Safeguarding *Drosophila* female germ cell identity depends on an H3K9me3 mini domain guided by a ZAD zinc finger protein

Laura Shapiro-Kulnane[1], Micah Selengut[1], and Helen K. Salz[1*]

[1]Department of Genetics and Genome Sciences, Case Western Reserve University School of Medicine, Cleveland, OH 44106-4955, USA.

*Corresponding author: hks@case.edu

Short title: H3K9me3 gene silencing in the Drosophila female germline

Key words: oogenesis, CG4936, epigenetic silencing, heterochromatin, SETDB1, *phf7*

Abstract

H3K9me3-based gene silencing is a conserved strategy for securing cell fate, but the mechanisms controlling lineage-specific installation of this epigenetic mark remain unclear. In *Drosophila*, H3K9 methylation plays an essential role in securing female germ cell fate by silencing lineage inappropriate *phf7* transcription. Thus, *phf7* regulation in the female germline provides a powerful system to dissect the molecular mechanism underlying H3K9me3 deposition onto protein coding genes. Here we used genetic studies to identify the essential cis-regulatory elements, finding that the sequences required for H3K9me3 deposition are conserved across *Drosophila* species. Transposable elements are also silenced by an H3K9me3-mediated mechanism. But our finding that *phf7* regulation does not require the dedicated piRNA pathway components, *piwi, aub, rhino, panx,* and *nxf2,* indicates that the mechanisms of H3K9me3 recruitment are distinct. Lastly, we discovered that an uncharacterized member of the zinc finger associated domain (ZAD) containing C2H2 zinc finger protein family, IDENTITY CRISIS (IDC; CG4936), is necessary for H3K9me3 deposition onto *phf7.* Loss of *idc* in germ cells interferes with *phf7* transcriptional regulation and H3K9me3 deposition, resulting in ectopic PHF7 protein expression. IDC's role is likely to be direct, as it localizes to a conserved domain within the *phf7* gene. Collectively, our findings support a model in which IDC guides sequence-specific establishment of an H3K9me3 mini domain, thereby preventing accidental female-to-male programming.

Author Summary

Tissue development and function relies on cells remembering their identity. A cell's identity is defined by the genes it expresses and those it does not. Recent work has shown that genes can be silenced by trimethylation of histone H3 lysine 9 (H3K9me3) marked chromatin, and that H3K9me3-mediated gene silencing is a vital strategy for securing cell fate. But there is very little information about how the machinery responsible for H3K9 methylation finds its target

genes. Here we explore this issue using the *Drosophila* female germline where a mini domain of repressive H3K9me3 chromatin secures female germ cell fate by silencing *phf7,* a gene normally expressed in male germ cells. Transposable elements are also silenced by H3K9me3 mini domains, but we find that the proteins involved in this process are not required for *phf7* silencing. Instead, we find that silencing requires a previously uncharacterized protein, we have named IDENTITY CRISIS. Our work provides evidence that IDENTITY CRISIS directs the H3K9 methylation machinery to build a mini domain at the *phf7* locus. Our results shed new light into how cells safeguard their identity by silencing cell type inappropriate genes, and more specifically how these genes are identified by the silencing machinery.

Introduction

Gene silencing is critical to establishing and maintaining cell fates. Once made, the decision to silence a gene is fortified by the acquisition of repressive histone modifications. While cell type specific epigenetic silencing is primarily associated with tri-methylation of histone H3 lysine 27 (H3K27me3)-marked chromatin, tissue specific genes can also be repressed by H3K9 methylation [1–4]. For example, in *S. pombe,* discrete H3K9me3 domains silence meiotic genes in vegetative cells (e.g., [5,6]). In *C. elegans,* H3K9 methylation silences inappropriate cell type specific genes, including germline genes, in somatic cells (e.g., [7,8]). In the *D. melanogaster* female germline, H3K9me3 silences male germline genes [9]. In the mouse, H3K9me3 silences germline genes during early embryonic development (e.g., [10,11]). Studies carried out in mammalian tissue culture systems further identify H3K9me3-mediated gene silencing as a conserved and vital strategy for maintaining cell fates in a wide range of tissues (e.g., [12–23]). However, the molecular mechanisms controlling tissue specific installation of this epigenetic modification onto protein-coding genes are largely unknown.

63 The repressive H3K9me3 histone modification has well characterized roles in
64 constitutive heterochromatin formation, and the transcriptional silencing of repetitive DNA
65 elements such as satellite repeats and transposable elements (TEs) [24–26]. These studies
66 identified two mechanisms of H3K9me3 recruitment. One mechanism involves small RNAs that
67 guide localization through a complementary base pairing mechanism. In *Drosophila* and
68 mammals, for example, the PIWI-associated small RNAs (piRNAs) guide the H3K9me3
69 silencing machinery to TEs [27]. A second mechanism involves sequence specific DNA binding
70 proteins. In mammals, for example, H3K9me3 can be guided to TEs by members of the
71 vertebrate specific Kruppel-associated box zinc finger (KRAB-ZNF) family of DNA binding
72 proteins [28–31]. An analogous mechanism might exist in *Drosophila*, as two zinc finger
73 proteins, KIPFERL and SMALL OVARY, were recently shown to have roles in TE silencing [32–
74 35]. Whether installation of this epigenetic modification onto protein-coding genes employs
75 similar mechanisms remains unclear.

76 In *Drosophila,* H3K9 methylation plays an essential role in securing female germ cell fate
77 by silencing lineage inappropriate *PHD finger protein 7* (*phf7*) transcription [9,36]. Thus, *phf7*
78 regulation in the female germline provides a powerful system to investigate how the H3K9me3
79 silencing mark is installed at protein-coding genes. *phf7* encodes a predicted chromatin reader,
80 first identified in a screen for genes expressed in male but not female embryonic germ cells [37].
81 Curiously, *phf7* is not essential for male fertility, but it is critical that female germ cells not
82 express the PHF7 protein. Forced expression of PHF7 activates a toxic gene expression
83 program enriched for genes usually restricted to the male germline [36]. Even though PHF7
84 protein is limited to male germ cells, *phf7* is transcribed in both male and female germ cells. Sex
85 specificity is achieved by alternative transcription start sites (TSS; **Fig 1A**). In testes,
86 transcription from the upstream TSS (TSS1) produces a long mRNA isoform that makes protein.
87 In ovaries, an H3K9me3 mini domain prevents the selection of TSS1. Instead, transcription

88 initiates from the downstream TSS (TSS2) to produce a short mRNA isoform that is not
89 translated and has no function. Of the three *Drosophila* enzymes known to methylate H3K9,
90 only SETDB1 has a specific and nonredundant role in repressing *phf7* [9]. Germ cell specific
91 loss of SETDB1, its binding partner ATF7IP, or the H3K9 reader HP1a produced ovarian germ
92 cell tumors that inappropriately express the PHF7 protein. Notably, derepression results from
93 losing the H3K9me3 mini domain. While these results establish the H3K9me3 mini domain
94 controls *phf7* transcription, the mechanisms controlling installation of this epigenetic modification
95 remains unclear.

96 We address the mechanism of targeted H3K9me3 deposition onto *phf7* by identifying
97 essential cis-regulatory elements and trans-acting factors. We establish that the required cis-
98 regulatory sequences are conserved across *Drosophila* species and find that the mechanism
99 governing *phf7* regulation is different from what has been described for piRNA-guided H3K9me3
100 deposition on TEs. Lastly, we discover that repression depends on the previously unknown
101 gene, *CG4936,* that we have named *identity crisis (idc). idc* encodes a zinc finger associated
102 domain (ZAD) containing C2H2 zinc finger protein. Notably, loss of *idc* in germ cells interferes
103 with *phf7* repression by reducing H3K9me3 deposition. Together with the observation that IDC
104 localizes to a conserved region within the *phf7* gene, our analysis supports a model in which
105 IDC guides H3K9me3 installation, thereby preventing accidental female-to-male programming.

106 Results

107 Identification of cis-regulatory elements required for H3K9me3 deposition

108 H3K9me3 accumulates over a three kb region within the *phf7* gene [9]. This discrete
109 peak covers the male TSS1, the first exon, and most of the first intron. Interestingly, the intron
110 contains seven copies of ~250 bp sequence not found anywhere else in the genome (element
111 R; **Fig 1A and Fig S1**). Repetitive elements can regulate gene expression in *cis* by serving
112 directly as H3K9me3 nucleation sites or in *trans* by encoding small RNAs that serve as guides

113 via a base pairing mechanism. We, therefore, hypothesized that the repeats play a role in
114 H3K9me3 deposition and silencing of TSS1. To test this idea, we used CRISPR editing to
115 produce animals deleted for the repeats in the endogenous locus (*phf7*$^{\Delta R}$) (**Fig 1A**). We were
116 surprised to discover that homozygous *phf7*$^{\Delta R}$ animals are viable and fertile, suggesting that the
117 repeats may not be essential for repression. To evaluate the impact of deleting the repeats in
118 more detail, we measured H3K9me3 accumulation in wild-type and *phf7*$^{\Delta R}$ homozygous ovaries
119 using chromatin immunoprecipitation followed by quantitative PCR (ChIP-qPCR). In wild-type
120 (control) ovaries, sequences in the intron positioned 319 bp downstream of TSS1 (within intronic
121 element A) are significantly enriched in the H3K9me3 modification when compared to the region
122 within the coding sequence (CDS; $p<0.0001$) (**Fig 1B**). Significant H3K9me3 enrichment within
123 element A compared to the CDS region ($p<0.0001$) was also observed in *phf7*$^{\Delta R}$ homozygous
124 ovaries. Interestingly, there was less H3K9me3 accumulation within element A in *phf7*$^{\Delta R}$ than in
125 controls ($p=0.0003$), raising the possibility that sex specific transcriptional regulation might be
126 disrupted. In testes, transcription from the upstream TSS (TSS1) produces a long mRNA
127 isoform, called *phf7-RC.* We therefore used RT-qPCR to assay for testis specific *phf7-RC*
128 expression levels in wild-type and *phf7*$^{\Delta R}$ gonads. Contrary to our expectations, we found that
129 *phf7*$^{\Delta R}$ did not disrupt sex specific transcription: As in wild-type gonads, *phf7-RC* expression was
130 detected in *phf7*$^{\Delta R}$ homozygous testis but not ovaries (**Fig 1C**). These results indicate that the
131 repeats are not essential for H3K9me3-mediated repression.

132 To complement these studies, we used a transgenic approach to identify the sequences
133 capable of promoting H3K9me3 to an ectopic location. For these assays we created transgenic
134 flies that carry portions of the *phf7* locus, beginning at TSS1 and ending at the beginning of the
135 open reading frame in exon 2, inserted into the same heterologous genomic location on the 3rd
136 chromosome using site specific integration. To assay for H3K9me3 accumulation on the
137 transgene, but not at the endogenous locus, we crossed each line into a *phf7*$^{\Delta 13}$ background.

