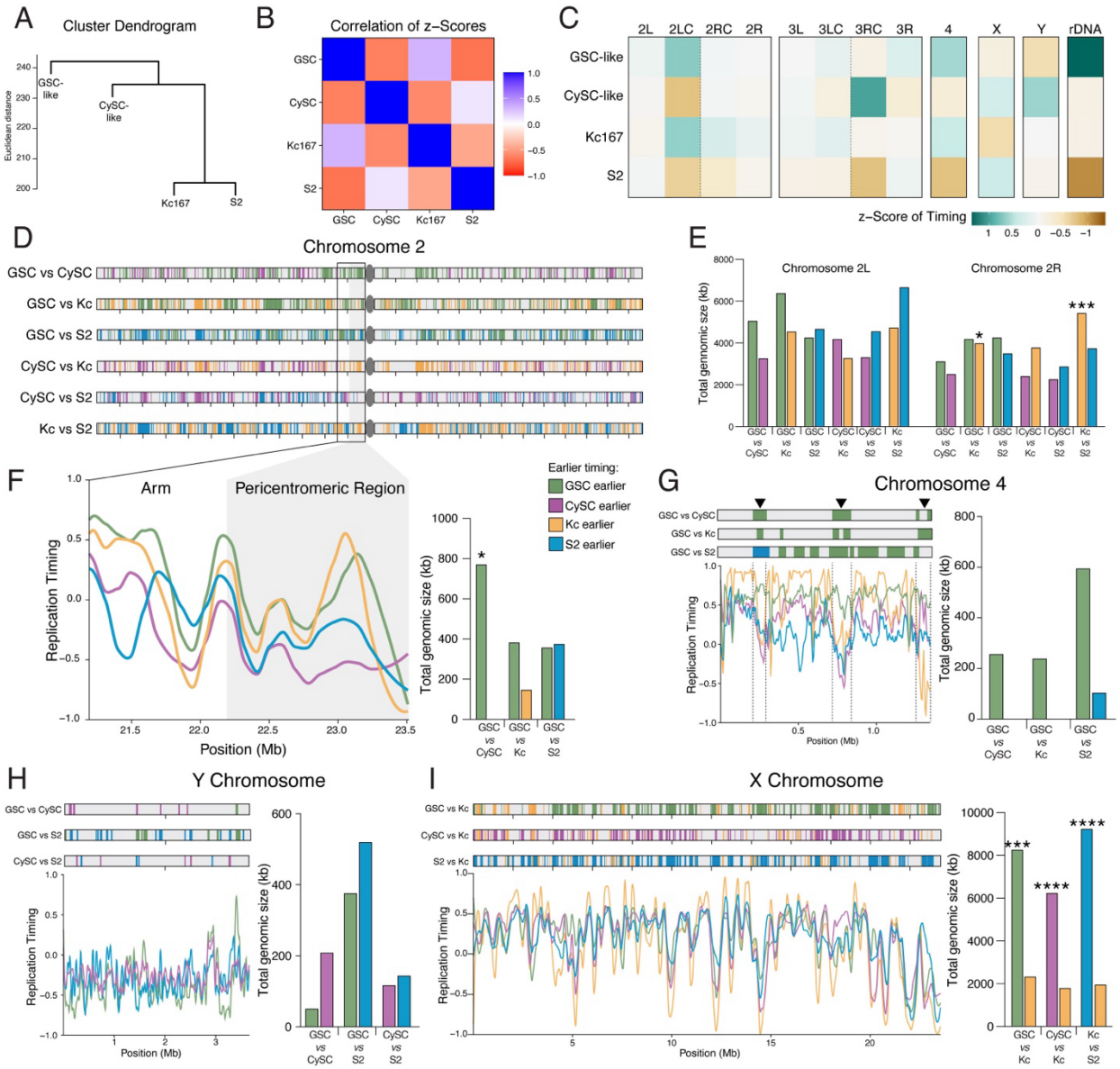


Figure 5: GSC-like cells have a distinct replication timing program from CySC-like and cultured *Drosophila* cells

A: Hierarchical comparison of replication timing indicates that the GSC-like cell replication program is dissimilar to the other three replication programs. On the y-axis, tree branches are shown at the root mean squared z-score differences of all genomic windows (Euclidean distance), and cell-cell distances at each tree branch are combined using the average method.

512 B: Replication timing (RT) z-scores were determined by standardizing variance at each window on the
513 genome, then averaged across 11 chromosomal regions and scored at the ribosomal DNA (rDNA) locus,
514 yielding cell-type profiles with a dynamic Pearson correlation. In terms of these shifts in timing of distinct
515 chromatin locations, the GSC-like and Kc167 replication programs share significant similarities (positive
516 correlation), as to a lesser extent do the programs of the CySC-like and S2 cells.
517 C: The mean replication timing z-score for 11 defined chromosomal regions, and the z-score of the rDNA
518 unit replication timing, indicate how this summary of the genome hierarchically explains the cell types'
519 epigenomes.
520 D: Pairwise comparison of replication timing differences for Chromosome 2, where domains (reaching a
521 size at least 20 kb) are evaluated against the null hypothesis of having equal replication timing. Shown is
522 the location and identification of regions that are replicating earlier in the indicated comparison.
523 E: Total genomic size of pairwise comparisons for total amount of the genome replicating earlier in a
524 specific cell type on the left and right arms of Chromosome 2.
525 F: Replication timing line plots for each cell type within the pericentromeric region of Chromosome 2
526 demonstrates the locations where the GSC-like cells replicate early in comparison to the other cell types.
527 A quantification of the total genomic size where GSC-like cells replicate earlier is presented on the right.
528 G: Compared to the other cell types, the 4th chromosome has more genomic regions that replicate earlier
529 in GSC-like cells. Three of these genomic windows are consistent when comparing to the other three cell
530 types and marked with an arrowhead.
531 H: Comparison of replication timing between male cell types for the Y chromosome shows the generally
532 late replication timing of this chromosome.
533 I: Compared to female Kc cells, all male cells have a greater total genomic size of earlier replicating DNA
534 on the X chromosome. These early-replicating domains span the entirety of the X chromosome.
535
536

537 **Cell-type-specific replication timing correlates with histone modification patterns**

538 The temporal order of genome duplication in eukaryotic cells is intricately linked to chromatin
539 structure and gene expression, both of which are cell-type-specific (Rhind and Gilbert 2013). To
540 investigate this relationship, we integrated our replication timing data with the cell-type-specific histone
541 modification profiles generated in this study. Additionally, we reanalyzed published Repli-seq and ChIP-
542 chip data from Kc167 and S2 cells to compare those *in vitro* datasets with our *in vivo* results (Kharchenko
543 et al. 2011; Lubelsky et al. 2014).

544 Consistent with previous findings in Kc167 and S2 cells (Eaton et al. 2011), H3K4me3 is
545 enriched at early replicating regions, coinciding with genes that have the highest expression levels (Figure
546 6A-B, S6A-C). Accordingly, elevated H3K4me3 is present at the TSSs of genes located in the early-
547 replicating fraction in both GSC-like and CySC-like cells (Figure S6E-F). In contrast, H3K9me3 is found
548 in gene-poor and late-replicating regions across all cell types (Figure 6A-B, S6A-B, S6D).

549 On the other hand, H3K27me3 displays distinct enrichment patterns at different S-phase stages
550 across cell types. In CySC-like cells, H3K27me3 is present throughout early and early-mid S-phase
551 (Figure 6B-C), overlapping with H3K4me3, and is enriched at TSSs of genes replicating in these stages
552 (Figure S6F). Notably, this H3K27me3 pattern is distinct from the other three cell types, further
553 supporting the observation that there is a unique H3K27me3 chromatin signature in CySC-like cells. In
554 contrast, H3K27me3 is predominantly enriched during mid S-phase in GSC-like, Kc167, and S2 cells
555 (Figure 6A, 6C, S6A-B). Consistently, in GSC-like cells, H3K27me3 is most enriched at TSSs of genes
556 replicated during early-mid S-phase (Figure S6E).

557 Our replication timing analysis revealed cell-type-specific timing regimes for pericentromeric
558 heterochromatin on Chromosome 2L and globally for the 4th, X, and Y chromosomes (Figure 5). We next
559 tested whether these regions have corresponding differences in chromatin landscapes. To compare histone
560 modification differences across genomic regions between GSC-like and CySC-like cells, we defined a
561 given locus as having cell-type-specific histone modification enrichment if its L_2FE was >0.2 in one cell

562 type and <-0.2 in the other. This analysis revealed strong overall correlations, suggesting broadly similar
563 histone modification landscapes (Figure S6G). However, distinct GSC- or CySC-enriched histone
564 modification regimes were also identified (Figure 6D, S6G, green and purple boxes). To quantify cell-
565 type-specific chromatin features, we calculated the percentage of specific genomic regions enriched with
566 each histone modification (Figure 6D). For example, 7.4% of the pericentromeric region of Chromosome
567 2L in GSC-like cells is enriched for H3K4me3, compared to only 3.2% in CySC-like cells (Figure 6D,
568 Supplemental Table 7), aligning with its earlier replication timing in GSC-like cells (Figure 5C). In
569 contrast, the 4th chromosome in GSC-like cells shows almost no specific enrichment of H3K4me3 or
570 H3K27me3 (Figure 6D, Supplemental Table 7), despite its markedly early replication in this cell type. In
571 CySC-like cells, a greater percentage of chromatin on the X and Y chromosomes is enriched with both
572 H3K4me3 and H3K27me3 (Figure 6D).

573 We next evaluated histone modification enrichment scores at regions with distinct replication
574 timing to evaluate correlations between cell-type-specific replication timing and chromatin features. In
575 GSC-like cells, earlier replicating regions in Chr2L pericentromeric heterochromatin displayed higher
576 H3K4me3 enrichment than in CySC-like cells (Figure 6E, Supplemental Table 5). Interestingly, regions
577 on the 4th chromosome that replicate earlier in GSC-like cells do not display a similar increase in
578 H3K4me3 L_2FE values when compared with CySC-like cells (Figure 6F, Supplemental Table 5),
579 suggesting that the early replication of this chromosome may not be linked to a chromatin mark like
580 H3K4me3.

581 In CySCs, the X and Y chromosomes replicate earlier (Figure 5C), and their earlier-replicating
582 regions show higher levels of H3K4me3 and H3K27me3 compared to those in GSC-like cells (Figures
583 6G-H, Supplemental Table 5). A similar observation was made for regions on the X and Y with earlier-
584 replication in GSC-like cells (Figure S6H, S6I, Supplemental Table 5), indicating a chromosome- and
585 cell-type-specific relationship between replication timing and histone modifications. In summary, this

integrative analysis reveals a relationship between cell type- and chromosome-specific chromatin landscapes and DNA replication programs in *Drosophila* testis stem cells.

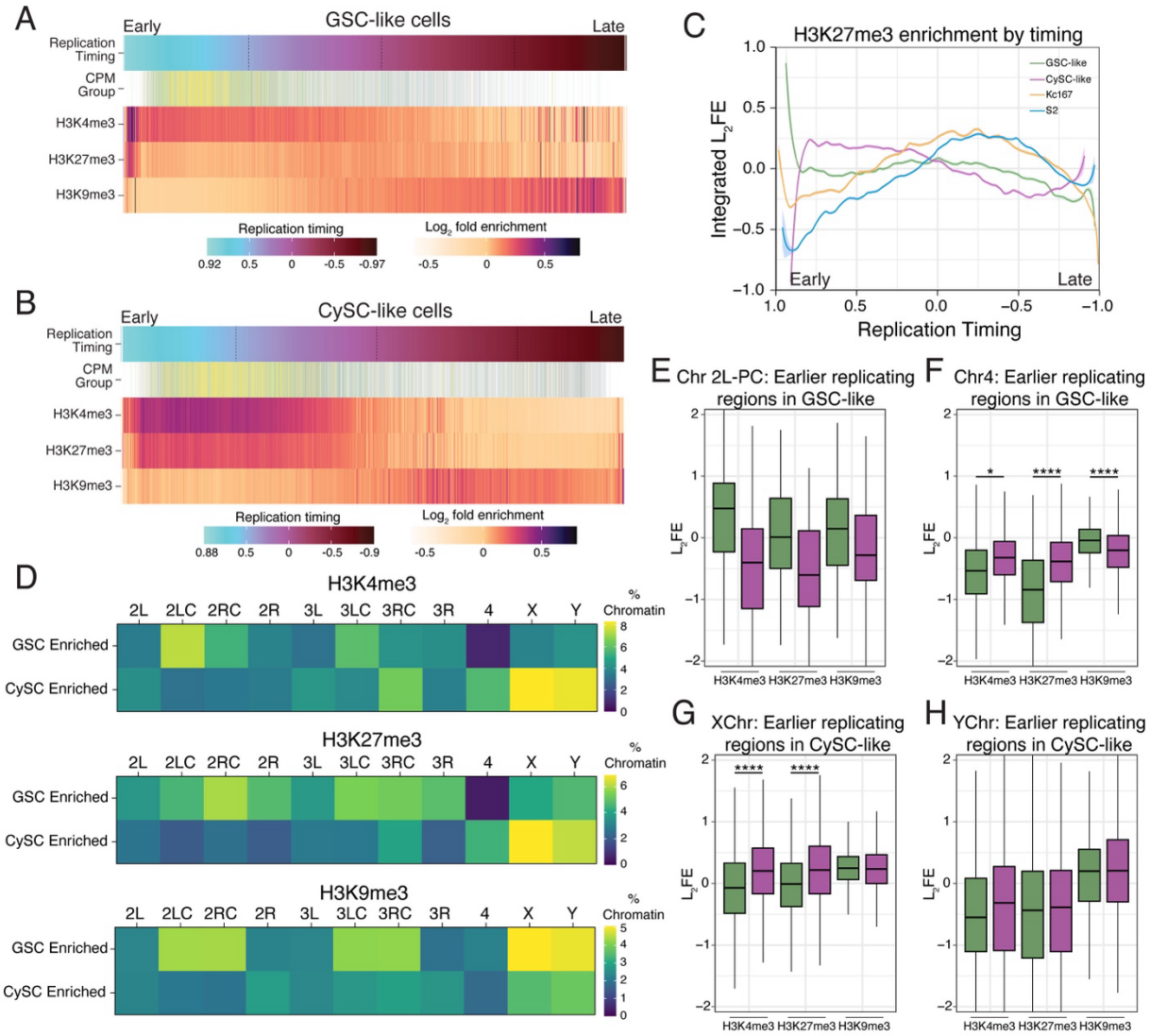


Figure 6: Cell-specific replication timing regimes correlate with cell-specific histone enrichment

A: For GSC-like cells, the Log_2 Fold Enrichment (L_2FE) values for H3K4me3, H3K27me3, and H3K9me3 are shown for genomic locations ordered from early to late by their replication timing value. The expression level for genes (categorized and colored by their CPM group as in Figure 3) is also shown indicating genes with the highest expression are replicated earlier in S-phase.

594 B: The L_2FE values for H3K4me3, H3K27me3, and H3K9me3 are shown for CySC-like cells within
595 genomic locations ordered by their replication timing (from early to late). Gene expression level
596 (categorized and colored by CPM group) is also shown indicating genes with the highest expression are
597 replicated earlier in S-phase.
598 C: The enrichment levels of H3K27me3 are plotted according to replication timing for GSC-like, CySC-
599 like, Kc167, and S2 cells, demonstrating increased H3K27me3 enrichment at regions that replicate early
600 in CySC-like cells compared with the other three cell types.
601 D: The calculated percent of chromatin within in each indicated chromosomal domain that is categorized
602 as enriched in either GSC-like or CySC-like cells. Cell-specific enriched regions are considered as having
603 L_2FE values >0.2 in one cell type and <-0.2 in the other.
604 E-H: Box and whisker plots of L_2FE for histone modifications in GSC-like (green) and CySC-like
605 (magenta) cells on the indicated chromosomes. The values here are a subset of enrichment scores found in
606 genome regions that replicate earlier in the indicated cell type. The '*' represents statistical significance
607 as measured by a T-test (Supplemental Table 5).
608
609 **Selected features within differentially replicating regions in germline and somatic lineages**
610 To explore the relationship between cell-type-specific replication timing and gene expression, we
611 assessed gene-specific replication timing and expression differences between GSC-like and CySC-like
612 cells within regions identified as differentially replicating. In the pericentromeric region of Chromosome
613 2L, several differentially expressed genes, such as *Marf1*, *mir-2490*, and *CR46253*, replicate earlier in
614 GSC-like cells than in CySC-like cells (Figure 7A-B). While *Marf1* represses *nanos* expression and
615 promotes meiotic progression during oogenesis in the female germline (Kawaguchi et al. 2020), the
616 functions of *Marf1*, *mir-2490,* and *CR46253* in the male germline remain unknown. Notably, the genomic
617 loci for these regulatory RNAs are associated with Orc2 binding sites, suggesting potential replication
618 origins within this region (Figure 7B). These observed differences in replication timing between GSC-like

619 and CySC-like cells may reflect cell-type-specific origin usage that could arise from transcriptional
620 differences within this region.

621 Extending beyond Chromosome 2L, the X chromosome also exhibits distinct replication timing
622 and histone modification enrichment in a cell-type-specific manner. In male cells, the X chromosome
623 generally replicates early, with CySC-like cells displaying higher H3K4me3 enrichment (Figure 6D). The
624 correlation between early replicating regions and active histone modification enrichment further supports
625 the connection between chromatin state and replication timing. Categorizing X-linked genes by cell-type-
626 specific replication timing revealed a strong alignment between earlier replication and higher expression
627 in the corresponding cell type (Figure 7C-D). Notably, GSC-like cells showed more non-coding genes in
628 early-replicating regions, suggesting a potential regulatory role for non-coding RNAs in this cell type
629 (Figure 7E-F).

630 Similarly, distinct replication patterns are also observed on the 4th chromosome. Several loci on
631 the 4th chromosome replicate earlier in GSC-like cells across all pair-wise comparisons (Figure 5G). One
632 64kb domain contains six tandem genes that synchronously replicate early (Figure 7G), clustering in the
633 Early GSC, Late CySC quadrant (Figure 7H). This region has five protein-coding genes and one
634 uncharacterized long non-coding RNA. This region is also associated with an Orc2 binding site,
635 suggesting cell-type-specific origin usage (Figure 7G).

636 The Y chromosome also follows distinct replication timing patterns, with certain genes exhibiting
637 early replication in a cell-type-specific manner despite its largely late-replicating nature. Among its 16
638 annotated protein-coding genes, functional data on these genes remain limited. For example, the *CCY*
639 gene replicates early in CySC-like cells (Figure 7I) and is essential for sperm individualization during late
640 spermiogenesis, though its role in somatic gonadal cells remains unknown (Hafezi et al. 2020; Zhang et
641 al. 2020). The presence of these genes in cell-type-specific early-replicating regions suggests their
642 unexplored roles within the somatic lineage.

Taken together, these findings reveal a broader pattern linking replication timing, chromatin state, and cell identity. In summary, our replication timing data aligns with the cell-type-specific chromatin landscape, highlighting distinct regulatory regimes for the X, Y, and 4th chromosomes in the soma *versus* germline. A broader pattern emerges: regions that are typically heterochromatic in differentiated cells replicate earlier in GSC-like cells. This is reminiscent of findings in *Drosophila* female GSCs and cultured mouse embryonic stem cells, which also exhibit reduced heterochromatic marks compared to differentiated cells (Pang et al. 2023; Rausch et al. 2020). Thus, the male germline provides a valuable system to study gene expression, chromatin regulation, and DNA replication in endogenous adult stem cell systems of a multicellular organism.

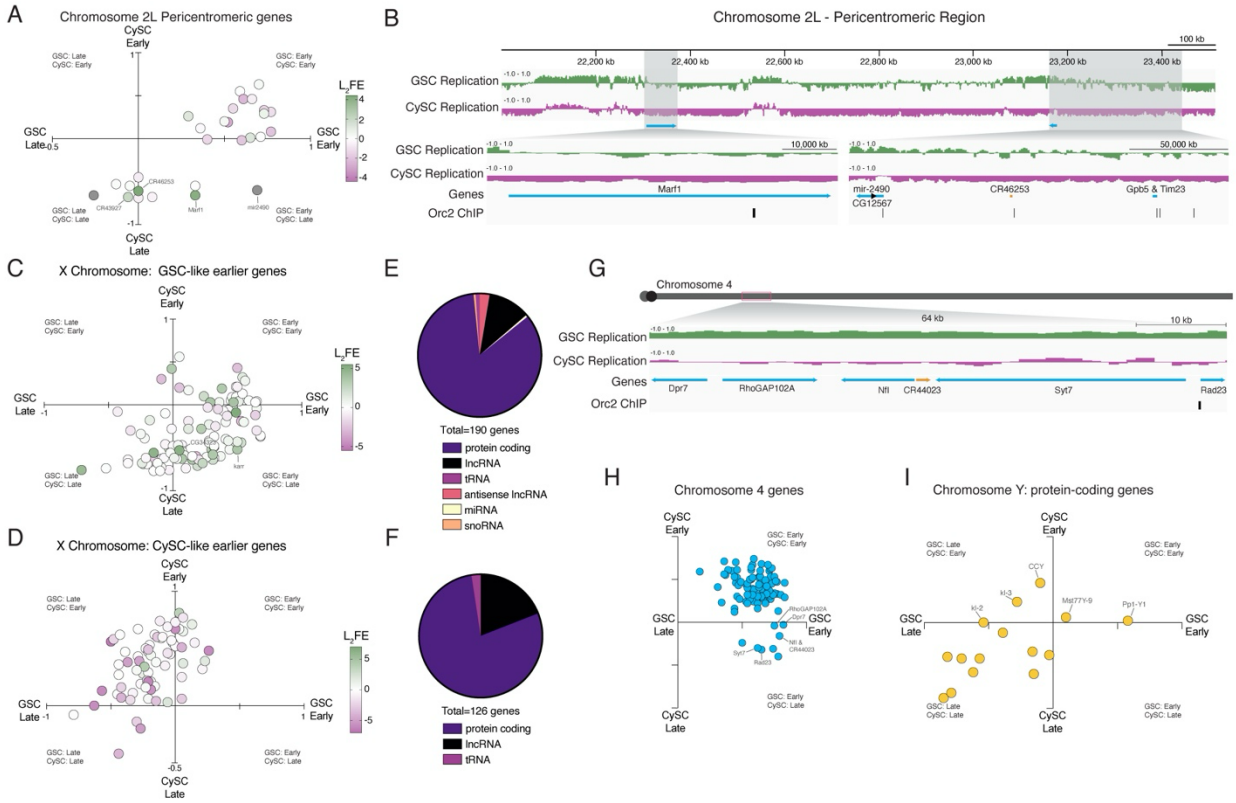


Figure 7: Selected genomic features within differentially replicating regions

A: Scatterplot of the GSC-like (X-axis) and CySC-like (Y-axis) replication timing values for genes located within the pericentromeric region of chromosome 2L. Each data point represents a gene, where