138 *phf7*$^{\Delta 13}$ is a 1.81 kb intragenic deletion allele, which allows us to measure the level of H3K9me3
139 deposition by ChIP-qPCR at a region that is present in the transgenes but absent in *phf7*$^{\Delta 13}$
140 (intronic element B, **Fig 2A**). In agreement with our previous findings, deleting the seven
141 tandem repeats (element R) from the first intron did not interfere with ectopic H3K9me3
142 accumulation. (**Fig 2A and Fig 2B**). In contrast, removing elements A and R abolished
143 H3K9me3 accumulation (**Fig 2A and Fig 2B**). These data demonstrate that cis-regulatory
144 determinants required for H3K9me3 recruitment are within element A.
145
146 **Conservation of non-coding cis-regulatory elements**
147 Cis-regulatory elements are often conserved. Thus, we might expect the essential
148 sequences within element A to be retained in evolution. To test this prediction, we chose to
149 examine the level of conservation between *D. melanogaster, D. simulans,* and *D. yakuba.*
150 Although *D. simulans* and *D. yakuba* are separated from *D. melanogaster* by 5-15 million years,
151 the *phf7* intron/exon structure is conserved (**Fig 3A**). Furthermore, RNA-seq data obtained from
152 wild-type testis and ovaries [38,39] shows that *phf7* expression is sex specific in *D. simulans*
153 and *D. yakuba* (**Fig 3A)**. Together these observations suggest that sex-specific transcriptional
154 regulation through alternative TSS selection is conserved between *D. melanogaster, D.*
155 *simulans,* and *D. yakuba*. When we compared the level of sequence conservation in the first
156 intron between *D. melanogaster, D. simulans,* and *D. yakuba,* we found that only the first 235 bp
157 of intron element A was conserved (**Fig 3B and Fig S2**). Interestingly, the first non-coding 174
158 bp exon also displays a high level of conservation, suggesting that additional cis-regulatory
159 elements may be present within the first non-coding exon (**Fig 3B and Fig S2**)
160
161 **Core components of the piRNA pathway are not required for *phf7* sex specific**
162 **transcriptional control**

163 Prior studies have shown that SETDB1 is required for both TE and *phf7* silencing in
164 germ cells [9,40–42]. Although there are no recognizable TE sequences at *phf7*, the piRNA
165 pathway may nevertheless play a role in repressing *phf7.* To test this possibility, we analyzed
166 published RNA-seq data sets from ovaries carrying mutations in genes encoding the dedicated
167 piRNA pathway components, *piwi, aubergine (aub), rhino, panoramix (panx,* also known as
168 *silenceo),* and *nuclear RNA export factor 2 (nxf2)* [43–45]. Loss of each component causes the
169 piRNA pathway to collapse. *piwi, aub* and *rhino* are essential for piRNA biogenesis [42,43,46].
170 *panx* and *nxf2,* while dispensable for piRNA biogenesis, are necessary for SETDB1 to deposit
171 H3K9me3 marks onto TEs [45,47–51]. In agreement with our prior studies [9], only the shorter
172 *phf7-RA* transcript is detectable in wild-type ovaries, but the longer testis specific *phf7-RC*
173 transcript is present in *setdb1* mutant ovaries **(Fig 4)**. In contrast to the *setdb1* mutant ovaries,
174 *piwi, aub, rhino, panx,* and *nxf2* mutant ovaries only express the shorter *phf7-RA* transcript,
175 indicating that transcriptional regulation is not disrupted. Based on these findings we conclude
176 that the mechanisms controlling H3K9me3 deposition onto *phf7* and TEs are distinct.
177
178 **The ZAD-ZNF protein IDC is required for *phf7* sex specific transcriptional control**
179 In mammalian cells, SETDB1 can be recruited to its targets by members of the large
180 KRAB-zinc finger family of sequence specific DNA binding proteins [28–30]. The KRAB family is
181 confined to mammals. It has been hypothesized that members of the insect specific ZAD-ZNF
182 family might be functional analogues [52–54]. We recently tested the function of 68 of the 93
183 ZAD-ZNF encoding genes in female germ cells by performing an RNAi screen [55]. These data
184 revealed that loss of CG4936 in germ cells phenocopies the ovarian defects caused by ectopic
185 PHF7 expression. CG4936, which we name Identity crisis (IDC), is a 521 amino acid (aa)
186 protein that contains an N-terminal ZAD (zinc finger associated domain, 22-95 aa), an
187 unstructured linker region, and a C-terminal domain that includes an array of 5 C2H2 zinc
188 fingers (386-491 aa). Studies of other ZAD-ZNF proteins suggest that the ZAD mediates

189 protein-protein interactions and the C2H2 zinc fingers bind DNA [56–60]. We, therefore,
190 hypothesize that IDC is required for *phf7* repression.
191 To investigate the possibility that loss of *idc* in female germ cells disrupts *phf7*
192 repression, we first used RT-qPCR to assay for the presence of the testis specific *phf7-RC*
193 transcript. Germ cell specific knock down (GLKD) was achieved using an inducible RNA
194 interference (RNAi) transgene and the germ cell specific *nos-GAL4* driver. To rule out indirect
195 effects arising from RNAi knockdown, we also knocked down *white*, which is not expressed in
196 germ cells, as a negative control. Using primer pairs capable of detecting *phf7-RC,* we found
197 that in contrast to control ovaries, *phf7-RC* is detectable in *idc* GLKD mutant ovaries (**Fig 5A**).
198 We then asked whether ectopic *phf7-RC* expression correlated with ectopic PHF7 protein
199 expression. Previous work has shown that a BAC transgene that encodes an N-terminally HA-
200 tagged PHF7 protein (HA-PHF7) is a faithful reporter of the sex specific protein expression
201 pattern of PHF7 [37,61]. We found that, in contrast to wild-type ovaries, the HA-PHF7 protein is
202 detectable in the cytoplasm and the nucleus of the *idc GLKD* mutant ovaries (**Fig 5B and Fig**
203 **5C**). We conclude that IDC is required for *phf7* regulation.
204 To test whether *idc* plays a role in controlling H3K9me3 deposition, we compared the
205 amount of H3K9me3 accumulation within element A in wild-type and mutant ovaries using ChIP-
206 qPCR. We found that H3K9me3 was significantly reduced in *idc* GLKD ovaries compared to
207 control ovaries ($p<0.0001$; **Fig 5D**). Based on these studies, we conclude that IDC regulates
208 *phf7* transcription by promoting H3K9me3 deposition.
209
210 **IDC is a nuclear protein that associates with DNA in ovaries**
211 To understand the role of IDC in H3K9me3 deposition, we first determined its
212 localization pattern in the adult ovary. We used a genomic fosmid transgene for these studies
213 that encodes a C-terminally GFP-tagged IDC protein (IDC-GFP). We inferred that the GFP tag
214 does not interfere with *idc* function because the *idc-GFP* rescues the idc^1 lethal phenotype (see

215 materials and methods). Each ovary comprises 16-20 ovarioles that contain germ cells
216 spanning the range of maturity from germline stem cells (GSCs) at the anterior end to fully
217 mature eggs at the posterior end [62]. We found that IDC-GFP is detectable in the nucleus of
218 several different cell types within the egg chambers, including nurse cells, oocytes, and somatic
219 follicle cells (**Fig 6A**). IDC-GFP is tightly associated with the polytene chromosomes in the nurse
220 cells, consistent with its presumed DNA binding activity (**Fig 6B**).

221 In the germarium, at the anterior end of the ovariole, we found that the IDC-GFP protein
222 accumulates to high levels in the nucleus of only 3 to 6 germ cells (boxed region; **Fig 6A)**. This
223 expression pattern is like that described for the female specific sex determination protein SEX-
224 LETHAL (SXL), which is also required for *phf7* H3K9me3 silencing [9,63,64]. Indeed, co-staining
225 experiments reveal that SXL accumulates in the cytoplasm of all IDC-GFP expressing germ
226 cells (**Fig 6C**). Studies have shown that SXL accumulates in GSCs and their immediate
227 daughter cells. When GSCs divide, the daughter cells at the tip remain a GSC, while the more
228 posterior daughter cells, called a cystoblast (CB), express the BAG OF MARBLES protein
229 (BAM). In agreement with published studies, BAM expression, assessed with a transgene that
230 encodes a C-terminally HA-tagged BAM protein expressed from a *bam* promoter is readily
231 detectable in the cytoplasm of just a few germ cells [65]. Co-staining experiments reveal that
232 IDC-GFP is expressed in the GSCs where BAM-HA is not expressed as well as in the adjacent
233 two to three cells where BAM-HA is detectable (**Fig 6D**). Based on these observations, we
234 conclude that the nuclear IDC protein is expressed in the GSCs and their immediate daughter
235 cells.

236

237 **IDC binds to sequences within the male specific exon**

238 IDC contains five zinc fingers suggesting that it has sequence specific DNA binding
239 activity. We speculated that IDC might promote H3K9me3 deposition by binding to the *phf7*
240 locus. To test this prediction by ChIP-qPCR, we designed primers along the length of the

241 conserved noncoding sequences in the first exon and the intron. These experiments revealed
242 that the IDC-GFP tagged protein associates with chromatin within the first non-coding exon, but
243 not outside of it (**Fig 7A**). Collectively, our studies indicate that IDC regulates *phf7* directly by
244 serving as an H3K9me3 guidance factor.
245

246 ## Discussion

247 SETDB1-controlled H3K9 methylation plays an essential role in securing female germ
248 cell fate by silencing lineage-inappropriate *phf7* transcription [9,36]. SETDB1 is also required for
249 TE silencing, where a piRNA-guided mechanism guides H3K9me3 deposition [40–42]. In this
250 work, we establish that *phf7* is silenced by a piRNA-independent mechanism and discover that
251 regulation depends on IDC, a previously uncharacterized member of the ZAD-ZNF family of
252 DNA binding proteins. Regulation appears direct, as IDC is required for H3K9me3 deposition
253 and localizes to a conserved region within the *phf7* locus. Collectively, our data establish that
254 the sequence specific DNA binding protein IDC guides the installation of the H3K9me3 mini-
255 domain onto the *phf7* locus, thereby preventing accidental female-to-male reprogramming. In
256 addition to extending our understanding of how female germ cell fate is maintained, our studies
257 provide one of the first examples of a ZAD-ZNF protein guiding H3K9me3-mediated gene
258 silencing.

259 Our work is consistent with a simple model in which IDC recruits the SETDB1 H3K9me3
260 methyltransferase to *phf7* (**Fig 7B**). However, we could not detect an interaction between IDC
261 and SETDB1 in co-immunoprecipitation experiments. It is, therefore, possible that another as
262 yet undiscovered protein recruits SETDB1 to *phf7* and that IDC acts at a later step in the
263 pathway to ensure the sequence specific establishment of the H3K9me3 mini domain.
264 The IDC protein is also expressed in male germ cells, suggesting that an accessory
265 protein is required to confer female specificity to *phf7* silencing. One attractive contender is

266 *stonewall (stwl)*. STWL is a female specific heterochromatin-associated protein that acts as a
267 transcriptional repressor *in vitro* [66], associates with SETDB1 in yeast [67] and localizes to the
268 *phf7* locus in S2 cells [68]. Importantly, loss of *stwl* in ovaries leads to the inappropriate
269 expression of the testis *phf7* transcript [68]. Together, these studies raise the possibility that *stwl*
270 collaborates with IDC to silence *phf7* in female germ cells.

271 We have previously established that the H3K9me3 reader protein, HP1a, is essential for
272 *phf7* silencing [9]. A requirement for HP1a is not surprising, as HP1a drives chromatin
273 compaction and transcriptional silencing [69]. HP1a also recruits H3K9 methyltransferases,
274 enabling the spreading of the H3K9me3 repressive domain through a positive feedback
275 mechanism. The spread of H3K9me3 over the three kb region within the *phf7* gene likely occurs
276 by an analogous mechanism. At the *phf7* locus, however, the H3K9me3 domain does not
277 extend into the open reading frame or the neighboring genes. Therefore, more work is needed
278 to understand what stops the spreading of this repressive chromatin domain.

279 The SETDB1-controlled H3K9 methylation silencing pathway in female germ cells is not
280 restricted to *phf7.* Genome-wide H3K9me3 profiling in wild-type and mutant ovaries has shown
281 that SETDB1 silences two classes of protein-coding genes. One type includes genes usually
282 expressed in the testis [9]. A second class includes genes that are typically expressed in
283 undifferentiated female germ cells but are silenced once the oocyte is specified, such as
284 *ribosomal protein S19b* [70]. These few examples of context dependent H3K9me3 gene
285 silencing suggest a finely tuned guidance mechanism in the female germline. It will be
286 interesting to explore whether other members of the ZAD-ZNF gene family serve as H3K9me3
287 guidance factors.