656 each gene is colored by its Log_2 fold change value calculated by differential gene expression analysis of
657 our scRNA-seq (Supplemental Table 2).
658 B: A genome-browser image of the pericentromeric region of chromosome 2L. In green are the
659 replication timing tracks for GSC-like cells, where the replication timing tracks for CySC-like cells are in
660 magenta. Shown highlighted are the genes found to have differential replication timing and gene
661 expression in GSC-like cells. Also indicated is the location of Orc2 binding sites, which represent
662 putative origins of replication.
663 C-D: Scatterplots of the GSC-like (X-axis) and CySC-like (Y-axis) replication timing values for genes
664 located on the X chromosome, separated by whether they replicate earlier in GSC-like or CySC-like cells.
665 Each data point represents a gene, where each gene is colored by its Log_2 fold change value which is
666 calculated using differential gene expression analysis of our scRNA-seq (Supplemental Table 2).
667 E-F: Shown is the proportion of gene types that replicate earlier in GSC-like or CySC-like cells. The
668 majority are protein-coding genes, where the next largest gene category is long non-coding RNAs.
669 G: A genome-browser image of the 4th chromosome with one 64kb region highlighted that is earlier
670 replicating in GSC-like cells. The replication timing tracks for GSC-like cells are in green and the
671 replication timing tracks for CySC-like are magenta. Shown is the location of six tandem genes residing
672 in this earlier replicating domain, including a non-coding RNA and Orc2 binding site.
673 H: Scatterplot of the GSC-like (X-axis) and CySC-like (Y-axis) replication timing values for genes
674 located on the 4th chromosome. Each data point represents a gene. Genes present within the 64kb region
675 cluster together as earlier replicating in GSC-like cells.
676 I: Scatterplot of the GSC-like (X-axis) and CySC-like (Y-axis) replication timing values for genes
677 located on the Y chromosome. Each data point represents a gene. Genes that replicate earlier in CySC-
678 like cells with potential roles in male fertility are indicated as marked.
679
680

681 **DISCUSSION**

682 Understanding the mechanisms underlying cell fate specification and fate transitions is essential
683 to developmental and stem cell biology. In this study, we integrated transcriptional profiling, chromatin
684 landscape analysis, and one of the first *in vivo* replication timing studies in adult stem cells to identify
685 both shared and distinct regulatory mechanisms of *Drosophila* male germline and somatic cyst stem cells.
686 We further compared our *in vivo* datasets with *in vitro* cultured *Drosophila* cell lines. In this way, we
687 uncovered stem cell-type-specific differences in genomic and epigenomic features.

688 Our transcriptional profiling highlights a fundamental regulatory dichotomy between GSC-like
689 and CySC-like cells. First, CySC-like cells rely on intercellular communication, with many highly
690 expressed genes associated with signaling pathways that coordinate germline-soma interactions (Vidaurre
691 and Chen 2021; Papagiannouli 2014). Second, GSC-like cells likely have a greater focus on regulating the
692 chromatin landscape, with an overrepresentation of DNA replication and chromatin-modifying factors.
693 This difference aligns with the well-established roles of intercellular signaling in coordinating GSC and
694 CySC identity and function. While intercellular communication appears an essential function for the
695 somatic lineage, the strong enrichment of chromatin factors in GSCs suggests that chromatin regulation
696 plays a more dominant role in regulating germline stem cell identity and activity.

697

698 **Cell-type-specific heterochromatin replication timing**

699 Our study reinforces the well-established relationship between chromatin features and replication
700 timing, where transcriptionally active euchromatic regions replicate early and transcriptionally inactive
701 heterochromatic regions replicate late. However, by comparing these dynamics between two distinct adult
702 stem cell populations *in vivo*, we find stem cell-type-specific differences. Notably, we found that
703 heterochromatic domains have locations that replicate earlier in GSC-like cells compared to CySC-like
704 and cultured cells, suggesting fundamental differences in heterochromatin regulation between germline
705 and somatic lineages.

706 A notable distinction between GSC-like and CySC-like cells is the abundance and broad
707 distribution of H3K27me3, a facultative heterochromatin mark established by the Polycomb-complex that
708 is typically associated with gene silencing and cell identity maintenance (Wiles and Selker 2016). In
709 CySC-like cells, H3K27me3 enrichment displayed a canonical pattern, where lower expressed genes
710 exhibited higher levels of this mark. It is possible that the extensive presence of H3K27me3 in CySCs
711 serves to restrict gene expression and reinforce the cyst cell fate. This may reflect their unique division
712 pattern - a single asymmetric cell division produces a self-renewed CySC daughter and a differentiating
713 cyst cell (CC) daughter. Along with another CC and without further cell division, the two CCs forms a
714 unit to support the differentiating germline for the duration of spermatogenesis, which takes
715 approximately 10 days at 25°C (Fabian and Brill 2012; Zoller and Schulz 2012). In this way, the
716 restriction of H3K27me3 reinforces somatic cyst cell identity and reflects an early commitment to a
717 terminal cell fate in this lineage.
718 In the female *Drosophila* germline lineage, constitutive pericentromeric heterochromatin
719 formation begins in the germline cyst progeny of GSCs, while Polycomb-mediated gene repression
720 suppresses male germline and somatic lineage genes (Pang et al. 2023). Further genomic and genetic
721 studies in the female germline revealed a surprisingly open chromatin environment in GSCs (Deluca et al.
722 2020). In female GSCs, there is a non-canonical H3K27me3 distribution, whereby this modification is
723 present in low amounts on active chromatin as well as on transcriptionally inactive loci (Deluca et al.
724 2020). We observe a similar trend in male GSCs whereby there is a roughly equivalent amount of
725 H3K27me3 across all gene categories regardless of their expression levels, suggesting that low, non-gene-
726 specific H3K27me3 may be an epigenetic feature of undifferentiated cells (Deluca et al. 2020).
727 Our finding that pericentromeric heterochromatin replicates early in GSCs aligns with published
728 studies in mouse embryonic stem cells, which demonstrate that pericentromeric heterochromatin shifts
729 from early to late replicating upon differentiation (Rausch et al. 2020). An additional key insight from that
730 work was the significantly higher proportion of unidirectional replication forks in mouse embryonic stem

731 cells (around 13%) compared to differentiated human cell lines (around 7%) (Rausch et al. 2020).
732 Similarly, a notable 35% of replication forks are unidirectional in the *Drosophila* germline compared to
733 ~17% in somatic cells (Wooten et al. 2019). This trend of higher unidirectional fork usage in less
734 differentiated states appears to be consistent across species.

735 In addition, activated origins in mouse embryonic stem cells are spaced at half the distance of that
736 in somatic cells, indicating that stem cells activate twice as many origins of replication (Rausch et al.
737 2020). Although future studies in *Drosophila* germline stem cells are needed to determine whether the
738 higher incidence of unidirectional replication forks corresponds to reduced inter-origin distances, the
739 elevated expression of DNA replication-related genes supports a model of increased replication demand
740 in these cells. Furthermore, globally reduced H3K27me3 may create a permissive chromatin environment
741 for increased unidirectional replication forks and origin activation. This supports the idea that unique
742 replication dynamics may underlie specialized genome regulation in stem cells. These features may
743 reflect developmental constraints on cell fate transitions or represent lineage-specific adaptations.
744 Notably, our observation of early replication at the ribosomal DNA (rDNA) locus, which is well-
745 established to replicate unidirectionally (Gerber et al. 1997), supports this model.

746 In parallel with these replication features, the *Drosophila* germline exhibits a striking pattern of
747 histone inheritance: old histones, synthesized in the previous cell cycle, are preferentially incorporated
748 onto the leading strand during DNA synthesis, while newly synthesized histones are biased toward the
749 lagging the strand (Wooten et al. 2019). This produces sister chromatids asymmetrically enriched with
750 different histone populations, which are then differentially segregated such that old-histone-enriched
751 sisters are retained in the self-renewing stem cell while the gonialblast inherits new-histone-enriched
752 sister chromatids (Tran et al. 2012). The elevated incidence of unidirectional forks may facilitate this
753 global histone asymmetry, enabling extended tracts of old or new histone deposition. Notably, early
754 unidirectional replication at the rDNA locus may serve as a key organizing feature in establishing such
755 genome-wide histone asymmetry in *Drosophila* GSCs. Together, these mechanisms may ensure the

756 faithful inheritance of epigenetic information during stem cell asymmetric cell division, by preserving
757 stem cell identity while allowing for flexible gene regulation in the differentiating progeny.
758 Recent studies in budding yeast suggest rDNA replication can influence anaphase onset, with
759 extended rDNA repeats potentially delaying replication to prevent premature chromosome segregation
760 (Kwan et al. 2023). In *Drosophila* germ cells, rDNA copy number decreases during aging, yet can be
761 restored in GSCs of young males in the following generation (Lu et al. 2018). The mechanism underlying
762 this generational recovery remains unknown. Early replication of the rDNA locus in GSC-like cells may
763 present a strategy for maintaining or restoring rDNA copy number for subsequent generations, which
764 might be less efficient if replication occurred later in S-phase. This preferential timing of rDNA
765 replication in stem cells may therefore serve a critical function in long-term germline and organismal
766 maintenance. Future studies can further clarify how this unique replication feature contributes to stem cell
767 genome integrity and lineage continuity.
768
769 **Chromosome-specific replication timing profiles and regulatory roles for non-coding RNAs**
770 Our work also uncovered chromosome-specific differences in histone modification enrichment
771 and replication timing, particularly for the X and 4th Chromosomes. Historically, the 4th chromosome is
772 considered as entirely heterochromatic, based on its high percentage of repetitive elements, lack of
773 recombination, and high levels of HP1 and H3K9me2 (Haynes et al. 2007; Riddle et al. 2008). Previous
774 studies in *Drosophila* suggest that the 4th chromosome largely shares chromatin features with pericentric
775 heterochromatin, although it also contains some euchromatic regions (Riddle et al. 2008; Haynes et al.
776 2007). Unexpectedly, we found this chromosome replicates earlier in GSC-like cells than in CySC-like
777 and other somatic cells. The biological significance of its early replication remains unclear. One
778 possibility is that early replication of the 4th chromosome in GSC-like cells reflects a regulatory role for
779 this chromosome that is specific to GSCs. Alternatively, this timing pattern may result from its similar
780 chromatin composition to pericentromeric heterochromatin, rather than a regulatory process. Further

781 studies are required to determine whether the early replication timing of the 4th chromosome in the male
782 GSC-like cells has biological relevance for genome stability or gene regulation.
783 Our work also identified regulatory non-coding RNAs enriched in genomic regions undergoing
784 coordinated changes in replication timing, chromatin state, and transcriptional activity, including those
785 within the 4th chromosome and the pericentromeric heterochromatin of Chromosome 2. Although
786 regulatory RNAs have historically been challenging to study, comparative analyses like ours allow for
787 identifying them for further functional characterization. Previous studies have shown that approximately
788 30% of *Drosophila* testis-specific long non-coding RNAs (lncRNAs) function during late spermiogenesis,
789 regulating nuclear condensation, sperm individualization, and other terminal differentiation processes
790 (Wen et al. 2016). However, their roles in early spermatogenesis, particularly in stem cells, remains
791 poorly understood. Notably, several non-coding RNAs located in differentially replicating regions were
792 also differentially expressed in GSC-like cells. It is possible that their transcription promotes an open
793 chromatin environment, facilitating origin recognition and early replication. Rather than acting through
794 their protein products, these transcripts may instead provide a mechanism to achieve the elevated origin
795 activation observed in stem cells. An intriguing possibility is that the inherent directionality of RNA
796 transcription could influence coordinated replication origin activation and directionality of fork
797 progression. By identifying candidate non-coding RNAs situated in regions with stem cell-specific
798 replication timing, our findings provide a foundation for investigating their functions in stem cell
799 maintenance, differentiation, and genome regulation.
800 Our findings reinforce our knowledge that the male X chromosome maintains a unique chromatin
801 environment that supports its well-established role in dosage compensation in somatic cells (Samata and
802 Akhtar 2018). In CySC-like cells, the X chromosome displays signatures of open chromatin consistent
803 with this function, including H3K4me3 (this study), H4K16ac, and early replication (Rastelli and Kuroda
804 1998; Anderson et al. 2023). Interestingly, in GSC-like cells where dosage compensation is inactive, the
805 X chromosome still replicates early. Previous studies in male and female *Drosophila* cell lines (Clone8