288 In summary, by focusing on a single biologically relevant gene, we discovered a putative
289 DNA binding protein that guides the installation of a H3K9me3 repressive domain onto a
290 protein-coding gene. These findings are reminiscent of how TEs can be silenced in mammals

291 wherein members of the KRAB-ZNF protein family recruit SETDB1 to establish epigenetic
292 repression [28–31]. Interestingly, the KRAB-ZNFs are vertebrate specific, and the ZAD-ZNFs
293 are insect specific. Yet, both gene families exhibit similar patterns of species-specific gene
294 expansion and diversification. These observations have fueled the speculation that ZAD-ZNFs
295 and KRAB-ZNFs perform similar functions [52–54]. In fact, IDC's closest human relative is
296 ZNF75D, a KRAB-ZNF protein of unknown function. Whether this or other members of the
297 KRAB-ZNF protein family are required for tissue specific H3K9me3 gene silencing will be an
298 exciting area of further exploration.
299

300 Materials and Methods

301 *Drosophila* stocks and culture conditions

302 Fly strains were kept on standard media at 25 °C unless otherwise noted. Wild-type
303 control flies were from the lab y^1 w^1 stock. Stocks used to report on protein gene activity include
304 the HA-tagged *phf7* transgene, *PBac{3XHA-PHF7}* [37]; the GFP-tagged *idc* transgene,
305 *FlyFos{fTRG00376.sfGFP-TVPTBF}* (VDRC #318556) described in [71] and the HA-tagged
306 *bam* transgene, *P{Bam::HA}* [65]. $phf7^{\Delta 13}$, $phf7^{\Delta R}$, the idc^1 null allele, and the ΔR and ΔAR
307 chromatin reporters were generated for this study, as described below.
308 Crosses for knock-down studies were set up at 29°C, and adults were aged 3-5 days
309 before gonad dissection. Knock-down studies were carried out with the following germline-
310 optimized lines generated by the *Drosophila* Transgenic RNAi Project [72]: *idc-*
311 *P{TRiP.HMC05569}* (BDSC #64550), and the control *w-P{TRiP.GL00094}* (BDSC #35573).
312 Expression of the UAS-RNAi transgenic lines was driven by the third chromosome
313 *P{GAL4::VP16-nos.UTR}* insertion (BDSC #4937), which is strongly expressed in germ cells.
314

315 Generation of the $phf7^{\Delta R}$ allele

316 We generated the *phf7*$^{\Delta R}$ allele in two steps. First, CRISPR was used to replace the
317 repeats within the first intron with a 3XP3::DsRed cassette by Rainbow Transgenic Flies, Inc.
318 The deletion was generated with the following guide RNAs: agttaaaaaaaatcaatcgatgg and
319 cgcagcgattgaatgttaatggg. 1 kb homology arms were generated through PCR and cloned into the
320 pHD-dsRed-attP (Addgene #51019) [73]. Guide RNAs and the donor vector were co-injected
321 into *nos-Cas9* embryos by Rainbow Transgenic Flies. Flies were screened for DsRed
322 expression in the eyes, and sequence verified for accuracy. *phf7*$^{\Delta R}$ was generated by removing
323 the DsRed cassette. Homozygous *phf7*$^{\Delta R+dsRED}$ females were crossed to a Cre-expressing fly
324 line (BDSC #1092), and the male progeny screened for the loss of dsRED expression in the
325 eyes, followed by sequence confirmation of precise tag excision.
326
327 **Generation of the *phf7*$^{\Delta 13}$ allele**
328 The *phf7*$^{\Delta 13}$ allele was generated by P-element mediated imprecise excision as follows:
329 females homozygous for the P-element insertion P{EPgy2}Phf7^{EY03023} (BDSC #15894) were
330 crossed with the Δ2-3 transposase expressing line (BDSC #3664). Potential excision lines were
331 established from the male progeny of the F1 females that had lost the *white*$^{+}$ marker carried by
332 the P element. Stocks that carried deletions were identified by PCR. The exact breakpoints
333 were determined by DNA sequencing the PCR amplified DNA fragments.
334
335 **Generation of the ΔR and ΔAR chromatin transgenic reporter lines**
336 Constructs were generated by VectorBuilder's custom cloning services
337 (https://en.vectorbuilder.com). The transgenes were constructed in their "user-defined promoter"
338 modification of the pUASTattB expression vector. The transgenic constructs were sent to
339 Rainbow Transgenic Flies Inc. for *phi-C31* catalyzed integration into the 65B2 PBac{y[+]-attP-
340 3B}VK00033 site. Transgenic flies were sequence verified for accuracy.

341

342 **Generation of the *idc*^{1} null allele and rescue by the GFP-tagged *idc* transgene**

343 The *idc*^{1} allele was generated by *in vivo* CRISPR mutagenesis [74]. Females expressing
344 two sgRNAs under control of the GAL4/UAS system, P{HD_CFD00598} (VDRC #341525), were
345 crossed at 25°C to M{vas-Cas9}; P{GAL4::VP16-nos.UTR} (BDSC #55821 + #4937) males. The
346 *vas-Cas9; nos>gRNA* male offspring were crossed to a third chromosome balancer line to
347 isolate and balance each putative *idc* allele in the next generation. *idc* alleles were identified by
348 the failure to complement Df(3R)ED6027 (BDSC #9479) and checked for the presence of the
349 predicted indel by PCR.

350 We found that *idc*^{1}*/Df(3R)ED6027* animals do not survive to adulthood (n>100). To verify
351 that the GFP-tagged *idc* transgene, PBac{fTRG00376.sfGFP-TVPTBF} (VDRC #318556),
352 rescues *idc*^{1}*,* we first made the double mutant *idc*^{1}*, idc-GFP.* To test for rescue, we crossed *idc*^{1}*,*
353 *idc-GFP/TM3* to *Df(3R)ED6027/TM3* males. This cross yielded 84% (31/37) of the expected
354 *idc*^{1}*, idc-GFP/ Df(3R)ED6027* progeny, all of which were fertile.

355

356 **qRT-PCR and data analysis**

357 RNA was extracted from dissected gonads using TRIzol (Thermo Fisher, cat#
358 15596026). Quantity and quality were measured using a NanoDrop ND-1000
359 spectrophotometer. The RNA was treated with DNase RQ1 (Promega, cat# M6101). cDNA was
360 generated by reverse transcription using a SuperScript First-Strand Synthesis System for RT-
361 PCR kit (Thermo Fisher, cat# 11904018) using random hexamers. qPCR was performed using
362 *Power* SYBR Green PCR Master Mix (Thermo Fisher, cat# 4368706) with an Applied
363 Biosystems QuantStudio 3 Real-Time PCR system. PCR steps were as follows: 95°C for 10
364 minutes followed by 40 cycles of 95°C for 15 seconds and 60°C for 1 minute. Melt curves were
365 generated with the following parameters: 95°C for 15 seconds, 60°C for 1 minute, 95°C for 15
366 seconds, and 60°C for 15 seconds. Measurements were taken in biological triplicate with two

367 technical replicates each. Relative transcript levels were calculated using the 2-ΔΔCt method
368 [75]. Using GraphPad Prism software, P values were calculated using unpaired two-tailed
369 Student's t-tests. The primers used for measuring RNA levels are listed in **Table S1**.
370
371 **ChIP-qPCR and data analysis**
372 ChIP was performed as described in [9] with some modifications. Buffers are listed in
373 **Table S2**. Briefly, ovaries from 200 1-2 day old adults were dissected in PBSP, homogenized
374 with a pellet pestle, and crosslinked with 1.8% methanol-free formaldehyde (Thermo Fisher,
375 cat# 28906) for 5 minutes. Fixation was quenched for 5 minutes by adding glycine to a final
376 concentration of 225mM. The solution was removed, and the pellet was washed in PBSP 3
377 times. Samples were resuspended in 500μl PBSP and stored at -80°C. Samples were thawed,
378 centrifuged, and resuspended in 1ml lysis buffer 1. Samples were homogenized using sterile
379 homogenizing beads in a Bullet Blender Homogenizer at 4°C for six cycles of 30 seconds on
380 and 1 minute off. Following centrifugation, cell pellets were washed at 4°C for 10 minutes in lysis
381 buffer 1, centrifuged, washed for 10 min at 4°C in lysis buffer 2, centrifuged and then
382 resuspended in 750μl lysis buffer 3. Chromatin was sheared to 100–500 base pairs at 4°C with
383 a Diagenode BioRuptor Pico for 15 cycles of 30 seconds on, and 30 seconds off.
384 For H3K9me3 ChIP, the chromatin lysates were precleared with a 1:1 mix of protein A/G
385 Dynabead magnetic beads (Thermo Fisher, cat# 10001D and 10003D) for 1 hour at 4°C and
386 then incubated overnight at 4°C with 1:1 mix of protein A/G Dynabead magnetic beads
387 conjugated to H3K9me3 antibody (Abcam, cat# 8898 RRID: AB_306848).
388 For GFP ChIP, the chromatin lysates were precleared with Chromotek Magnetic Binding
389 Control Agarose Beads (Proteintech, cat# bmab) for 1 hour at 4°C, and then incubated
390 overnight at 4°C with Chromotek GFP-Trap Magnetic Agarose beads (Proteintech, cat# gtma).

391 Following immunoprecipitation, the samples were washed six times with ChIP-RIPA
392 buffer, two times with ChIP-RIPA/500 buffer, two times with ChIP-LiCl buffer, and twice with TE
393 buffer. DNA was eluted and reverse crosslinked from beads in 200µl elution buffer overnight at
394 65°C. Following RNaseA and proteinase K treatment, DNA was recovered by phenol-chloroform
395 extraction and used for qPCR.
396 qPCR was performed using *Power* SYBR Green PCR Master Mix (Thermo Fisher, cat#
397 4367659) with an Applied Biosystems QuantStudio 3 Real-Time PCR system. PCR steps were
398 as follows: 95ºC for 10 minutes followed by 40 cycles of 95ºC for 15 seconds and 60ºC for 1
399 minute. Melt curves were generated with the following parameters: 95ºC for 15 seconds, 60ºC
400 for 1 minute, 95ºC for 15 seconds, and 60ºC for 15 seconds. Primers as listed in **Table S1**.
401 ChIP experiments were performed on three independent biological samples. qPCR
402 measurements on each sample were performed on at least two technical replicates. For
403 H3K9me3 ChIP-qPCR, the ChIP to input enrichment (presented as percent input) is normalized
404 to a negative control genomic region in the *rp49* locus. Statistical analysis was carried out by the
405 GraphPad Prism software. The P values were calculated using unpaired two-tailed Student's t
406 tests.
407
408 **Immunofluorescence and image analysis**
409 Females were aged 3-5 days before gonad dissection. Ovaries were fixed and stained
410 according to standard procedures with the following primary antibodies: mouse anti-*Drosophila*
411 SXL (1:100, Developmental Studies Hybridoma Bank, cat# M18, RRID: AB_528464), rabbit
412 anti-GFP (1:2500, Thermo Fisher, cat# A-11122, RRID: AB_221569) or FITC conjugated goat
413 anti-GFP (1:750, Abcam, cat# ab6662, RRID: AB_305635), and rat anti-HA high affinity (1:500,
414 Sigma, cat# 11867423001, RRID: AB_390919). The following secondary antibodies were used
415 at 1:200: Alexa Fluor 555 goat anti-rat (Thermo Fisher, cat# A-21434, RRID: AB_2535855),
416 FITC goat anti-mouse (Jackson Immunoresearch Laboratories, cat# 115-095-003, RRID:

417 AB_2338589), or FITC goat anti-rabbit (Jackson Immunoresearch Laboratories, cat# 111-095-
418 003, RRID: AB_2337972). TO-PRO-3 Iodide carbocyanine monomer nucleic acid stain (1:1000,
419 Thermo Fisher, cat# T3605) was used to stain DNA. Images were taken on a Leica SP8
420 confocal with 1024x1024 pixel dimensions, a scan speed of 600 Hz, and a frame average of 3.
421 Sequential scanning was done for each channel, and three Z-stacks were combined for each
422 image. Processed images were compiled with Microsoft PowerPoint.
423

424 **Analysis of publicly available RNA-seq data sets**

425 **Table S3** lists the sources of the publicly available RNA-seq data generated from wild-type and
426 mutant *Drosophila* tissues. The data were downloaded from the SRA database at NCBI and
427 analyzed using the tools available on the Galaxy web platform (https://usegalaxy.org/). FastQC
428 assessed read quality and STAR was used to align the reads to the *Drosophila* reference
429 genome (UCSC dm6). Screenshots of the expression data displayed on The Integrated
430 Genome Viewer (IGV) are shown.

431 *D. simulans* and *D. yakuba phf7* orthologs were identified using the ortholog database at
432 www.flybase.org. RNA-seq data from wild-type *D. simulans* and *D. yakuba* gonads was
433 downloaded from the SRA database at NCBI (**Table S4**). We uploaded the sequencing data to
434 the Galaxy web platform (https://usegalaxy.org/), assessed read quality with FastQC and used
435 STAR to align the reads to their respective reference genomes (listed in **Table S4**). Browser
436 screenshots of the expression data are displayed on the Integrated Genome Viewer (IGV).

437 **DNA sequence alignments**

438 Multiple sequence alignments were performed in SnapGene (https://www.snapgene.com) using
439 MUSCLE with default options. Pairwise sequence alignments were performed in BLASTn
440 (https://blast.ncbi.nlm.nih.gov).
441

Acknowledgments

We would like to thank the members of the Cleveland fly community for helpful discussions, Jane Heatwole for fly food, and FlyBase as an essential resource. Stocks obtained from the Bloomington (NIH P40D018537) and The Vienna *Drosophila* Stock Centers were used in this study. Imaging was performed using equipment purchased through NIH Grant S10-OD016164, housed in the CWRU School of Medicine Imaging Core.

Author contributions

HS conceived this study. LSK and HS generated reagents, designed and executed the experiments, and interpreted the data. MS performed the evolutionary analysis. HS wrote the manuscript with input from LSK and MS.

Declaration of interests

The authors declare no competing interests

Availability of data and material

All genomic sequence and expression data were obtained from public databases and are described in **Table S3** and **S4**. *Drosophila* strains are available upon request.

References

1. Becker JS, Nicetto D, Zaret KS. H3K9me3-Dependent Heterochromatin: Barrier to Cell Fate Changes. Trends in genetics : TIG. 2016;32: 29–41. doi:10.1016/j.tig.2015.11.001

2. Nicetto D, Zaret KS. Role of H3K9me3 heterochromatin in cell identity establishment and maintenance. Current Opinion in Genetics & Development. 2019;55: 1–10. doi:10.1016/j.gde.2019.04.013

467 3. Ninova M, Tóth KF, Aravin AA. The control of gene expression and cell identity by H3K9 trimethylation.
468 Development (Cambridge, England). 2019;146. doi:10.1242/dev.181180

469 4. Padeken J, Methot SP, Gasser SM. Establishment of H3K9-methylated heterochromatin and its
470 functions in tissue differentiation and maintenance. Nat Rev Mol Cell Bio. 2022; 1–18.
471 doi:10.1038/s41580-022-00483-w

472 5. Martienssen R, Moazed D. RNAi and Heterochromatin Assembly. Csh Perspect Biol. 2015;7: a019323.
473 doi:10.1101/cshperspect.a019323

474 6. Zofall M, Yamanaka S, Reyes-Turcu FE, Zhang K, Rubin C, Grewal SIS. RNA elimination machinery
475 targeting meiotic mRNAs promotes facultative heterochromatin formation. Science. 2012;335: 96–100.
476 doi:10.1126/science.1211651

477 7. Rechtsteiner A, Costello ME, Egelhofer TA, Garrigues JM, Strome S, Petrella LN. Repression of
478 Germline Genes in Caenorhabditis elegans Somatic Tissues by H3K9 Dimethylation of Their Promoters.
479 Genetics. 2019;212: 125–140. doi:10.1534/genetics.118.301878

480 8. Methot SP, Padeken J, Brancati G, Zeller P, Delaney CE, Gaidatzis D, et al. H3K9me selectively
481 blocks transcription factor activity and ensures differentiated tissue integrity. Nat Cell Biol. 2021;23: 1163–
482 1175. doi:10.1038/s41556-021-00776-w

483 9. Smolko AE, Shapiro-Kulnane L, Salz HK. The H3K9 methyltransferase SETDB1 maintains female
484 identity in Drosophila germ cells. Nature communications. 2018;9: 4155. doi:10.1038/s41467-018-06697-
485 x

486 10. Mochizuki K, Sharif J, Shirane K, Uranishi K, Bogutz AB, Janssen SM, et al. Repression of germline
487 genes by PRC1.6 and SETDB1 in the early embryo precedes DNA methylation-mediated silencing. Nat
488 Commun. 2021;12: 7020. doi:10.1038/s41467-021-27345-x

489 11. Karimi MM, Goyal P, Maksakova IA, Bilenky M, Leung D, Tang JX, et al. DNA Methylation and
490 SETDB1/H3K9me3 Regulate Predominantly Distinct Sets of Genes, Retroelements, and Chimeric
491 Transcripts in mESCs. Cell Stem Cell. 2011;8: 676–687. doi:10.1016/j.stem.2011.04.004

492 12. Wu K, Liu H, Wang Y, He J, Xu S, Chen Y, et al. SETDB1-Mediated Cell Fate Transition between 2C-
493 Like and Pluripotent States. Cell Reports. 2020;30: 25-36.e6. doi:10.1016/j.celrep.2019.12.010

494 13. Jiang Y, Loh Y-HE, Rajarajan P, Hirayama T, Liao W, Kassim BS, et al. The methyltransferase
495 SETDB1 regulates a large neuron-specific topological chromatin domain. Nature genetics. 2017;49:
496 1239–1250. doi:10.1038/ng.3906

497 14. Du D, Katsuno Y, Meyer D, Budi EH, Chen S-H, Koeppen H, et al. Smad3-mediated recruitment of
498 the methyltransferase SETDB1/ESET controls Snail1 expression and epithelial-mesenchymal transition.
499 EMBO reports. 2018;19: 135–155. doi:10.15252/embr.201744250

500 15. Koide S, Oshima M, Takubo K, Yamazaki S, Nitta E, Saraya A, et al. Setdb1 maintains hematopoietic
501 stem and progenitor cells by restricting the ectopic activation of nonhematopoietic genes. Blood.
502 2016;128: 638–649. doi:10.1182/blood-2016-01-694810

503 16. Schultz DC, Ayyanathan K, Negorev D, Maul GG, Rauscher FJ. SETDB1: a novel KAP-1-associated
504 histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic
505 genes by KRAB zinc-finger proteins. Genes & Development. 2002;16: 919–932. doi:10.1101/gad.973302

506 17. Tan S-L, Nishi M, Ohtsuka T, Matsui T, Takemoto K, Kamio-Miura A, et al. Essential roles of the
507 histone methyltransferase ESET in the epigenetic control of neural progenitor cells during development.
508 Development (Cambridge, England). 2012;139: 3806–3816. doi:10.1242/dev.082198

509 18. Cheng E-C, Hsieh C-L, Liu N, Wang J, Zhong M, Chen T, et al. The Essential Function of SETDB1 in
510 Homologous Chromosome Pairing and Synapsis during Meiosis. Cell Reports. 2021;34: 108575.
511 doi:10.1016/j.celrep.2020.108575

512 19. Lohmann F, Loureiro J, Su H, Fang Q, Lei H, Lewis T, et al. KMT1E mediated H3K9 methylation is
513 required for the maintenance of embryonic stem cells by repressing trophectoderm differentiation. Stem
514 cells (Dayton, Ohio). 2010;28: 201–212. doi:10.1002/stem.278

515 20. Bilodeau S, Kagey MH, Frampton GM, Rahl PB, Young RA. SetDB1 contributes to repression of
516 genes encoding developmental regulators and maintenance of ES cell state. Genes & Development.
517 2009;23: 2484–2489. doi:10.1101/gad.1837309

518 21. McCarthy RL, Kaeding KE, Keller SH, Zhong Y, Xu L, Hsieh A, et al. Diverse heterochromatin-
519 associated proteins repress distinct classes of genes and repetitive elements. Nat Cell Biol. 2021;23:
520 905–914. doi:10.1038/s41556-021-00725-7

521 22. Nicetto D, Donahue G, Jain T, Peng T, Sidoli S, Sheng L, et al. H3K9me3-heterochromatin loss at
522 protein-coding genes enables developmental lineage specification. Science. 2019;363: 294–297.
523 doi:10.1126/science.aau0583

524 23. Biferali B, Bianconi V, Perez DF, Kronawitter SP, Marullo F, Maggio R, et al. Prdm16-mediated H3K9
525 methylation controls fibro-adipogenic progenitors identity during skeletal muscle repair. Sci Adv. 2021;7:
526 eabd9371. doi:10.1126/sciadv.abd9371

527 24. Marsano RM, Dimitri P. Constitutive Heterochromatin in Eukaryotic Genomes: A Mine of
528 Transposable Elements. Cells. 2022;11: 761. doi:10.3390/cells11050761

529 25. Wallrath LL, Rodriguez-Tirado F, Geyer PK. Shining Light on the Dark Side of the Genome. Cells.
530 2022;11: 330. doi:10.3390/cells11030330

531 26. Janssen A, Colmenares SU, Karpen GH. Heterochromatin: Guardian of the Genome. Annual review
532 of cell and developmental biology. 2018;34: 265–288. doi:10.1146/annurev-cellbio-100617-062653

533 27. Aravin AA. Pachytene piRNAs as beneficial regulators or a defense system gone rogue. Nat Genet.
534 2020;52: 644–645. doi:10.1038/s41588-020-0656-8

535 28. Yang P, Wang Y, Macfarlan TS. The Role of KRAB-ZFPs in Transposable Element Repression and
536 Mammalian Evolution. Trends in genetics : TIG. 2017;33: 871–881. doi:10.1016/j.tig.2017.08.006

537 29. Ecco G, Imbeault M, Trono D. KRAB zinc finger proteins. Development (Cambridge, England).
538 2017;144: 2719–2729. doi:10.1242/dev.132605

539 30. Bruno M, Mahgoub M, Macfarlan TS. The Arms Race Between KRAB–Zinc Finger Proteins and
540 Endogenous Retroelements and Its Impact on Mammals. Annu Rev Genet. 2019;53: 1–24.
541 doi:10.1146/annurev-genet-112618-043717

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

544 32. Benner L, Castro EA, Whitworth C, Venken KJT, Yang H, Fang J, et al. Drosophila Heterochromatin
545 Stabilization Requires the Zinc-Finger Protein Small Ovary. Genetics. 2019;213: 877–895.
546 doi:10.1534/genetics.119.302590

547 33. Jankovics F, Bence M, Sinka R, Faragó A, Bodai L, Pettkó-Szandtner A, et al. Drosophila small ovary
548 gene is required for transposon silencing and heterochromatin organization, and ensures germline stem
549 cell maintenance and differentiation. Development (Cambridge, England). 2018;145: dev170639.
550 doi:10.1242/dev.170639

551 34. Andreev VI, Yu C, Wang J, Schnabl J, Tirian L, Gehre M, et al. Panoramix SUMOylation on chromatin
552 connects the piRNA pathway to the cellular heterochromatin machinery. Nat Struct Mol Biol. 2022;29:
553 130–142. doi:10.1038/s41594-022-00721-x