806 and Kc cells, respectively) similarly demonstrated that the X chromosome replicates entirely during early
807 S-phase, even at genes enriched with H4K16ac that did not undergo gene dosage compensation
808 (Schwaiger et al. 2009). Based on these findings, H4K16ac was proposed to associate more with early
809 replication than with transcriptional activity. However, our data show that early replication persists in
810 male GSCs even in the absence of this mark. Together, these results suggest that replication timing of the
811 X chromosomes is more closely linked to chromatin accessibility than to either transcriptional output or
812 H4K16ac enrichment (Schwaiger et al. 2009). Our findings support a model that the X chromosome's
813 early replication timing in the male germline is maintained through dosage compensation-independent
814 mechanisms, likely involving chromatin accessibility features.
815 In conclusion, by integrating transcriptional, chromatin, and replication timing analyses, our data
816 provide a comprehensive framework for understanding the unique genomic regulation of male germline
817 and somatic stem cells. Altogether, our findings reveal distinct fundamental divergence in regulatory
818 mechanisms, including heterochromatin replication timing programs, chromosome-specific epigenetic
819 differences, and potentially novel roles for lncRNAs in stem cell genome organization. These findings
820 provide new insights into stem-cell specific genome regulation and establish a foundation for future
821 studies on the epigenetic and replication dynamics of *in vivo* adult stem cells.
822
823 **METHODS**
824 Fly strains husbandry
825 Fly stocks and crosses were maintained on standard molasses food using the Bloomington recipe, reared
826 at 25°C. See Supplemental Materials for a complete list of fly stocks and their sources used in this study.
827 For fly lines used in all genomic experiments, standard fly genetics was used to introduce the switchable
828 dual-tagged histone cassettes into a genetic background containing a *UAS-upd* transgene, which was
829 kindly provided by Dr. Stephen Dinardo (University of Pennsylvania, USA). After eclosion, male
830 progenies with the appropriate genotype were aged for ~2 days before dissection and sample preparation

831 for either scRNA-seq, ChIC-ChIP-seq, or Repli-seq. For males used in the immunostaining experiments,
832 young males (1-5 days old) were collected from fresh flips of the stock.
833
834 Immunostaining (whole mount method)
835 Testes from 1-5 day old adult male flies were dissected in pre-warmed Schneider's media (Gibco
836 #21720-024) before fixing at room temperature in 4% formaldehyde (Cell Signaling #12606) in 1X PBS
837 containing 0.1% TritonX-100 (PBST; Quality Biologicals #119-069-131, Fisher #BP151) for 10 minutes
838 with rotation. Following fixation, samples were briefly rinsed twice in PBST followed by three 5-minute
839 PBST washes with rotation. Primary antibodies were prepared at appropriate dilutions in 5% normal goat
840 serum (NGS) (Cell Signaling #5425) in PBST and the testes incubated for three nights at 4°C. See
841 Supplemental Materials for a complete list of all antibodies used in this study.
842 After incubation, the primary antibody solution was removed and two brief PBST rinses
843 performed. The testes were next washed three times in PBST for 5-minutes each with rotation. Secondary
844 antibodies were diluted 1:1000 in 5%NGS in PBST and testes were incubated for two nights, shielded
845 from light at 4°C with rotation. At the end of the secondary incubation, Hoechst 33342 (Invitrogen
846 #H3570) solution was added to secondary solution at a dilution of 1:1000, and the testes rotated for 10
847 minutes. Afterward, the secondary antibody solution was removed and two brief PBST washes
848 performed. This was followed with three 5-minute PBST washes with rotation. Testes were then mounted
849 on a Superfrost Plus microscope slide (Fisher #1255015) using Vectasheild PLUS Antifade mounting
850 media (Vector Laboratories #H-1900). Imaging was performed on a Leica Stellaris5 Confocal with
851 405nm and white light laser.
852
853 Immunostaining (squash method)
854 *Upd*-tumor testes (*nos*>Upd) were dissected in warm Schneider's media. Approximately 5-6
855 individual testes were transferred to a microscope slide, and excess media removed. A 10uL drop of 1X

856 PBS was placed over the tumors before gently rupturing each tissue. A coverslip was next placed over the
857 ruptured testes to squash and spread the tissue. Slides were next flash frozen in liquid nitrogen for 2
858 minutes. Following freezing, the coverslips were removed with a razor blade and the slides immediately
859 placed in a coplin jar containing pre-chilled 95% ethanol (Fisher #BP2818). The jar was incubated at -
860 20°C for 10minutes. Following incubation, slides were removed from the jar. Excess ethanol was wiped
861 from the slide and the tissue outlined with a hydrophobic pen. A 50uL drop of 1% formaldehyde in 1X
862 PBS with 0.1% TritonX-100 (PBST) was placed on top of the tissue and covered with parafilm. The
863 slides were fixed in a humid chamber for 3 minutes, after which the slides were rinsed quickly in PBST
864 thrice.
865 Following washes, the slides were permeabilized in a PBST/DOC (1X PBS, 0.5% TritonX-100,
866 0.5% Sodium Deoxycholate) solution for 15 minutes at room temperature with gentle rocking. This was
867 repeated once for a total of 2 washes. After the second wash, a PBST wash was performed for 10 minutes.
868 Slides were next incubated with 20uL of appropriate primary antibodies diluted in 3% BSA in
869 PBST (Cell Signaling #9998) overnight in a humid chamber at 4°C. The next day, slides were washed
870 three times, 10 minutes each in PBST. Secondary antibodies were diluted 1:1000 in 5%NGS/PBST and
871 incubated overnight at 4°C in a humid chamber. The following day, slides were washed three times for 5
872 minutes each with gentle rocking. Hoechst 33342 was added to the final wash at a concentration of
873 0.5ug/mL in PBST. Finally, slides were mounted in vectashield and imaged on a Leica Stellaris5
874 Confocal with 405nm and white light laser.
875
876 <u>Intensity quantification for GFP-tagged fusion proteins</u>
877 The image analysis software Imaris was used to quantify the amount of GFP signal in
878 endogenously tagged Orc1, MCM2, and Caf1-p105 containing GSCs and CySCs. First, machine learning
879 was used to identify the nuclei of individual GSCs and CySCs as surfaces within the tissue. GSCs were
880 determined by their adjacent proximity to the stem cell niche and presence of the spectrosome, whereas

881 CySCs were identified using the expression of Zfh1. For each nucleus (surface), the intensity sums of
882 both Hoechst signal and GFP were calculated using the statistics option. To normalize GFP intensity to
883 DNA content for an individual nucleus, the intensity sum of GFP signal was divided by the intensity sum
884 of Hoechst signal. The $\log_2$ of this value was calculated and plotted, followed by a two-tailed Mann-
885 Whitney test in GraphPad.
886
887 <u>10X Genomics single cell RNA-sequencing</u>
888 *Preparation of single cell solution from Upd-tumor testes for single cell RNA-sequencing*
889 Tumor testes from either *nos>Upd, H3-GFP* or *tj>Upd, H3-GFP* were dissected in warm,
890 filtered Schneider's media supplemented with 10% FBS (Thermo Fisher #16140071). After collecting
891 ~18-20 pairs, testes were transferred to ~250uL of digestion buffer (TrypLE Express, Gibco #12605-101
892 with 2mg/ml collagenase, Sigma #C9407). The tissues were incubated in a 37°C waterbath, with
893 intermittent agitation to dissociate the tissue to completion. Next, the cells were filtered first through a
894 40um filter (Corning #352340), followed by a 10um filter (pluriSelect, #43-10010-40). Digestion buffer
895 was removed after pelleting the cell suspension by centrifugation at 1,200 rpm for at least 7 minutes. The
896 cell pellet was washed twice in 500uL HBSS (Gibco #14170-112), before resuspending in 20uL HBSS.
897 Cell count and viability was obtained at this point using Trypan blue (Gibco #15250-061), before
898 proceeding to the 10X Genomics single cell RNA-sequencing protocol.
899
900 *10X genomics sequencing*
901 Single cell RNA-sequencing libraries were generated following the manufacturer's guidelines for
902 the Chromium Next GEM Single Cell 3'Kit v3.1 (#1000268). Final libraries were pooled to 4nM and
903 sequenced on the Illumina NovaSeq sequencer, using the S1 100 cycle within the Johns Hopkins
904 Genomics Core.
905

906 Cell-specific sequential Chromatin Immunocleavage – Chromatin Immunoprecipitation
907 (ChIC-ChIP)
908 *Prepare magnetic beads for immunoprecipitation:*
909 The evening prior to dissection, immunoprecipitation (IP) beads were prepared for blocking. Per
910 desired IP, 10uL of magnetic protein-G beads (Invitrogen, #10003D) were incubated in 200uL antibody
911 binding buffer containing 5% BSA (10mM Tris-HCl, 1mM EDTA, pH7.5, 150mM NaCl, 0.1% TritonX-
912 100). The beads were rotated end-over-end overnight at 4°C.
913 The next morning, blocked beads were prepared for immunoprecipitation by washing the beads
914 twice with 300uL antibody binding buffer containing 0.05% BSA. To the washed beads, 1ug of ChIP
915 antibody in 200uL 0.1% BSA in antibody binding buffer was added per IP. The beads were incubated on
916 rotating wheel at 4°C for 6 hours while preparing tissue for dissection and chromatin extraction.
917 Antibodies for the ChIC and ChIP experiments can be found in Supplemental Materials.
918
919 *Preparation of single cell solution from Upd-tumor testes for ChIC-ChIP-seq*
920 Approximately 18-20 tumor testes from either *nos>Upd, H3-GFP* or *tj>Upd, H3-GFP* were
921 dissected and dissociated as outlined for single cell RNA-sequencing. However, instead of resuspending
922 the cell pellet in HBSS following removal of digestion buffer, cells to be used for ChIC were resuspended
923 in 400uL of Schneider's media with FBS. To this, 16% formaldehyde (Thermo #28906) was added at
924 1/15 the total volume to obtain a 1% final concentration. The cells were fixed by end-over-end rotation
925 for 5 minutes at room temperature. The fixation reaction was quenched by the addition of 1/10 volume of
926 1.25M glycine and additional incubation for 5 minutes. Following this, the cells were pelleted by
927 centrifugation for 7 minutes at 1,320rpm at 4°C. Afterward, the fix solution was removed and the cell
928 pellet washed with 500uL cold, sterile 1X PBS. The 1X PBS wash was repeated once, after which all
929 liquid was removed. The cells were resuspended in 20uL 1X PBS to obtain a cell count.
930

931 *Preparation of ProteinA-MNase-Antibody complex and chromatin*

932 For a single IP, 4uL antibody binding buffer, 1uL anti-GFP antibody, and 3uL ProteinA (PA)-
933 MNase were prepared. The cocktail was incubated on ice for 30 minutes. Meanwhile, chromatin from
934 formaldehyde fixed cells was prepared for binding by the PA-MNase-Ab complex by treating cells with
935 500uL RIPA buffer (10mM Tris-HCl, 1mM EDTA, 150mM NaCl, 0.2% SDS, 0.1% Sodium
936 Deoxycholate, 1% TritonX-100) for 10 minutes at room temperature. Following incubation, cells were
937 pelleted by a 5-minute centrifugation at 3,000 rpm and the RIPA was removed. The cell pellet was
938 washed once with 500uL antibody binding buffer, before resuspending in 100uL antibody binding buffer.

939 *Perform PA-MNase-Ab incubation, cell washes, and prepare antibody-coupled beads*

940 The 100uL of resuspended fixed cells were added to the microcentrifuge tube containing 8uL of
941 prepared PA-MNase-Ab complex. After gentle mixing, the tube was incubated on ice for 60 minutes.
942 Following the incubation, excess PA-MNase-Ab was removed through a series of washes. First, antibody
943 binding buffer was removed after spinning cells at 3,000rpm for 5 minutes at 4°C. The cell pellet was
944 then washed with 500uL high salt wash (10mM Tris, 1mM EDTA, 400mM NaCl, 1% TritonX-100),
945 incubating for 1 minute at room temperature before spinning. The high salt wash was repeated for a total
946 of three washes. Next, the cell pellet was washed with 200uL rinsing buffer (10mM Tris pH 7.5, 10mM
947 NaCl, 0.1% TritonX-100), immediately spinning down for 5 minutes. After removing the rinsing buffer,
948 the pellet was resuspended in 40uL RSB (20mM Tris pH 7.5, 10mM NaCl, 2mM $CaCl_2$, 0.1% TritonX-
949 100).