554 35. Baumgartner L, Handler D, Platzer SW, Yu C, Duchek P, Brennecke J. The Drosophila ZAD zinc
555 finger protein Kipferl guides Rhino to piRNA clusters. Elife. 2022;11. doi:10.7554/elife.80067

556 36. Smolko AE, Shapiro-Kulnane L, Salz HK. An autoregulatory switch in sex-specific phf7 transcription
557 causes loss of sexual identity and tumors in the Drosophila female germline. Development (Cambridge,
558 England). 2020. doi:10.1242/dev.192856

559 37. Yang SY, Baxter EM, Doren M van. Phf7 controls male sex determination in the Drosophila germline.
560 Developmental Cell. 2012;22: 1041–1051. doi:10.1016/j.devcel.2012.04.013

561 38. Rogers RL, Shao L, Sanjak JS, Andolfatto P, Thornton KR. Revised Annotations, Sex-Biased
562 Expression, and Lineage-Specific Genes in the Drosophila melanogaster Group. G3 Genes Genomes
563 Genetics. 2014;4: 2345–2351. doi:10.1534/g3.114.013532

564 39. VanKuren NW, Vibranovski MD. A Novel Dataset for Identifying Sex-Biased Genes in Drosophila. J
565 Genom. 2014;2: 64–67. doi:10.7150/jgen.7955

566 40. Osumi K, Sato K, Murano K, Siomi H, Siomi MC. Essential roles of Windei and nuclear
567 monoubiquitination of Eggless/SETDB1 in transposon silencing. Embo Rep. 2019;20: e48296.
568 doi:10.15252/embr.201948296

569 41. Rangan P, Malone CD, Navarro C, Newbold SP, Hayes PS, Sachidanandam R, et al. piRNA
570 production requires heterochromatin formation in Drosophila. Current biology : CB. 2011;21: 1373–1379.
571 doi:10.1016/j.cub.2011.06.057

572 42. Sienski G, Dönertas D, Brennecke J. Transcriptional silencing of transposons by Piwi and maelstrom
573 and its impact on chromatin state and gene expression. Cell. 2012;151: 964–980.
574 doi:10.1016/j.cell.2012.10.040

575 43. Mohn F, Sienski G, Handler D, Brennecke J. The rhino-deadlock-cutoff complex licenses
576 noncanonical transcription of dual-strand piRNA clusters in Drosophila. Cell. 2014;157: 1364–1379.
577 doi:10.1016/j.cell.2014.04.031

578 44. Teixeira FK, Okuniewska M, Malone CD, Coux R-X, Rio DC, Lehmann R. piRNA-mediated regulation
579 of transposon alternative splicing in the soma and germ line. Nature. 2017;8: 272.
580 doi:10.1038/nature25018

581 45. Fabry MH, Ciabrelli F, Munafò M, Eastwood EL, Kneuss E, Falciatori I, et al. piRNA-guided co-
582 transcriptional silencing coopts nuclear export factors. Elife. 2019;8: e47999. doi:10.7554/elife.47999

583 46. Théron E, Maupetit-Mehouas S, Pouchin P, Baudet L, Brasset E, Vaury C. The interplay between the
584 Argonaute proteins Piwi and Aub within Drosophila germarium is critical for oogenesis, piRNA biogenesis
585 and TE silencing. Nucleic Acids Res. 2018;46: gky695-. doi:10.1093/nar/gky695

586 47. Yu Y, Gu J, Jin Y, Luo Y, Preall JB, Ma J, et al. Panoramix enforces piRNA-dependent
587 cotranscriptional silencing. Science. 2015;350: 339–342. doi:10.1126/science.aab0700

588 48. Murano K, Iwasaki YW, Ishizu H, Mashiko A, Shibuya A, Kondo S, et al. Nuclear RNA export factor
589 variant initiates piRNA-guided co-transcriptional silencing. Embo J. 2019;38: e102870.
590 doi:10.15252/embj.2019102870

591 49. Sienski G, Batki J, Senti K-A, Dönertas D, Tirian L, Meixner K, et al. Silencio/CG9754 connects the
592 Piwi-piRNA complex to the cellular heterochromatin machinery. Genes & Development. 2015;29: 2258–
593 2271. doi:10.1101/gad.271908.115

594 50. Batki J, Schnabl J, Wang J, Handler D, Andreev VI, Stieger CE, et al. The nascent RNA binding
595 complex SFiNX licenses piRNA-guided heterochromatin formation. Nat Struct Mol Biol. 2019;26: 720–
596 731. doi:10.1038/s41594-019-0270-6

597 51. Zhao K, Cheng S, Miao N, Xu P, Lu X, Zhang Y, et al. A Pandas complex adapted for piRNA-guided
598 transcriptional silencing and heterochromatin formation. Nat Cell Biol. 2019;21: 1261–1272.
599 doi:10.1038/s41556-019-0396-0

600 52. Chung H-R, Schäfer U, Jäckle H, Böhm S. Genomic expansion and clustering of ZAD-containing
601 C2H2 zinc-finger genes in Drosophila. EMBO reports. 2002;3: 1158–1162. doi:10.1093/embo-
602 reports/kvf243

603 53. Chung H-R, Löhr U, Jäckle H. Lineage-specific expansion of the zinc finger associated domain ZAD.
604 Molecular biology and evolution. 2007;24: 1934–1943. doi:10.1093/molbev/msm121

605 54. Kasinathan B, Colmenares SU, McConnell H, Young JM, Karpen GH, Malik HS. Innovation of
606 heterochromatin functions drives rapid evolution of essential ZAD-ZNF genes in Drosophila. eLife.
607 2020;9. doi:10.7554/elife.63368

608 55. Shapiro-Kulnane L, Bautista O, Salz HK. An RNA-interference screen in Drosophila to identify ZAD-
609 containing C2H2 zinc finger genes that function in female germ cells. G3 Genes Genomes Genetics.
610 2020;11: jkaa016. doi:10.1093/g3journal/jkaa016

611 56. Jauch R, Bourenkov GP, Chung H-R, Urlaub H, Reidt U, Jäckle H, et al. The zinc finger-associated
612 domain of the Drosophila transcription factor grauzone is a novel zinc-coordinating protein-protein
613 interaction module. Structure (London, England : 1993). 2003;11: 1393–1402.
614 doi:10.1016/j.str.2003.09.015

615 57. Zolotarev N, Fedotova A, Kyrchanova O, Bonchuk A, Penin AA, Lando AS, et al. Architectural
616 proteins Pita, Zw5,and ZIPIC contain homodimerization domain and support specific long-range
617 interactions in Drosophila. Nucleic Acids Res. 2016;44: 7228–7241. doi:10.1093/nar/gkw371

618 58. Maksimenko O, Kyrchanova O, Klimenko N, Zolotarev N, Elizarova A, Bonchuk A, et al. Small
619 Drosophila zinc finger C2H2 protein with an N-terminal zinc finger-associated domain demonstrates the
620 architecture functions. Biochimica Et Biophysica Acta Bba - Gene Regul Mech. 2020;1863: 194446.
621 doi:10.1016/j.bbagrm.2019.194446

622 59. Bonchuk AN, Boyko KM, Nikolaeva AY, Burtseva AD, Popov VO, Georgiev PG. Structural insights
623 into highly similar spatial organization of zinc-finger associated domains with a very low sequence
624 similarity. Structure. 2022;30: 1004-1015.e4. doi:10.1016/j.str.2022.04.009

625 60. Bonchuk A, Boyko K, Fedotova A, Nikolaeva A, Lushchekina S, Khrustaleva A, et al. Structural basis
626 of diversity and homodimerization specificity of zinc-finger-associated domains in Drosophila. Nucleic
627 Acids Res. 2021;49: 2375–2389. doi:10.1093/nar/gkab061

628 61. Shapiro-Kulnane L, Smolko AE, Salz HK. Maintenance of Drosophila germline stem cell sexual
629 identity in oogenesis and tumorigenesis. Development (Cambridge, England). 2015;142: 1073–1082.
630 doi:10.1242/dev.116590

631 62. Hinnant TD, Merkle JA, Ables ET. Coordinating Proliferation, Polarity, and Cell Fate in the Drosophila
632 Female Germline. Frontiers in cell and developmental biology. 2020;8: 19. doi:10.3389/fcell.2020.00019

633 63. Chau J, Kulnane LS, Salz HK. Sex-lethal facilitates the transition from germline stem cell to committed
634 daughter cell in the Drosophila ovary. Genetics. 2009;182: 121–132. doi:10.1534/genetics.109.100693

635 64. Chau J, Kulnane LS, Salz HK. Sex-lethal enables germline stem cell differentiation by down-
636 regulating Nanos protein levels during Drosophila oogenesis. Proceedings of the National Academy of
637 Sciences of the United States of America. 2012;109: 9465–9470. doi:10.1073/pnas.1120473109

638 65. Li Y, Minor NT, Park JK, Mckearin DM, Maines JZ. Bam and Bgcn antagonize Nanos-dependent
639 germ-line stem cell maintenance. Proceedings of the National Academy of Sciences. 2009;106: 9304–
640 9309. doi:10.1073/pnas.0901452106

641 66. Yi X, Vries HI de, Siudeja K, Rana A, Lemstra W, Brunsting JF, et al. Stwl modifies chromatin
642 compaction and is required to maintain DNA integrity in the presence of perturbed DNA replication.
643 Molecular biology of the cell. 2009;20: 983–994. doi:10.1091/mbc.e08-06-0639

644 67. Shokri L, Inukai S, Hafner A, Weinand K, Hens K, Vedenko A, et al. A Comprehensive Drosophila
645 melanogaster Transcription Factor Interactome. Cell Reports. 2019;27: 955-970.e7.
646 doi:10.1016/j.celrep.2019.03.071

647 68. Zinshteyn D, Barbash DA. Stonewall prevents expression of ectopic genes in the ovary and
648 accumulates at insulator elements in D. melanogaster. Plos Genet. 2022;18: e1010110.
649 doi:10.1371/journal.pgen.1010110

650 69. Allshire RC, Madhani HD. Ten principles of heterochromatin formation and function. Nature reviews
651 Molecular cell biology. 2018;19: 229–244. doi:10.1038/nrm.2017.119

652 70. Sarkar K, Kotb NM, Lemus A, Martin ET, McCarthy A, Camacho J, et al. A feedback loop between
653 heterochromatin and the nucleopore complex controls germ-cell to oocyte transition during Drosophila
654 oogenesis. Biorxiv. 2021; 2021.10.31.466575. doi:10.1101/2021.10.31.466575

655 71. Sarov M, Barz C, Jambor H, Hein MY, Schmied C, Suchold D, et al. A genome-wide resource for the
656 analysis of protein localisation in Drosophila. Elife. 2016;5: e12068. doi:10.7554/elife.12068

657 72. Hu Y, Comjean A, Rodiger J, Liu Y, Gao Y, Chung V, et al. FlyRNAi.org—the database of the
658 Drosophila RNAi screening center and transgenic RNAi project: 2021 update. Nucleic Acids Res.
659 2020;49: gkaa936-. doi:10.1093/nar/gkaa936

660 73. Gratz SJ, Ukken FP, Rubinstein CD, Thiede G, Donohue LK, Cummings AM, et al. Highly specific and
661 efficient CRISPR/Cas9-catalyzed homology-directed repair in Drosophila. Genetics. 2014;196: 961–971.
662 doi:10.1534/genetics.113.160713

663 74. Port F, Strein C, Stricker M, Rauscher B, Heigwer F, Zhou J, et al. A large-scale resource for tissue-
664 specific CRISPR mutagenesis in Drosophila. Elife. 2020;9: e53865. doi:10.7554/elife.53865

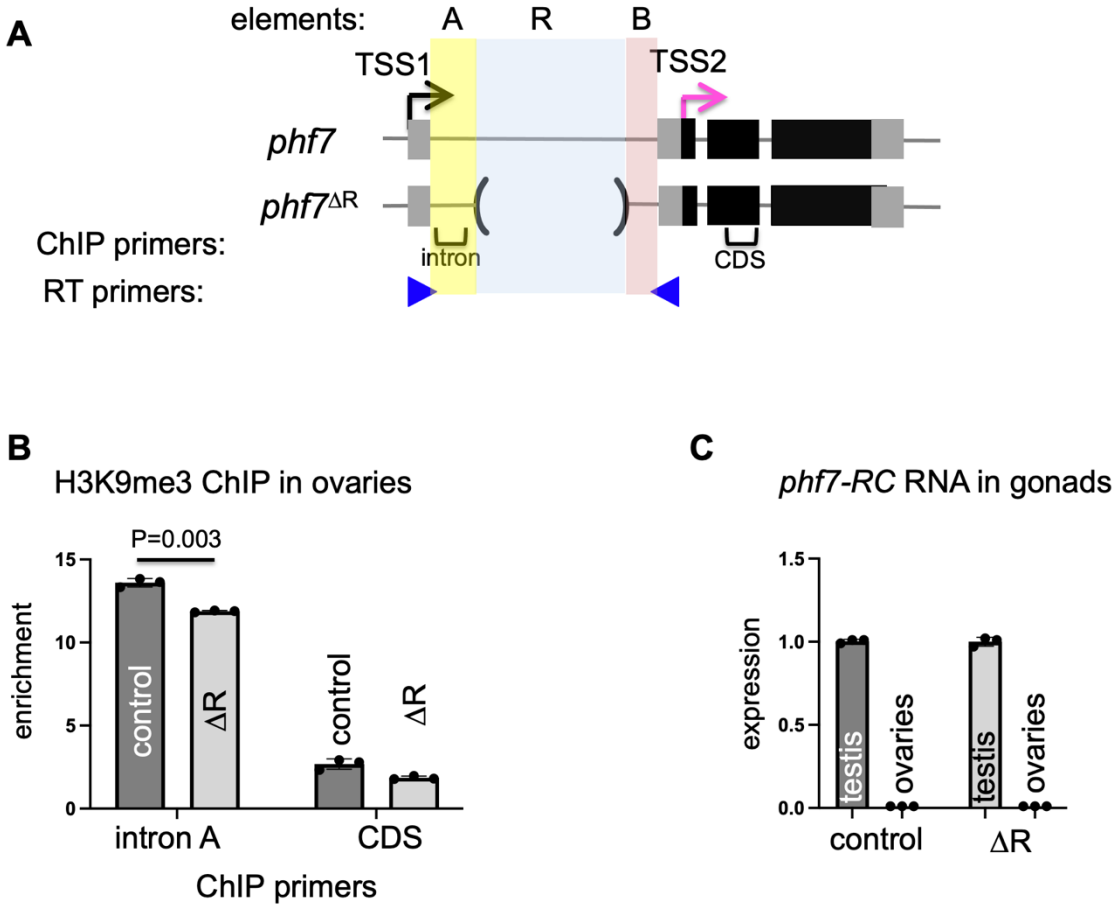
665 75. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR
666 and the 2(-Delta Delta C(T)) Method. Methods (San Diego, Calif). 2001;25: 402–408.
667 doi:10.1006/meth.2001.1262

668

669

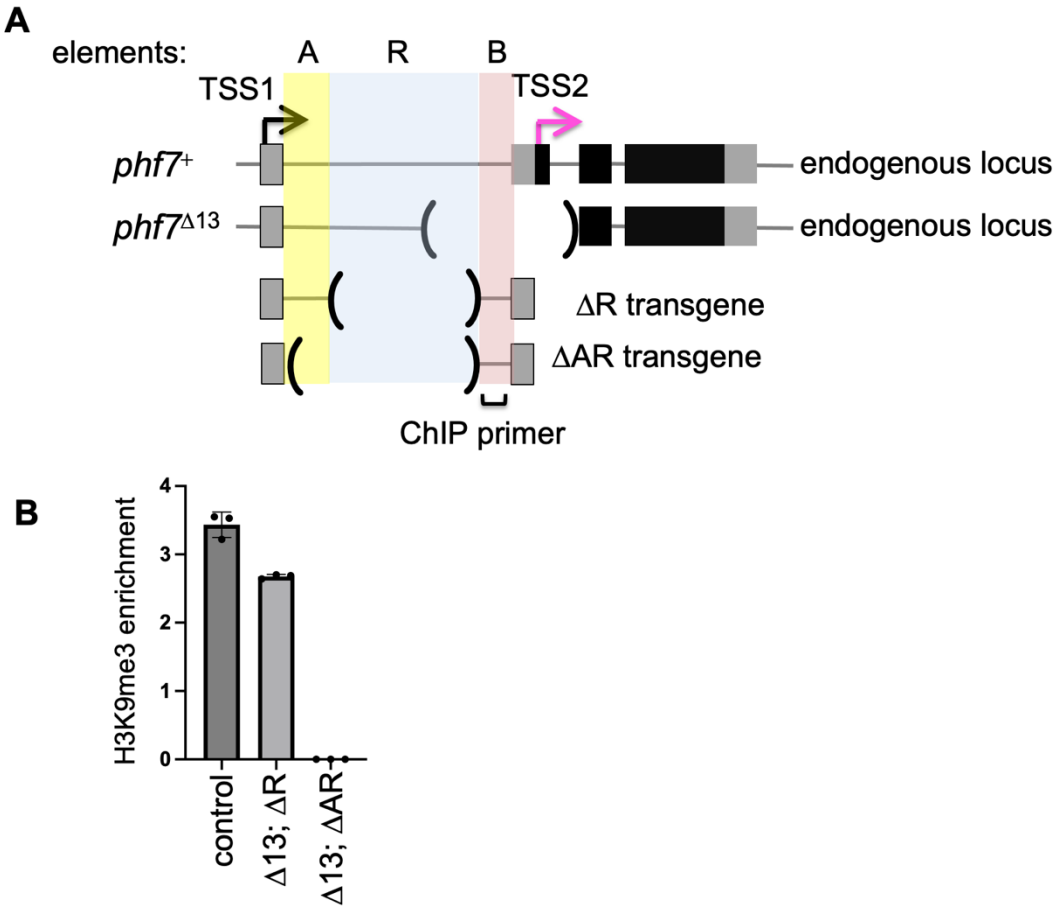
670 # Figures

671 **Fig 1. H3K9me3 deposition and sex specific transcriptional regulation are maintained in**
672 ***phf7$^{\Delta R}$* ovaries.**
673 **(**A) Diagram of the *phf7*$^{+}$ and *phf7*$^{\Delta R}$ alleles. Exons are represented by boxes, flanking DNA, and
674 introns by lines, black boxes are coding sequences, and grey boxes are untranslated regions
675 (UTRs). The sequence elements A, R, and B location within the first intron are shaded in yellow,
676 blue, and pink. In ovaries, transcription initiates from TSS2 (pink arrow), whereas an H3K9me3-
677 mediated mechanism represses transcription from TSS1 (black arrow). Locations of primers for
678 ChiP-qPCR and RT-qPCR are indicated by brackets & blue arrowheads. **(B)** H3K9me3
679 accumulation at the endogenous *phf7* locus in control (*y*^{1} *w*^{1}*)* and mutant ovaries. ChIP-qPCR
680 measured H3K9me3 signal at the *phf7* locus. ChIP to input enrichment is normalized to *rp49.*
681 **(C)** Expression of the testis specific *phf7* transcript (*phf7-RC)* in control (*y*^{1} *w*^{1}*)* and mutant
682 ovaries. Expression is measured by RT-qPCR and is shown as the fold change relative to the
683 testis. Expression is normalized to the total level of *phf7.* For both RT-qPCR and ChIP-qPCR
684 experiments. Error bars represent the standard error of the mean (SD) of three biological
685 replicates. A two-tailed Student's t-test estimates statistical significance.
686



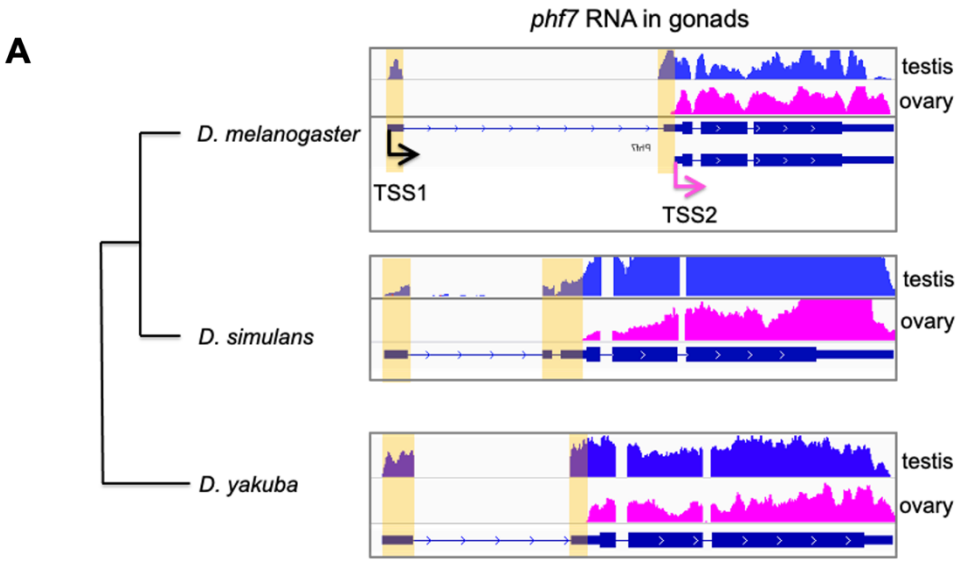
687
688

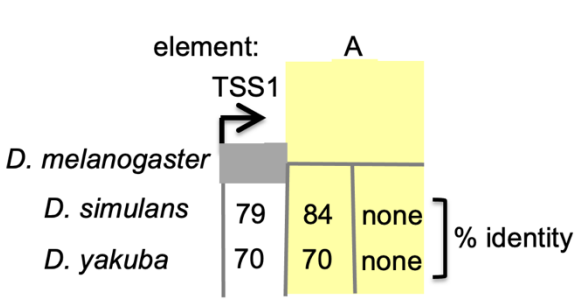
689 **Fig 2. Non-coding sequences within the first intron are sufficient for H3K9me3**
690 **deposition.**
691 **(A)** Diagram of *phf7*$^{+}$, *phf7*$^{\Delta 13}$, and the genomic transgenes. **(B)** H3K9me3 accumulation on the
692 transgenes. ChIP-qPCR measured signal, and the ChIP to input enrichment, normalized to
693 *rp49.* Error bars represent the standard error of the mean (SD) of three biological replicates.
694



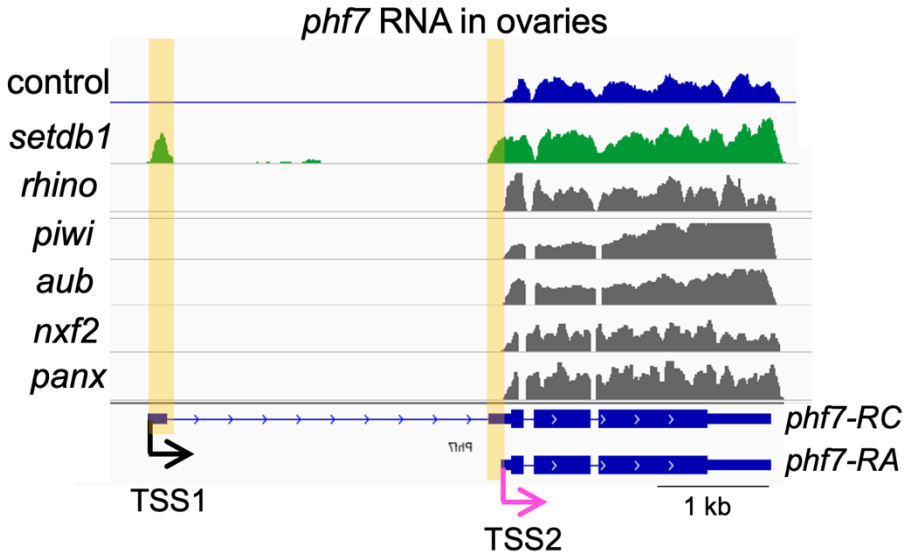
695

Fig 3. Sex specific transcriptional regulation is conserved
(A) Genome browser views of RNA-seq data aligned to the *phf7* locus from *D. melanogaster, D. simulans,* and *D. yakuba* ovaries and testis. **(B)** The first exon and a portion of the first intron are conserved. Pairwise alignments between the first exon and the adjacent intron were identified using the Basic Local Alignment Search Tool (BLASTn) available at (https://blast.ncbi.nlm.nih.gov). % identity to the Drosophila melanogaster *phf7* gene is shown.