950

951 The blocked, antibody-coupled beads were washed simultaneously with the cell pellets. First,
952 beads were collected and unbound antibody removed. The beads were then washed a total of three times
953 with 500uL antibody binding buffer, rotating end-over-end for 5 minutes at 4°C. After the last wash,
954 beads were resuspended in 10uL of rinsing buffer per IP.

955

956 *MNase digestion:*

957 To perform the MNase digestion, the 40uL of cells in RSB were placed in a 37°C water bath for 3
958 minutes. Afterward, the reaction was stopped by adding 80uL of stop buffer (20mM Tris pH8, 10mM
959 EGTA, 20mM NaCl, 0.2% SDS). Following digestion, the insoluble chromatin was spun down at 13,000
960 rpm for 10 minutes at 4°C. The soluble chromatin (~120uL) was moved to a new tube, where protease
961 inhibitor (Cell signaling #5872) was added to final concentration of 1X. Next, 5% of the soluble
962 chromatin was reserved as the input fraction.

963

964 *Chromatin IP, Bead preparation, and washing:*

965 For each IP, 10uL of antibody-bound beads was added and the mixture incubated overnight at
966 4°C with end-over-end rotation.

967 The next day, the IP solution was cleared on a magnetic rack and the supernatant removed. The
968 protein bound beads were then washed three times by adding 500uL of low salt IP wash (20mM Tris-HCl
969 pH8, 150mM NaCl, 0.1% SDS, 2mM EDTA, 1% TritonX-100) to the beads and incubating at 4°C with
970 end-over-end rotation for 5 minutes. After the three low salt IP washes, the beads were washed once with
971 a high salt IP wash (20mM Tris-HCl pH8, 500mM NaCl, 0.1% SDS, 2mM EDTA, 1% TritonX-100).
972 Following the washes, DNA was eluted from the beads by adding 150uL of ChIP elution buffer (1% SDS,
973 0.1M Sodium Bicarbonate) to each IP and incubating for 30 minutes in a 65°C thermomixer with
974 vortexing at 1,200 rpm. After separating the beads on a magnetic rack, eluted DNA was transferred to a
975 clean DNA LoBind tube. At this point, 150uL ChIP elution buffer was added to the 5% input samples.
976 For all input and IP samples, fixed crosslinks were reversed by adding 5M NaCl (final 187mM NaCl) and
977 40ug Proteinase K (NEB #P8107S) and incubating overnight at 65°C.

978

979 *Purify immunoprecipitated DNA and Input samples:*

980 The DNA from Input and IP samples was purified using a Qiagen MinElute Reaction Cleanup Kit
981 (Qiagen #28204) with the following modification. First, 5 volumes of ERC (750uL) were added to each
982 sample, mixed well, and transferred to a Qiagen MinElute spin column. The sample was centrifuged at
983 13,000 rpm for 1 min, and flow through discarded. Next, the column was washed with 700uL PE and
984 centrifuged once again. After transferring the column to a new DNA LoBind tube, 10uL 10mM Tris,
985 pH8.0 was added to the spin column. Following a 1 minute room temperature incubation, the sample was
986 eluted by spinning at 13,000 rpm for 1 minute. Eluted DNA was quantified using the Qubit dsDNA high
987 sensitivity kit (Thermo #Q33231) before proceeding to library preparation using the instructions included
988 in NEBNext Ultra II DNA Library Prep Kit for Ilumina (NEB #7645). Finalized libraries were sequenced
989 on a NovaSeq sequencer at the National Institutes of Health.
990
991 Nuclei isolation, sorting, and DNA purification for Repli-seq
992 Tumor testes from either *nos>Upd, H3-GFP* or *tj>Upd, H3-GFP* were dissected in warm,
993 filtered Schneider's media, to obtain 100-200 total pairs. For downstream sample preparation, only
994 tumors shaped like a large ball were collected. The testes were then incubated with rotation for 15
995 minutes in Schneider's media containing 100uM BrdU.
996 After the BrdU incubation, the Schneiders' media was removed, and the testes were washed once
997 with 1X PBS. Nuclei isolation was performed following a published protocol, with experimental
998 considerations (McLaughlin et al. 2022). The dissected testes were first resuspended in 1mL
999 homogenization buffer (250mM Sucrose, 10mM 1M Tris, pH 8.0, 25mM KCl, 5mM $MgCl_2$, 0.1% Triton
1000 X-100, and freshly added 1X protease inhibitors (Cell signaling #5872) and 0.1mM DTT – omit RNase
1001 Inhibitors) and transferred to a 1mL dounce homogenizer. The tissues were dissociated with 20 strokes of
1002 the loose pestle (A), followed by 40 strokes of the tight pestle (B) before pipetting gently ~20 times with
1003 wide bore tips. The nuclei were next filtered through a 40um cell strainer, (Corning #352340), followed
1004 by a 10um filter (pluriSelect, #43-10010-40) into a fresh tube. The nuclei were then spun down at 1,000xg

1005 for 10 minutes at 4°C to remove the homogenization buffer. Following pelleting, nuclei were resuspended
1006 in 400uL resuspension buffer (1X PBS, 0.5%BSA), containing final concentrations of 250ug/mL RNaseA
1007 (NEB #T3018L) and 50 ug/mL propidium iodide (Sigma #P4864). The nuclei were incubated on ice, in
1008 the dark for at least 30 minutes before proceeding to the cell sorter.

1009 Nuclei were sorted on a MoFlo XDP cell sorter (Beckman Coulter) by gating for GFP+ cells and
1010 dividing S-phase cells into 4 equal fractions corresponding to the Early, Early-Mid, Late-Mid and Late S-
1011 phase timepoints. Total number and volume of sorted nuclei was noted.

1012 Following the sort, components of SDS-PK were added individually to keep volume low,
1013 allowing for accurate pipetting of 10,000-20,000 nuclei per fraction (SDS-PK: 50mM Tris-HCl pH8,
1014 10mM EDTA, 1M NaCl, 0.5% SDS, 0.2mg/mL proteinase K). Samples were incubated at 56°C for 2
1015 hours. Afterward, nuclei were divided into separate tubes containing ~10,000 nuclei each. The DNA from
1016 ~10,000 nuclei was then purified following the protocol for cell suspensions and proteinase K digested
1017 samples from the Zymo Quick DNA MicroPrep kit (Zymo #D3021). Purified DNA was stored at -20°C.
1018

1019 Repli-sequencing

1020 *Fragmentation and library construction:*

1021 The preparation of samples for Repli-seq mostly followed published protocols, with some
1022 modifications(Marchal et al. 2018). For each of the four S-phase fractions, two tubes of 10,000 nuclei
1023 were processed as technical replicates for Repli-seq. After thawing, the volume of DNA was brought to
1024 100uL with UltraPure water (UPW, Invitrogen, #10977-015) and samples transferred to Diagenode
1025 0.5mL microtubes for the Bioruptor Plus (Diagenode #C30010013). The DNA was then sonicated in a
1026 cold waterbath for 1 hour on low intensity (30sec on, 90 sec off), checking every 15 minutes for ice level.
1027 The sheared DNA was then concentrated to 15uL of UPW using the DNA Clean & Concentrator-5 kit
1028 (Zymo Research #D4003) following the instructions for DNA fragments.

1029 The concentrated DNA was then brought to a volume of 50uL with UPW prior to starting library
1030 construction using the NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB #7645). The NEBNext
1031 End Prep (step 1) and Adapter Ligation (step 2), were performed following the kit instructions. After
1032 incubation with the USER enzyme, the DNA was purified again using the DNA Clean and Concentrator-5
1033 kit following the instructions for DNA fragments and eluting in 50uL UPW.
1034
1035 *Immunoprecipitation of BrdU containing DNA:*
1036 To the 50uL of clean adapter ligated DNA, 450uL of TE (10mM Tris pH8, 1mM EDTA) was
1037 added. The DNA was then denatured by heating at 95°C for 5 minutes, followed by cooling on ice for 2
1038 minutes. The denatured DNA was then transferred to a tube prepared with 60uL 10X IP buffer (100mM
1039 Sodium Phosphate, pH 7, 1.4M NaCl, 0.5% TritonX-100). To this, 40uL of 12.5ug/mL anti-BrdU
1040 antibody (BD Biosciences #555627) was added and the sample rocked at room temperature for 20
1041 minutes. Next, 20ug of rabbit anti-mouse IgG antibody (Sigma #M7023) was added and the sample
1042 rocked again at room temperature for 20 minutes. Following this, tubes were centrifuged at 16,000xg for
1043 5 minutes at 4°C to pellet the BrdU containing DNA. After removing the supernatant, the DNA pellet was
1044 washed with 750uL pre-chilled 1X IP buffer (10mM Sodium Phosphate, pH 7, 140mM NaCl, 0.05%
1045 TritonX-100). Centrifugation was repeated and the supernatant removed. The DNA pellet was then
1046 resuspended in 200uL digestion buffer containing 0.25mg/mL proteinase K (50mM Tris-HCl, 10mM
1047 EDTA, 0.5% SDS) and incubated overnight at 37°C in an air incubator. The next day, 1.25uL of
1048 20mg/mL proteinase K was added to each tube before incubating for 60 minutes at 56°C. Next, the
1049 immunoprecipitated DNA was purified using the DNA Clean & Concentrator-5 kit following the
1050 instructions for DNA fragments and eluting in 16uL UPW.
1051
1052 *Measuring BrdU-IP efficiency and library amplification:*

1053 Using a serial dilution of known DNA concentrations, the amount of BrdU immunoprecipitated
1054 DNA was quantified using a standard qPCR protocol with included melt curve. The calculated total
1055 amount of adapter ligated DNA was then used to determine the number of PCR cycles required for library
1056 amplification. PCR enrichment of DNA was performed following the instructions for the NEBNext Ultra
1057 II DNA Library Prep Kit for Illumina. The final libraries were pooled to 4nM and sequenced on a
1058 NovaSeq X Plus.
1059
1060 *Bioinformatics*
1061 Our genomic sources for all analyses are: the FlyBase dm6 chromosomes (release 6.47) (Öztürk-
1062 Çolak et al. 2024); the sequence for H3-GFP added to the genome of the flies of both genotypes; an
1063 interval of FlyBase chromosome 2L which serves as our reference for the histone unit tandem repeat;
1064 and/or the FlyBase transposon sequence set (*Drosophila* genus). Single-cell samples are quantified using
1065 FlyBase release 6.47, concatenated with the H3-GFP sequence, and using cellranger mkref and cellranger
1066 count (version 7.0.0). ChIC-seq and Repli-Seq samples use FlyBase release 6.47, passed to the bowtie2-
1067 build command. However, repetitive sequences analysis proceeds differently, using RepeatMasker
1068 (version 4.1.6) to find occurrences of the *Drosophila* transposons, or the histone repeat unit, on the
1069 chromosomes, masking occurrences. Reads were mapped using Bowtie 2 (version 2.4.5). In the masked-
1070 mapped ChIC-seq samples, n = 119 transposons, as well as each of the histone genes, were covered by
1071 uniquely-mapped fragments (MAPQ >= 20).
1072 Regions of interest in the dm6 reference are as follows. For rDNA, we created a window that
1073 bounds the genes with a Feature Type of rRNA gene in Flybase, with pseudogenes generally falling
1074 outside of this window, then rounded the size of the window to 1 significant figure (rDNA:
1075 36901..76900). For chromosome 2 and 3, some of the centromere-adjacent chromatin was previously
1076 classified as classic pericentric heterochromatin, with coverage becoming sparse towards the centromere
1077 (Filion et al. 2010). We extended the closest-to-centromere segments of heterochromatin to the Crick end

1078 of the dm6 2L and 3L, and to the Watson end of the dm6 2R and 3R, calling these regions: 2LC, 2RC,
1079 3LC, 3RC. Where these regions are plotted, the mostly euchromatic chromosome arms not in this region
1080 have their values given as the 2L, 2R, 3L, and 3R values.
1081
1082 *Single-cell RNA-seq analysis*
1083 The single-cell RNA-seq experiment quantifies the GSC- and CySC-like cell transcriptomes and
1084 supports our hypothesis that these tumor cell populations are largely homogenous. Analysis is
1085 downstream of Cell Ranger feature matrix quantification (v7.0.0) with default parameters. First, we
1086 reported Log Normalized gene expression, which rescales the sum of transcripts from the cell's total UMI
1087 counts to a constant (Seurat's default parameter of 10,000) and applies shifted logarithm (Ahlmann-Eltze
1088 and Huber 2023). On the other hand, cells are plotted (in UMAP) and identified (as clusters) using
1089 Seurat's SCTransform integrated assay, where the cell embedding is computed from residuals from a
1090 Negative Binomial model, followed by fitting a nonlinear model to remove much of the within-batch
1091 effect (Hafemeister and Satija 2019; Butler et al. 2018).
1092 Before quantifying transcripts per-cluster, we removed the following effects. To shrink a
1093 correlation between log(percent of grand total UMIs in a particular gene) and log(gene CDS length), we
1094 counted only the UMIs tagged as having an exonic alignment. We also limited UMIs to a particular
1095 compatible isoform (TX) per gene to further shrink spurious alignments to full isoforms. Finally, we
1096 identified ambient RNA contamination from sparse expression of *Nanos* in the CySC-like cluster, and
1097 *Traffic Jam* and *roX* transcripts in the GSC-like cluster. Therefore, we applied decontX inference to factor
1098 our TX matrix into real-valued nonnegative ambient RNA contamination and true expression estimate
1099 matrices (Yang et al. 2020). However, to retain more transcripts at single-cell resolution, the figures plot
1100 Log Normalized decontX true expression estimates computed from the original Cell Ranger feature
1101 matrix with default parameters.

1102 The key transcript values are drawn from a log-quantification model for secondary analyses, and
1103 additionally (for enrichment analyses only) from a log-fold-change Negative Binomial regression model.
1104 Log-CPM is computed via linear model on pseudobulked, then log-transformed, real-valued decontX
1105 estimates (Ritchie et al. 2015). Log-CPM coefficients (per cluster) require a shift to be added, such that
1106 the sum of the exponentiated coefficients is 1 million. Genes are cut at 5 CPM to reveal "on" or "off"
1107 genes, chosen to fit a bimodal violin plot where genes start piling up at ~1 CPM, but the "on" genes
1108 produce a mode of positively quantified genes that is between 10 and 100 CPM. On the other hand, we
1109 did not present CPM ratios for differential gene expression. Instead, our transcriptome log-fold-change is
1110 a Bayesian regression coefficient estimate, with heavy-tailed prior (Zhu et al. 2019). To apply the benefits
1111 of decontX for estimating the correct log-fold-change, we added decontX's contamination score, crossed
1112 with cell type, as covariates. Our GSC-vs-CySC coefficient DEG significance threshold is an *s*-value
1113 (posterior two-tailed test transformed to have FDR-like power) less than 10^{-4}. DEGs are also filtered by
1114 log effect size (L2FC) at least 1.5, producing GSC and CySC gene lists to which we apply GO term
1115 enrichment (Thomas et al. 2022).
1116
1117 *ChIC-ChIP-seq analysis*
1118 ChIC-seq reveals nucleosome occupancy, H3K4me3 enrichment, H3K27me3 enrichment, and
1119 H3K9me3 enrichment, in the bulk GSC-like cell and CySC-like cell genomic DNA libraries. In our
1120 ChIC-seq regression, we estimate the intercept (H3 input log-enrichment), control vs treatment (log-fold
1121 change in the mark antibody ChIP treatment), and batch effect: depending on the biological library that
1122 the input and IP come from, there is a random batch effect added to the intercept. This design was chosen
1123 as each experiment has replicates, and the replicate input and ChIP are paired. The variable is the count of
1124 fragments tested against a fixed-sized sliding window on the reference.
1125 ChIC-seq analysis proceeds to produce fine versions of the H3 control and treatment
1126 quantification tracks (20 bp step), as well as broad H3 control and treatment quantifications (100 bp step)

1127 with Negative Binomial distribution parameter estimate (Ahlmann-Eltze and Huber 2021). Monosomes
1128 are visible as local maxima, producing the finished ChIC-seq tracks, after estimating feature density on
1129 the Negative Binomial estimates: smoothing using a Gaussian kernel with sigma = 40 bp (Boyle et al.
1130 2008). The log-fold enrichment of the IP "F-Seq" (Gaussian) track, relative to the H3 input track, reveals
1131 antibody binding to the monosome, with appropriate (20 bp) resolution to reveal the monosome midpoint
1132 location.
1133
1134 *Repli-seq analysis*
1135 Repli-seq identifies genomic loci that are enriched in a certain interval of S-phase (where the
1136 interval of S-phase is marked by molecular mass of DNA). Reading our ChIC-seq and Repli-seq ingestion
1137 of fragments, in all cases, we mark the single-end 100 bp read or the properly paired alignment (up to 500
1138 bp wide) at a 1 bp midpoint location. We counted Repli-seq fragments assigned to 1 kb bins with 1 kb
1139 step (non-overlapping), and found that consecutive bins were reasonably variable and could function as
1140 internal replicates of largely different BrdU-labeled DNA molecules (individual observations for our
1141 regression). In a 3 kb window, we used the observations of 3 bins of DNA in either 6 or 8 samples to infer
1142 a Bayesian logistic regression parameter (logit timing score), with a novel, vector-valued and Beta
1143 function-based link function. We applied a minimum-mean-square-error estimator (posterior expected
1144 value) which naturally shrinks towards zero (mean timing; prior expected value) in the case of poor
1145 coverage or poor model fit.
1146 The Bayesian Repli-seq leads to a principled argument to reject the null hypothesis of a static
1147 replication regime. The null hypothesis predicts a zero between-cell-type effect on the timing parameter.
1148 A full and reduced model for HTS data are designed to produce a likelihood ratio, which allows us to
1149 reject the null hypothesis because the additional model complexity significantly enhances the model fit
1150 (Love et al. 2014). In our case, we apply numerical integration to the full and reduced model likelihoods,
1151 summarized as a Bayes factor, which is a posterior odds ratio for the alternative hypothesis. We chose to

1152 reject the null hypothesis for differentially timed loci of size at minimum 20 kb. Then, we summarized the
1153 cell-type effect for the different track values within the locus, calling one of the cell types as earlier than
1154 the other.

1155 We analyzed Repli-seq using a 3 kb window, which is several times smaller than the recognizable
1156 replication regimes. The Repli-seq estimator is smoothed using LOESS to reveal peaks and valleys of
1157 early or late replication. The smooth Repli-seq permits z-scoring at every location on the window, to
1158 suggest dynamic timing in chromosomal regions using more continuous rather than discrete analysis. We
1159 identified 12 regions or loci described in the supplement, including the autosomal pericentric
1160 heterochromatin established from modENCODE epigenomics. As these are the regions that had different
1161 replication prioritization in four cell types (GSC-like, CySC-like, Kc167, and S2), the z-scores can be
1162 zero-centered and weighted (multiplied) so that the contribution to a Pearson R^2 is normalized to the
1163 number of observations in the region or locus (reciprocal). We simplified the presentation to use a vector
1164 of 12 mean timings, for the 4 cell types.

1165

1166 **DATA ACCESS**

1167 All raw and processed sequencing data generated in this study have been submitted to the NCBI
1168 Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/) under accession numbers:

1169 scRNA-seq, GSE289063 Reviewer token: olcnywycrtmvtkb

1170 ChIC-seq, GSE290500 Reviewer token: kfwhoumenledrgf

1171 Repli-seq, GSE291715 Reviewer token: qfczymekbjidax

1172 Other datasets used in this study are available under accession numbers:

1173 Repli-seq Kc167, GSE41349

1174 Repli-seq S2, GSE41350

1175 modENCODE Kc, H3K4me3: GSE45088; H3K27me3: GSE45083; H3K9me3: GSE27796

1176 modENCODE S2, H3K4me3: GSE20787; H3K27me3: GSE20781; H3K9me3: GSE20794

1177 modENCODE Orc2 ChIP meta-peaks: GSE27123

1178 The genomics analysis scripts have been deposited in GitHub: https://github.com/ringw/Upd-Germline-

1179 Genomics

1180

1181 **COMPETING INTEREST STATEMENT**

1182 The authors declare that there are no competing interests related to this work.

1183

1184 **ACKNOWLEDGEMENTS**

1185 We thank Dr. Hao Zhang from the Flow Cytometry and Cell Sorting Core Facility within the Department

1186 of Molecular Microbiology and Immunology at the Johns Hopkins Bloomberg School of Public Health

1187 for contributing to the cell sorting experiments required for Repli-seq. We are also grateful to Hanhvy Bui

1188 in the John Hopkins Integrated Imaging Center for her initial efforts in establishing a cell sorting protocol

1189 for our tissue. We also thank the Johns Hopkins Genetics Resources Core Facility for consultation and

1190 performing next generation sequencing. Many thanks to several members of Dr. Keji Zhao's group who

1191 helped with establishing the ChIC-seq protocol, aided with sequencing submission, and data storage. This

1192 includes, Drs. Lixia Pan, Wai Lim Ku, Guangzhe Ge, Kairong Cui, Qingsong Tang, and Meiyuan Ji.

1193

1194 **AUTHOR CONTRIBUTIONS**

1195 Project conceptualization: JAU, XC

1196 Experimentation: JAU

1197 Experimentation and validation of system: RG, WX

1198 Bioinformatic data analysis: DR, JMU

1199 Writing/Editing manuscript; JAU, XC

1200 Funding and experimental support: JAU, KZ, XC

REFERENCES

Ahlmann-Eltze C, Huber W. 2023. Comparison of transformations for single-cell RNA-seq data. *Nature Methods 2023 20:5* **20**: 665–672. https://www.nature.com/articles/s41592-023-01814-1 (Accessed March 11, 2025).

Ahlmann-Eltze C, Huber W. 2021. glmGamPoi: fitting Gamma-Poisson generalized linear models on single cell count data. *Bioinformatics* **36**: 5701–5702. https://pubmed.ncbi.nlm.nih.gov/33295604/ (Accessed March 11, 2025).

Amoyel M, Bach EA. 2012. Functions of the Drosophila JAK-STAT pathway: Lessons from stem cells. *JAKSTAT* **1**: 176. https://pmc.ncbi.nlm.nih.gov/articles/PMC3670241/ (Accessed October 23, 2024).

Anderson JT, Henikoff S, Ahmad K. 2023. Chromosome-specific maturation of the epigenome in the Drosophila male germline. *Elife* **12**.

Armstrong RL, Penke TJR, Strahl BD, Matera AG, McKay DJ, MacAlpine DM, Duronio RJ. 2018. Chromatin conformation and transcriptional activity are permissive regulators of DNA replication initiation in Drosophila. *Genome Res* **28**: 1688–1700. https://genome.cshlp.org/content/28/11/1688.full (Accessed November 18, 2024).

Boyle AP, Guinney J, Crawford GE, Furey TS. 2008. F-Seq: a feature density estimator for high-throughput sequence tags. *Bioinformatics* **24**: 2537–2538. https://pubmed.ncbi.nlm.nih.gov/18784119/ (Accessed March 11, 2025).

Brawley C, Matunis E. 2004. Regeneration of male germline stem cells by spermatogonial dedifferentiation in vivo. *Science (1979)* **304**: 1331–1334. https://www.science.org/doi/10.1126/science.1097676 (Accessed March 19, 2025).

Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. 2018. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nature Biotechnology 2018 36:5* **36**: 411–420. https://www.nature.com/articles/nbt.4096 (Accessed March 11, 2025).

Cheng J, Tiyaboonchai A, Yamashita YM, Hunt AJ. 2011. Asymmetric division of cyst stem cells in Drosophila testis is ensured by anaphase spindle repositioning. *Development* **138**: 831–837. https://pubmed.ncbi.nlm.nih.gov/21303845/ (Accessed March 19, 2025).

Cherbas L, Gong L. 2014. Cell lines. *Methods* **68**: 74–81.

Chereji R V., Ramachandran S, Bryson TD, Henikoff S. 2018. Precise genome-wide mapping of single nucleosomes and linkers in vivo. *Genome Biol* **19**: 1–20. https://genomebiology.biomedcentral.com/articles/10.1186/s13059-018-1398-0 (Accessed October 24, 2024).

Deluca SZ, Ghildiyal M, Pang LY, Spradling AC. 2020. Differentiating drosophila female germ cells initiate polycomb silencing by regulating PRC2-interacting proteins. *Elife* **9**: 1–33.

Eaton ML, Prinz JA, MacAlpine HK, Tretyakov G, Kharchenko P V., MacAlpine DM. 2011. Chromatin signatures of the Drosophila replication program. *Genome Res* **21**: 164–174. https://genome.cshlp.org/content/21/2/164.full (Accessed January 13, 2025).

Eun SH, Feng L, Cedeno-Rosario L, Gan Q, Wei G, Cui K, Zhao K, Chen X. 2017. Polycomb group gene E(z) is required for spermatogonial dedifferentiation in Drosophila adult testis. *J Mol Biol* **429**: 2030. https://pmc.ncbi.nlm.nih.gov/articles/PMC5516936/ (Accessed January 5, 2025).

Eun SH, Shi Z, Cui K, Zhao K, Chen X. 2014. A non-cell-autonomous role of E(z) to prevent germ cells from turning on a somatic cell marker. *Science* **343**: 1513. https://pmc.ncbi.nlm.nih.gov/articles/PMC4040133/ (Accessed January 5, 2025).

Fabian L, Brill JA. 2012. Drosophila spermiogenesis: Big things come from little packages. *Spermatogenesis* **2**: 197. https://pmc.ncbi.nlm.nih.gov/articles/PMC3469442/ (Accessed March 25, 2025).

Fabrizio JJ, Boyle M, DiNardo S. 2003. A somatic role for eyes absent (eya) and sine oculis (so) in Drosophila spermatocyte development. *Dev Biol* **258**: 117–128. https://pubmed.ncbi.nlm.nih.gov/12781687/ (Accessed October 23, 2024).

Filion GJ, van Bemmel JG, Braunschweig U, Talhout W, Kind J, Ward LD, Brugman W, de Castro IJ, Kerkhoven RM, Bussemaker HJ, et al. 2010. Systematic Protein Location Mapping Reveals Five Principal Chromatin Types in Drosophila Cells. *Cell* **143**: 212–224. https://www.cell.com/action/showFullText?pii=S0092867410010573 (Accessed March 11, 2025).

Formosa T, Winston F. 2020. The role of FACT in managing chromatin: disruption, assembly, or repair? *Nucleic Acids Res* **48**: 11929–11941. https://dx.doi.org/10.1093/nar/gkaa912 (Accessed October 23, 2024).

Fuller MT, Spradling AC. 2007. Male and female Drosophila germline stem cells: Two versions of immortality. *Science (1979)* **316**: 402–404. https://www.science.org/doi/10.1126/science.1140861 (Accessed March 19, 2025).

Gan Q, Schones DE, Ho Eun S, Wei G, Cui K, Zhao K, Chen X. 2010. Monovalent and unpoised status of most genes in undifferentiated cell-enriched Drosophila testis. *Genome Biol* **11**: 1–17. https://genomebiology.biomedcentral.com/articles/10.1186/gb-2010-11-4-r42 (Accessed September 17, 2024).

Gerber JK, Gögel E, Berger C, Wallisch M, Müller F, Grummt I, Grumm F. 1997. Termination of Mammalian rDNA Replication: Polar Arrest of Replication Fork Movement by Transcription Termination Factor TTF-I. *Cell* **90**: 559–567.

Gönczy P, Dinardo S. 1996. The germ line regulates somatic cyst cell proliferation and fate during Drosophila spermatogenesis. *Development* **122**: 2437–2447. https://dev.biologists.org/content/122/8/2437.long (Accessed February 23, 2020).

Hafemeister C, Satija R. 2019. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol* **20**: 1–15. https://genomebiology.biomedcentral.com/articles/10.1186/s13059-019-1874-1 (Accessed March 11, 2025).

Hafezi Y, Sruba SR, Tarrash SR, Wolfner MF, Clark AG. 2020. Dissecting Fertility Functions of Drosophil a Y Chromosome Genes with CRISPR. *Genetics* **214**: 977–990. https://pubmed.ncbi.nlm.nih.gov/32098759/ (Accessed February 16, 2025).

Halder G, Callaerts P, Gehring WJ. 1995. Induction of Ectopic Eyes by Targeted Expression of the eyeless Gene in Drosophila. *Science (1979)* **67**: 1788–1792. https://www.science.org/doi/10.1126/science.7892602 (Accessed October 23, 2024).

Haynes KA, Gracheva E, Elgin SCR. 2007. A Distinct Type of Heterochromatin Within Drosophila melanogaster Chromosome 4. *Genetics* **175**: 1539–1542. https://dx.doi.org/10.1534/genetics.106.066407 (Accessed January 13, 2025).

Hiller MA, Lin TY, Wood C, Fuller MT. 2001. Developmental regulation of transcription by a tissue-specific TAF homolog. *Genes Dev* **15**: 1021. https://pmc.ncbi.nlm.nih.gov/articles/PMC312677/ (Accessed October 23, 2024).

Hu X, Li M, Hao X, Lu Y, Zhang L, Wu G. 2021. The Osa-Containing SWI/SNF Chromatin-Remodeling Complex Is Required in the Germline Differentiation Niche for Germline Stem Cell Progeny Differentiation. *Genes 2021, Vol 12, Page 363* **12**: 363. https://www.mdpi.com/2073-4425/12/3/363/htm (Accessed October 29, 2024).

Hudson AG, Parrott BB, Qian Y, Schulz C. 2013. A temporal signature of epidermal growth factor signaling regulates the differentiation of germline cells in testes of Drosophila melanogaster. *PLoS One* **8**. https://pubmed.ncbi.nlm.nih.gov/23940622/ (Accessed April 16, 2025).

Kawaguchi S, Ueki M, Kai T. 2020. Drosophila MARF1 ensures proper oocyte maturation by regulating nanos expression. *PLoS One* **15**: e0231114. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0231114 (Accessed March 25, 2025).

Kawase E, Wong MD, Ding BC, Xie T. 2004. Gbb/Bmp signaling is essential for maintaining germline stem cells and for repressing bam transcription in the Drosophila testis. *Development* **131**: 1365–1375.

Kelly KK, Meadows SM, Cripps RM. 2002. Drosophila MEF2 is a direct regulator of Actin57B transcription in cardiac, skeletal, and visceral muscle lineages. *Mech Dev* **110**: 39–50.
Kharchenko P V., Alekseyenko AA, Schwartz YB, Minoda A, Riddle NC, Ernst J, Sabo PJ, Larschan E, Gorchakov AA, Gu T, et al. 2011. Comprehensive analysis of the chromatin landscape in Drosophila melanogaster. *Nature* **471**: 480–485. http://www.nature.com/doifinder/10.1038/nature09725 (Accessed November 25, 2016).
Kiger AA, Jones DL, Schulz C, Rogers MB, Fuller MT. 2001. Stem Cell Self-Renewal Specified by JAK-STAT Activation in Response to a Support Cell Cue. *Science (1979)* **294**: 2542–2545. http://www.ncbi.nlm.nih.gov/pubmed/11752574 (Accessed January 6, 2018).
Kiger AA, White-Cooper H, Fuller MT. 2000. Somatic support cells restrict germline stem cell self-renewal and promote differentiation. *Nature* **407**: 750–754. https://pubmed.ncbi.nlm.nih.gov/11048722/ (Accessed April 16, 2025).
Kouzarides T. 2007. Chromatin Modifications and Their Function. *Cell* **128**: 693–705. https://www.sciencedirect.com/science/article/pii/S0092867407001845 (Accessed March 30, 2018).
Kuhn R, Kuhn C, Börsch D, Glätzer KH, Schäfer U, Schäfer M. 1991. A cluster of four genes selectively expressed in the male germ line of Drosophila melanogaster. *Mech Dev* **35**: 143–151.
Kwan EX, Alvino GM, Lynch KL, Levan PF, Amemiya HM, Wang XS, Johnson SA, Sanchez JC, Miller MA, Croy M, et al. 2023. Ribosomal DNA replication time coordinates completion of genome replication and anaphase in yeast. *Cell Rep* **42**: 112161.
Leatherman JL, Di Nardo S. 2008. Zfh-1 controls somatic stem cell self-renewal in the Drosophila testis and nonautonomously influences germline stem cell self-renewal. *Cell Stem Cell* **3**: 44–54.
Leatherman JL, Dinardo S. 2010. Germline self-renewal requires cyst stem cells and stat regulates niche adhesion in Drosophila testes. *Nat Cell Biol* **12**: 806–811. /pmc/articles/PMC2917891/ (Accessed September 17, 2024).
Lin H, Yue L, Spradling AC. 1994. The Drosophila fusome, a germline-specific organelle, contains membrane skeletal proteins and functions in cyst formation. *Development* **120**: 947–956. https://dx.doi.org/10.1242/dev.120.4.947 (Accessed October 23, 2024).
Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 1–21. https://genomebiology.biomedcentral.com/articles/10.1186/s13059-014-0550-8 (Accessed March 11, 2025).
Lu KL, Nelson JO, Watase GJ, Warsinger-Pepe N, Yamashita YM. 2018. Transgenerational dynamics of rDNA copy number in drosophila male germline stem cells. *Elife* **7**.
Lubelsky Y, Prinz JA, DeNapoli L, Li Y, Belsky JA, MacAlpine DM. 2014. DNA replication and transcription programs respond to the same chromatin cues. *Genome Res* **24**: 1102–1114. https://pubmed.ncbi.nlm.nih.gov/24985913/ (Accessed May 30, 2021).
Marchal C, Sasaki T, Vera D, Wilson K, Sima J, Rivera-Mulia JC, Trevilla-García C, Nogues C, Nafie E, Gilbert DM. 2018. Genome-wide analysis of replication timing by next-generation sequencing with E/L Repli-seq. *Nature Protocols 2018 13:5* **13**: 819–839. https://www.nature.com/articles/nprot.2017.148 (Accessed November 21, 2024).
Matunis EL, Stine RR, de Cuevas M. 2012. Recent advances in Drosophila male germline stem cell biology. *Spermatogenesis* **2**: 137–144.
McKearin DM, Spradling AC. 1990. bag-of-marbles: a Drosophila gene required to initiate both male and female gametogenesis. *Genes Dev* **4**: 2242–51. http://www.ncbi.nlm.nih.gov/pubmed/2279698 (Accessed May 8, 2019).
McLaughlin CN, Qi Y, Quake SR, Luo L, Li H. 2022. Isolation and RNA sequencing of single nuclei from Drosophila tissues. *STAR Protoc* **3**. https://pubmed.ncbi.nlm.nih.gov/35620068/ (Accessed November 21, 2024).
Meiklejohn CD, Landeen EL, Cook JM, Kingan SB, Presgraves DC. 2011. Sex Chromosome-Specific Regulation in the Drosophila Male Germline But Little Evidence for Chromosomal Dosage

Compensation or Meiotic Inactivation. *PLoS Biol* **9**: e1001126. https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.1001126 (Accessed March 5, 2025).