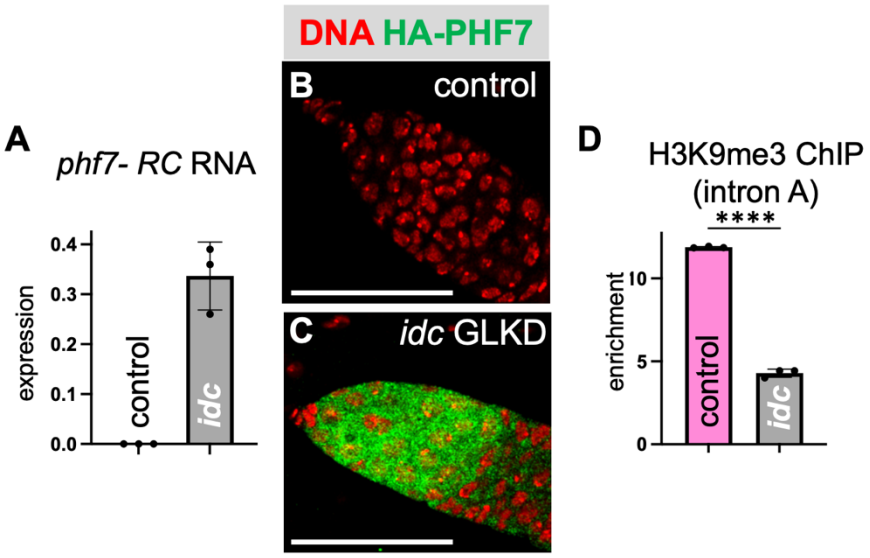


705 **Fig 4. Depletion of *setdb1,* but not the dedicated piRNA-pathway components, leads to**
706 **female-to-male reprogramming at the *phf7* locus.**
707 Genome browser views of the *phf7* locus. Tracks show RNA-seq reads from control and mutant
708 ovaries aligned to the Drosophila genome (UCSC dm6). The screenshot is reversed so that the
709 5' end of the gene is on the left. The RNA-seq reads that are unique to *setdb1* mutant ovaries
710 are highlighted. In wild-type control ovaries, only the *phf7*-RA transcript is detected (TSS2, pink
711 arrow), whereas loss of *setdb1* leads to ectopic expression of the testis specific *phf7-RC* isoform
712 (TSS1, black arrow). In contrast, mutations in the *rhino, piwi, aub, nxf2* and *panx* genes do not
713 disrupt sex specific *phf7* transcription. The complete mutant genotypes and accession numbers
714 of the RNA-seq data sets are listed in Table S3.
715

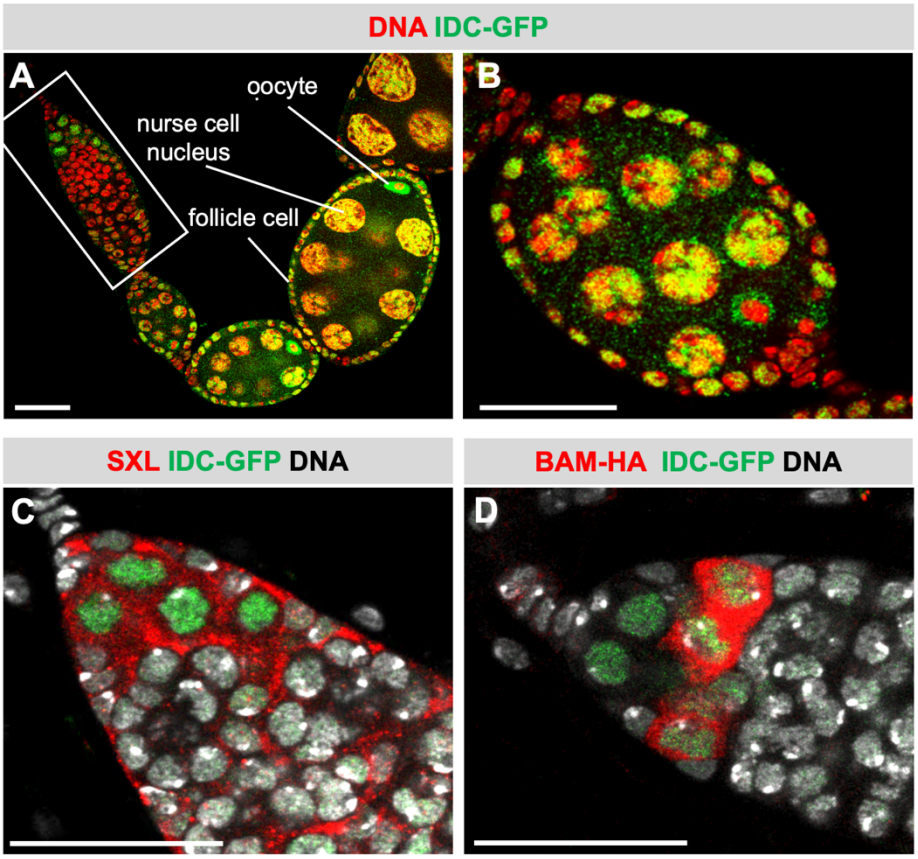


716

Fig 5. *idc* germ cell specific knock-down disrupts *phf7* regulation.
(A) *idc* GLKD germ cells (*nos>idc-RNAi*) express the testis *phf7-RC* transcript. RT-qPCR measures expression normalized to the total level of *phf7* in RNA extracted from control (*nos>white-RNAi*) and *idc* GLKD (*nos>idc-RNAi*) ovaries. The histograms show the mean ± SD of three biological replicates. **(B, C)** *idc* GLKD germ cells inappropriately express the PHF7 protein. Confocal images of ovaries were dissected from control (*nos>white-RNAi*) and mutant (*nos>idc-RNAi*) females carrying an HA-PHF7 transgene and stained for HA (green) and DNA (red). Scale bar 50μm. **(D)** H3K9me3 accumulation at the endogenous *phf7* locus is reduced in *idc* GLKD ovaries. ChIP-qPCR measured H3K9me3 signal at the *phf7* locus (primer in region A) in control (*nos>white-RNAi)* and mutant (*nos>idc-RNAi*) ovaries. ChIP to input enrichment is normalized to *rp49.* The histogram shows the mean ± SD of three biological replicates. A two-tailed Student's t-test estimates statistical significance where ****p<0.0001.

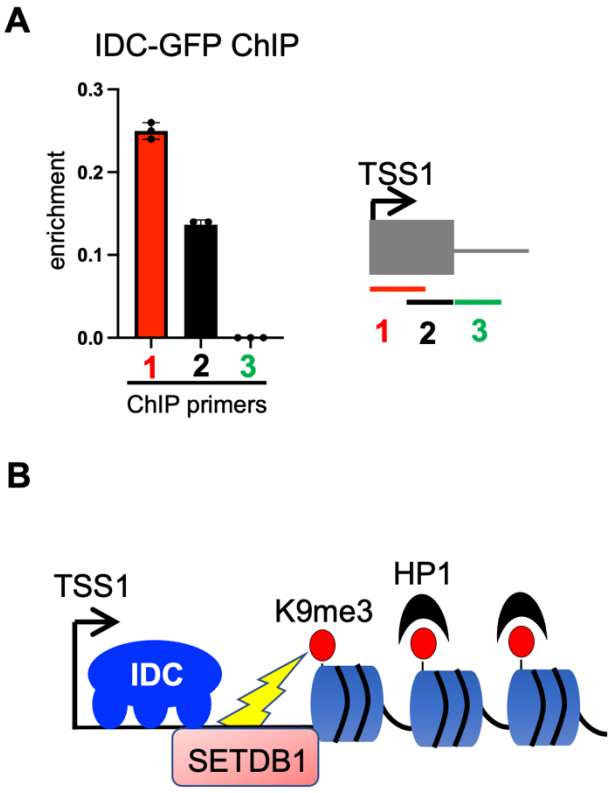


732
733
734 **Fig 6. IDC is nuclear protein expressed in both somatic and germline cells.**
735 **(A)** Confocal image of an ovariole dissected from a female carrying the IDC-GFP rescue
736 transgene stained for GFP (green) and DNA (red). The germarium at the anterior end of the
737 ovariole is boxed in. An oocyte, a nurse cell, and a somatic follicle cell are indicated. Scale bar
738 25μm. **(B)** Confocal image of an early-stage egg chamber dissected from a female carrying the
739 IDC-GFP transgene stained for GFP (green) and DNA (red). Scale bar 25μm. **(C)** Confocal
740 image of the anterior end of a germarium dissected from a female expressing the IDC-GFP
741 fusion protein and co-stained for GFP (green), DNA (white), and SXL (red). Scale bar 25 μm.
742 **(D)** Confocal image of the anterior end of a germarium dissected from a female expressing the
743 IDC-GFP and the BAM-HA fusion proteins co-stained for GFP (green), DNA (white), and HA
744 (red). Scale bar 25μm.
745



746

747
748 **Fig 7. IDC binds to *phf7.***
749 **(**A) ChIP-qPCR shows that IDC binds to the *phf7* first exon. Left: The histogram shows the
750 mean ± SD of three biological replicates. Right: Cartoon showing the location of the primers
751 used for ChIP-qPCR. **(B)** Model of H3K9me3-mediated silencing of *phf7.* IDC binds to *phf7,*
752 directing H3K9me3 deposition by SETDB1 H3K9 methytransferase. HP1 binds to H3K9me3,
753 resulting in transcriptional silencing.
754



755
756

Figure S1. The first intron of *phf7* contains seven copies of an ~250 bp DNA sequence
Sequence alignment of the 7 ~250 bp sequences generated with the multiple sequence alignment tool MUSCLE embedded in the SnapGene software (snapgene.com).

Figure S2 The first exon and a portion of the first intron are conserved between *D. melanogaster, D. simulans* and *D. yakuba.* Pairwise alignments between the first exon and the adjacent intron were identified using the Basic Local Alignment Search Tool (BLASTn) available at (https://blast.ncbi.nlm.nih.gov).