Ng CL, Qian Y, Schulz C. 2019. Notch and Delta are required for survival of the germline stem cell lineage in testes of Drosophila melanogaster. *PLoS One* **14**. https://pubmed.ncbi.nlm.nih.gov/31513679/ (Accessed April 16, 2025).

Öztürk-Çolak A, Marygold SJ, Antonazzo G, Attrill H, Goutte-Gattat D, Jenkins VK, Matthews BB, Millburn G, dos Santos G, Tabone CJ, et al. 2024. FlyBase: updates to the Drosophila genes and genomes database. *Genetics* **227**. https://dx.doi.org/10.1093/genetics/iyad211 (Accessed March 11, 2025).

Pang LY, De Luca S, Zhu H, Urban JM, Spradling AC. 2023. Chromatin and gene expression changes during female Drosophila germline stem cell development illuminate the biology of highly potent stem cells. *Elife* **12**. https://pubmed.ncbi.nlm.nih.gov/37831064/ (Accessed February 16, 2025).

Papagiannouli F. 2014. Male stem Cell Niche and Spermatogenesis in the Drosophila testis — A Tale of Germline-Soma Communication. *Adult Stem Cell Niches*. https://www.intechopen.com/chapters/47159 (Accessed April 8, 2025).

Parrott BB, Hudson A, Brady R, Schulz C. 2012. Control of germline stem cell division frequency--a novel, developmentally regulated role for epidermal growth factor signaling. *PLoS One* **7**. https://pubmed.ncbi.nlm.nih.gov/22586473/ (Accessed April 16, 2025).

Pratto F, Brick K, Cheng G, Lam KWG, Cloutier JM, Dahiya D, Wellard SR, Jordan PW, Camerini-Otero RD. 2021. Meiotic recombination mirrors patterns of germline replication in mice and humans. *Cell* **184**: 4251-4267.e20. http://www.cell.com/article/S0092867421007935/fulltext (Accessed January 10, 2022).

Rastelli L, Kuroda MI. 1998. An analysis of maleless and histone H4 acetylation in Drosophila melanogaster spermatogenesis. *Mech Dev* **71**: 107–117.

Rausch C, Weber P, Prorok P, Hörl D, Maiser A, Lehmkuhl A, Chagin VO, Casas-Delucchi CS, Leonhardt H, Cristina Cardoso M. 2020. Developmental differences in genome replication program and origin activation. *Nucleic Acids Res* **48**: 12751–12777. https://dx.doi.org/10.1093/nar/gkaa1124 (Accessed March 6, 2025).

Raz AA, Vida GS, Stern SR, Mahadevaraju S, Fingerhut JM, Viveiros JM, Pal S, Grey JR, Grace MR, Berry CW, et al. 2023. Emergent dynamics of adult stem cell lineages from single nucleus and single cell RNA-Seq of Drosophila testes. *Elife* **12**. https://pubmed.ncbi.nlm.nih.gov/36795469/ (Accessed September 17, 2024).

Rhind N, Gilbert DM. 2013. DNA replication timing. *Cold Spring Harb Perspect Biol* **5**. https://pubmed.ncbi.nlm.nih.gov/23838440/ (Accessed January 5, 2025).

Riddle NC, Leung W, Haynes KA, Granok H, Wuller J, Elgin SCR. 2008. An investigation of heterochromatin domains on the fourth chromosome of Drosophila melanogaster. *Genetics* **178**: 1177–1191. https://www.genetics.org/content/178/3/1177 (Accessed June 6, 2021).

Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. 2015. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**: e47. https://pubmed.ncbi.nlm.nih.gov/25605792/ (Accessed March 11, 2025).

Samata M, Akhtar A. 2018. Dosage Compensation of the X Chromosome: A Complex Epigenetic Assignment Involving Chromatin Regulators and Long Noncoding RNAs. *Annu Rev Biochem* **87**: 323–350. https://www.annualreviews.org/content/journals/10.1146/annurev-biochem-062917-011816 (Accessed March 26, 2025).

Schulz C, Wood CG, Jones DL, Tazuke SI, Fuller MT. 2002. Signaling from germ cells mediated by the rhomboid homolog stet organizes encapsulation by somatic support cells. *Development* **129**: 4523–4534. https://pubmed.ncbi.nlm.nih.gov/12223409/ (Accessed October 23, 2024).

Schwaiger M, Stadler MB, Bell O, Kohler H, Oakeley EJ, Sehübeler D. 2009. Chromatin state marks cell-type- and gender-specific replication of the Drosophila genome. *Genes Dev* **23**: 589. https://pmc.ncbi.nlm.nih.gov/articles/PMC2658520/ (Accessed March 10, 2025).

Sheng XR, Brawley CM, Matunis EL. 2009. Dedifferentiating spermatogonia outcompete somatic stem cells for niche occupancy in the Drosophila testis. *Cell Stem Cell* **5**: 191–203. https://pubmed.ncbi.nlm.nih.gov/19664993/ (Accessed March 19, 2025).

Shevelyov YY, Ulianov S V., Gelfand MS, Belyakin SN, Razin S V. 2022. Dosage Compensation in Drosophila: Its Canonical and Non-Canonical Mechanisms. *International Journal of Molecular Sciences 2022, Vol 23, Page 10976* **23**: 10976. https://www.mdpi.com/1422-0067/23/18/10976/htm (Accessed March 19, 2025).

Shi Z, Lim C, Tran V, Cui K, Zhao K, Chen X. 2020. Single-cyst transcriptome analysis of Drosophila male germline stem cell lineage. *Development* **147**. https://doi.org/10.1242/dev.184259.

Shivdasani AA, Ingham PW. 2003. Regulation of stem cell maintenance and transit amplifying cell proliferation by tgf-beta signaling in Drosophila spermatogenesis. *Curr Biol* **13**: 2065–2072. https://pubmed.ncbi.nlm.nih.gov/14653996/ (Accessed September 17, 2024).

Thomas PD, Ebert D, Muruganujan A, Mushayahama T, Albou LP, Mi H. 2022. PANTHER: Making genome-scale phylogenetics accessible to all. *Protein Sci* **31**: 8–22. https://pubmed.ncbi.nlm.nih.gov/34717010/ (Accessed March 11, 2025).

Tran V, Lim C, Xie J, Chen X. 2012. Asymmetric division of Drosophila male germline stem cell shows asymmetric histone distribution. *Science (1979)* **338**: 679–82. http://www.ncbi.nlm.nih.gov/pubmed/23118191 (Accessed March 30, 2018).

Tulina N, Matunis E. 2001. Control of stem cell self-renewal in Drosophila spermatogenesis by JAK-STAT signaling. *Science (1979)* **294**: 2546–2549. http://www.ncbi.nlm.nih.gov/pubmed/11752575 (Accessed January 6, 2018).

Tyagi M, Imam N, Verma K, Patel AK. 2016. Chromatin remodelers: We are the drivers!! *Nucleus* **7**: 388–404. https://www.tandfonline.com/doi/abs/10.1080/19491034.2016.1211217 (Accessed October 23, 2024).

Urban J, Kuzu G, Bowman S, Scruggs B, Henriques T, Kingston R, Adelman K, Tolstorukov M, Larschan E. 2017. Enhanced chromatin accessibility of the dosage compensated Drosophila male X-chromosome requires the CLAMP zinc finger protein ed. A. Akhtar. *PLoS One* **12**: e0186855. http://www.ncbi.nlm.nih.gov/pubmed/29077765 (Accessed October 30, 2017).

Vedelek V, Bodai L, Grézal G, Kovács B, Boros IM, Laurinyecz B, Sinka R. 2018. Analysis of Drosophila melanogaster testis transcriptome 06 Biological Sciences 0604 Genetics. *BMC Genomics* **19**: 1–19. https://bmcgenomics.biomedcentral.com/articles/10.1186/s12864-018-5085-z (Accessed September 17, 2024).

Vidaurre V, Chen X. 2021. Epigenetic regulation of drosophila germline stem cell maintenance and differentiation. *Dev Biol* **473**: 105–118. https://www.sciencedirect.com/science/article/pii/S0012160621000361.

Wen K, Yang L, Xiong T, Di C, Ma D, Wu M, Xue Z, Zhang X, Long L, Zhang W, et al. 2016. Critical roles of long noncoding RNAs in Drosophila Spermatogenesis. *Genome Res* **26**: 1233–1244. https://pmc.ncbi.nlm.nih.gov/articles/PMC5052038/ (Accessed March 10, 2025).

Wiles ET, Selker EU. 2016. H3K27 methylation: a promiscuous repressive chromatin mark. *Curr Opin Genet Dev* **43**: 31. https://pmc.ncbi.nlm.nih.gov/articles/PMC5447479/ (Accessed April 9, 2025).

Williams RW, Rubin GM. 2002. ARGONAUTE1 is required for efficient RNA interference in Drosophila embryos. *Proc Natl Acad Sci U S A* **99**: 6889–6894. https://www.pnas.org/doi/abs/10.1073/pnas.072190799 (Accessed October 27, 2024).

Witt E, Benjamin S, Svetec N, Zhao L. 2019. Testis single-cell RNA-seq reveals the dynamics of de novo gene transcription and germline mutational bias in drosophila. *Elife* **8**.

Wooten M, Snedeker J, Nizami ZF, Yang X, Ranjan R, Urban E, Kim JM, Gall J, Xiao J, Chen X. 2019. Asymmetric histone inheritance via strand-specific incorporation and biased replication fork

movement. *Nat Struct Mol Biol* **26**: 732–743. http://www.nature.com/articles/s41594-019-0269-z (Accessed July 28, 2019).

Yang S, Corbett SE, Koga Y, Wang Z, Johnson WE, Yajima M, Campbell JD. 2020. Decontamination of ambient RNA in single-cell RNA-seq with DecontX. *Genome Biol* **21**: 1–15. https://genomebiology.biomedcentral.com/articles/10.1186/s13059-020-1950-6 (Accessed March 11, 2025).

Young MD, Willson TA, Wakefield MJ, Trounson E, Hilton DJ, Blewitt ME, Oshlack A, Majewski IJ. 2011. ChIP-seq analysis reveals distinct H3K27me3 profiles that correlate with transcriptional activity. *Nucleic Acids Res* **39**: 7415. https://pmc.ncbi.nlm.nih.gov/articles/PMC3177187/ (Accessed October 24, 2024).

Yu J, Li Z, Fu Y, Sun F, Chen X, Huang Q, He L, Yu H, Ji L, Cheng X, et al. 2023. Single-cell RNA-sequencing reveals the transcriptional landscape of ND-42 mediated spermatid elongation via mitochondrial derivative maintenance in Drosophila testes. *Redox Biol* **62**: 102671.

Zhang J, Luo J, Chen J, Dai J, Montell C. 2020. The Role of Y Chromosome Genes in Male Fertility in Drosophila melanogaster. *Genetics* **215**: 623–633. https://pubmed.ncbi.nlm.nih.gov/32404399/ (Accessed February 16, 2025).

Zhang Y, Huang Y, Dong Y, Liu X, Kou X, Zhao Y, Zhao A, Sun J, Su Z, Li Z, et al. 2021. Unique Patterns of H3K4me3 and H3K27me3 in 2-Cell-like Embryonic Stem Cells. *Stem Cell Reports* **16**: 458–469. http://www.cell.com/article/S2213671121000503/fulltext (Accessed October 29, 2024).

Zhu A, Ibrahim JG, Love MI. 2019. Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. *Bioinformatics* **35**: 2084–2092. https://pubmed.ncbi.nlm.nih.gov/30395178/ (Accessed March 11, 2025).

Zoller R, Schulz C. 2012. The Drosophila cyst stem cell lineage: Partners behind the scenes? *Spermatogenesis* **2**: 145–157. http://www.ncbi.nlm.nih.gov/pubmed/23087834 (Accessed March 25, 2019).