Query: *D. melanogaster phf7* intron compared to *D. simulans* intron

Score	Expect	Identities	Gaps	Strand
224 bits(248)	1e-62	179/213(84%)	7/213(3%)	Plus/Minus

```
Query  1     GTAAGTTTTCATAATGAAATAGCATTGTGCAATGAATTGTAACACAAATACATATATCTA  60
             ||||||||| ||| |||| || ||||||||||| |||||| ||||||||||  |||||
Sbjct  1032  GTAAGTTTTTATAGTGAATTACCATTGTGCAAAGAATTGAAACACAAATATATATA----  977

Query  61    CTTGAAAAGGCATTAATTAGAACCCCTATACACATACTTTGAAAATAAAGGACTGCAACT  120
             ||||||||| ||| | || ||||  |||||||| ||||||||| ||||||||||||||
Sbjct  976   CTTGAAAAGGAATTTAATAAAACC--TATACACACACTTTGAAAAAAAAGGACTGCAACT  919

Query  121   TTCATTCAATTCGCATTTGCCCCAGGTTACGGCAATTTTCG-AGGAAATTCGTAGAAACA  179
             | ||||||||||| |||||||| |  ||| |  |||||||| | ||| |||||||||||| |
Sbjct  918   TCCATTCAATTCCCATTTGCCGCGAGTTGCACCAATTTTTGAAGGGAATTCGTAGAAATA  859

Query  180   GAAATAAAACCGAGAGTTTGACGCTTGCCAACT  212
             ||||||||| ||| | ||||||||||||||||||
Sbjct  858   GAAATAAAGCCGCGCGTTTGACGCTTGCCAACT  826
```

Query: *D. melanogaster phf7* intron compared to *D. yakuba* intron

Range 1: 1 to 223 Graphics ▼ Next Match ▲ Previou

Score	Expect	Identities	Gaps	Strand
101 bits(111)	3e-25	169/241(70%)	24/241(9%)	Plus/Plus

```
Query  1    GTAAGTTTTCATAATGAAATAGCATTGTGCAATGAATTGTAACACAAATACATATATCTA  60
            ||||||||||||  ||| | ||| |||||||||||||| ||| ||||||           ||
Sbjct  1    GTAAGTTTTCATAGAGAATTGGCAATGTGCAATGAATTGGAACGCAAATA--------TA  52

Query  61   CTTGAAAAGGCATTAATTAGAACCCCTATACACATACTTTGAAAATAAAGGAC--TGCAA  118
            | ||| |||||  | ||| |||||  || | |||||| | || |||||  | |||||  | |||
Sbjct  53   TTTGTAAAG--ACTAAATAGA--CCTTTTACACAAATTTAGAAAAGGAGGGACTTTCCAA  108

Query  119  CTTTCATTCAATTCGCATTTGCCCCAGGTTACGGC-AATTTTCGAGGAAATTC---GTAG  174
            ||| | |||        |||||   | ||| | || | |||||| | ||  |||||    | |
Sbjct  109  CTTGCTTTC------TATTTCTTCAAGGCTGCGCCAAATTTCCTAGTGAATTCTAAGAAA  162

Query  175  AAACAGAAATAAAACCGAGAGTTTGACGCTTGCCAACTCTGGCAGTAGGGTTGCCCCTAA  234
            ||| | |||||||||| | | ||||| ||| |  |||||||||||||||||| || || |
Sbjct  163  AAATACAAATAAAACTGTGCGTTTGGCGCATAGCAACTCTGGCAGTAGGGTTCCCTCTCA  222

Query  235  G  235
            |
Sbjct  223  G  223
```

Query: *D. melanogaster phf7* 1st exon compared to *D. simulans* 1st exon

Range 1: 31 to 199 GenBank Graphics ▼ Next Match ▲ Previous

Score	Expect	Identities	Gaps	Strand
135 bits(149)	1e-35	137/173(79%)	8/173(4%)	Plus/Plus

```
Query  5    AGCTCACAGGTCAGAGAAGAAATTTTCGGAATaaaaaaaTAAA---TATAACAAGTTTTC  61
            ||| |||||| || |||||| |||||||||||||||||||| | |   | ||| |||||||
Sbjct  31   AGCACACAGCTCGGAGAATAAATTTTCGGAATAAAAAAAAATACTTTGTAA-AAGTTTA-  88

Query  62   AAAAATAACTCCGCACCAAACCCATTTAACCAAATCAAAAGCTTAAAGCAGTGAATAGTG  121
            |||||||  ||  |  ||||||||||||||| ||| | |||||| ||||||||||||||||
Sbjct  89   AAAAATAGTTCACC--CAAACCCATTTAACAAAAACCAAAGCATAAAGCAGTGAATAGTG  146

Query  122  TTATTTTTAAAAATTTAGCAC-AAATAAAAAGTTCGGGAATTCAACGCTTTTT  173
              |||| | |||||||| ||| ||| || ||||||  |||||||| | ||| |||
Sbjct  147  CCATTTATCAAAATTTACCACAAAACAACAAGTTTTGGAATTCCATGCTGTTT  199
```

Query: *D. melanogaster phf7* 1st exon compared to *D. yakuba* 1st exon

Range 1: 34 to 174 GenBank Graphics ▼ Next Match ▲ Previous

Score	Expect	Identities	Gaps	Strand
64.4 bits(70)	1e-14	103/147(70%)	11/147(7%)	Plus/Plus

```
Query  9    CACAGGTCAGAGAAGAAATTTTCGGAATaaaaaaaTAAA-----TATAACAAGTTTTCAA  63
            ||||| |||| |||| | |     ||| |||||| || |     | | ||||||||| ||
Sbjct  34   CACAGCTCAGTGAAGTACT-----GAAAAAAAAATTATACTTTTTTTTACAAGTTTTTAA  88

Query  64   AAATAACTCCGCACCAAACCCATTTAACCAAATCAAAAGCTTAAAGCAGTGAATAGTGTT  123
            ||  |   ||| || ||| |  | ||| ||||||||||||| || |||| |||| ||| |
Sbjct  89   AATTTT-TCCACATCAACTCTATTGAACCAAATCAAACGCATAAAACAGTAAATTTTACC  147

Query  124  ATTTTTAAAAATTTAGCACAAATAAAA  150
            || ||| ||||||||| |||| | ||||
Sbjct  148  ATATTTCAAAATTTACCACACAAAAAA  174
```

778 **Table S1 Primers**
779
780 For RT-qPCR experiments
781

name	forward	reverse
phf7-RC	AGTTCGGGAATTCAACGCTT	GAGATAGCCCTGCAGCCA
total *phf7*	GAGCTGATCTTCGGCACTGT	GCTTCGATGTCCTCCTTGAG
rp49	ATCGGTTACGGATCGAACAA	GACAATCTCCTTGCGCTTCT

782
783 For ChIP-qPCR experiments

name	forward	reverse
Intron element A (Fig 1 & 5)	GAGAGTTTGACGCTTGCCAA	CATGGCGGGTTTCAGTTGAA
CDS (Fig 1)	CTGCGACACGTACAACTCGT	AGCGGGCACTTGAAGAAGTA
Rp49 (Fig 1, 2 & 5)	ATCGGTTACGGATCGAACAA	GACAATCTCCTTGCGCTTCT
Intron element B (Fig 2)	AAATTCCCGATTTGCTTGTG	GCGCAGCGATTGAATGTTA
1 (Fig 7)	CAGAAGCTCACAGGTCAGAG	TAAATGGGTTTGGTGCG
2 (Fig 7)	CTCCGCACCAAACCCATTTA	AGCGTTGAATTCCCGAACTT
3 (Fig 7)	GAAATAGCATTGTGCAATGAAT	GCAAATGCGAATTGAATGAAAGT

784
785
786

787
788 **Table S2 Buffers for ChIP**
789

Buffer	Components
PBS (phosphate buffered saline) pH 7.2	137mM sodium chloride
	2.7mM potassium chloride
	10mM sodium phosphate dibasic
	1.8mM potassium phosphate monobasic
PBSP	PBS
	cOmplete mini EDTA free Protease Inhibitor Cocktail tab (Sigma cat# 11836170001)
Lysis Buffer 1	50mM HEPES-KOH (Thermo Fisher cat# 15630080)
	140mM sodium chloride
	1mM EDTA (Thermo Fisher cat#AM9261)
	10% glycerol
	0.5% Igepal (NP-40)
	0.25% Triton X-100
	cOmplete mini EDTA free Protease Inhibitor Cocktail tab (Sigma cat# 11836170001)
Lysis Buffer 2	10mM TRIS pH 8 (Sigma cat# T1503)
	200mM sodium chloride
	1mM EDTA (Thermo Fisher cat#AM9261)
	0.5mM EGTA
	cOmplete mini EDTA free Protease Inhibitor Cocktail tab (Sigma cat# 11836170001)
Lysis Buffer 3	10mM TRIS pH 8 (Sigma cat# T1503)
	100mM sodium chloride
	1mM EDTA (Thermo Fisher cat#AM9261)
	0.5mM EGTA
	0.1% sodium deoxycholate (Sigma cat# D6750)
	0.5% N-lauroylsarcosine (Sigma cat# L7414)
	cOmplete mini EDTA free Protease Inhibitor Cocktail tab (Sigma cat# 11836170001)
ChIP-RIPA	10mM TRIS pH 8 (Sigma cat# T1503)
	1mM EDTA (Thermo Fisher cat#AM9261)
	140mM sodium chloride
	1% Triton X-100
	0.1% sodium dodecyl sulfate (Fisher cat# BP-1311-200)
	0.1% sodium deoxycholate (Sigma cat# D6750)
	cOmplete mini EDTA free Protease Inhibitor Cocktail tab (Sigma cat# 11836170001)
ChIP-RIPA/500	ChIP-RIPA
	500mM sodium chloride
	cOmplete mini EDTA free Protease Inhibitor Cocktail tab (Sigma cat# 11836170001)
ChIP-LiCl	10mM TRIS pH 8 (Sigma cat# T1503)
	1mM EDTA (Thermo Fisher cat#AM9261)
	250mM Lithium Chloride

	0.5% Igepal (NP-40) 0.5% sodium deoxycholate (Sigma cat# D6750) cOmplete mini EDTA free Protease Inhibitor Cocktail tab (Sigma cat# 11836170001)
TE	10mM TRIS pH 8 (Sigma cat# T1503) 1mM EDTA (Thermo Fisher cat#AM9261)
Elution Buffer	10mM TRIS pH 8 (Sigma cat# T1503) 5mM EDTA (Thermo Fisher cat#AM9261) 300mM sodium chloride 0.1% sodium dodecyl sulfate (Fisher cat# BP-1311-200)

790
791

792

793 **Table S3 The published RNA-seq data sets utilized in Figure 4**

Genotype	SRA Run ID	Reference
wild-type (y^1, w^1)	SRR6742503	Smolko et al., 2018 [9]
setdb1 germline RNAi knockdown	GSM2971276	Smolko et al., 2018 [9]
rhino germline RNAi knockdown	GSM3285257	Mohn et. al., 2014 [43]
aub^{HN2}/aub^{QC42}	GSM3284029	Teixeira et al., 2017 [44]
$piwi^1$/$piwi^2$	GSM3284031	Teixeira et al., 2017 [44]
$nxf2^{delta1}$/$nxf1^{F10}$	GSM3443092	Fabry et al., 2019 [45]
$panx^{M1}$/$panx^{M4}$	GSM3443094	Fabry et al., 2019 [45]

794

795

796 **Table S4 The published RNA-seq data sets utilized in Figure 3**

Species	Tissue	SRA Run ID	Reference Genome	Reference
D. melanogaster	ovary	SRR6742503	dm6	Smolko et al., (2018) [9]
D. melanogaster	testis	SRR070422	dm6	http://modeencode.org
D. simulans	ovary	SRR1511609	ASM175419v2 (GCF_000754195.2)	Rogers et al., (2014) [38]
D. simulans	testis	SRR1520537	ASM175419v2 (GCF_000754195.2)	Rogers et al., (2014) [38]
D. yakuba	ovary	SRR1024009	dyak_caf1 (GCF_000005975.2)	VanKuren and Vibranovski (2014) [39]
D. yakuba	testis	SRR1024006	dyak_caf1 (GCF_000005975.2)	VanKuren and Vibranovski (2014) [39]

797
798