Germline piRNAs counteract endogenous retrovirus invasion from somatic cells.

Marianne Yoth[1], Nathalie Gueguen[1], Silke Jensen[1]*, Emilie Brasset[1]*

[1] iGReD, Université Clermont Auvergne, CNRS, INSERM, Faculté de Médecine, 63000 Clermont-Ferrand, France.

* Correspondence: emilie.brasset@uca.fr, silke.jensen@uca.fr

Abstract

In metazoan gonads, transposable elements (TEs) mobilization is limited by PIWI-interacting RNAs (piRNAs). These small RNAs originate from specific source loci, the piRNA clusters. piRNAs are known to silence TEs in the cells where they are produced. Endogenous retroviruses (ERVs), a subclass of TEs, pose a particular threat because they are capable of transiting from cell to cell. In this study, we reveal that piRNAs produced locally in germ cells counteract invasion by ERVs arriving from adjacent somatic cells. We reactivated the Drosophila ERV *ZAM* in somatic gonadal cells by deleting, using CRISPR-Cas9 genome editing, its single copy in the somatic *flamenco* piRNA cluster, while keeping the piRNA pathway fully functional. Upon reactivation, *ZAM* invaded the oocytes, resulting in transposition and severe fertility defects. We show that once *ZAM*-piRNAs are produced in germ cells they counter the invasion. Our study sheds new light on the mechanisms of recognition and regulation of invasive genetic elements, which is essential for the maintenance of genome integrity.

60 # Introduction

61 The gonad is a major organ in animals because of its role in producing germ cells, the only cells
62 that can transfer genetic information to the next generation. Germ cells require a set of
63 supporting somatic cells, called gonadal somatic cells. Studies in model organisms, such as
64 *Caenorhabditis elegans*, mice and Drosophila, revealed the essential role of somatic cells in
65 regulating germ cell development [1–5].

66 Transposable elements (TEs), discovered by Barbara McClintock in the 1940s [6], are mobile
67 genetic entities that are present in nearly all organisms. Genome integrity relies on
68 transcriptional and post-transcriptional repression of TE activity. In animal gonads, the genome
69 defense against TEs is based essentially on the PIWI-interacting RNA (piRNA) pathway.
70 piRNAs, the central players in this regulatory pathway, are small (23 to 29 nucleotides) RNAs
71 that interact with proteins of the PIWI family [7–9]. They originate from transposon-rich genomic
72 loci called piRNA clusters [10–12]. In Drosophila ovaries, there are significant differences of TE
73 control between germline and somatic cells. Transposons have been classified as "germline",
74 "somatic" or "intermediate" TEs, according to the origin of the piRNAs that target them [13].
75 Indeed, there are germline- and soma-expressed piRNA clusters. In germ cells, most piRNA
76 clusters are expressed in a bidirectional manner and produce piRNAs from both genomic
77 strands. Other clusters are specifically expressed in the somatic cells that surround the ovary in
78 a unidirectional manner and produce piRNAs that are predominantly antisense to transposons
79 [10]. Nonetheless, the relationship between piRNA clusters and TE control is incompletely
80 understood. Surprisingly, while depletion of piRNA pathway components in the germline or
81 somatic lineage leads to TE derepression and sterility, the deletion of the three largest germline
82 piRNA clusters (*42AB*, *38C*, and *20A*) does not have any impact on transposon regulation or on
83 fertility [14–16]. Conversely, *flamenco (flam)*, the major somatic piRNA cluster, is directly

84 involved in the regulation of at least three TEs: *ZAM*, *Idefix,* and *Gypsy* [17–19] and female fertility
85 is affected when *flam* is mutated [20–22].
86 In the RevI-H2 line, in which the proximal part of *flam* is deleted (i.e., the region containing its
87 only *ZAM* insertions), *ZAM* is derepressed in the somatic cells of the ovary. We previously
88 showed that *ZAM* somatic piRNA production is abolished, whereas germline sense and
89 antisense *ZAM*-derived piRNAs are newly produced [23,24]. As *Gypsy, ZAM* is an endogenous
90 retrovirus (ERV), a class of TEs capable of transiting from cell to cell. However, whether and
91 how germline piRNAs contribute to protect the germline against invasion of TEs coming from
92 the surrounding somatic cells had not been investigated yet.
93
94 Here we report, using *ZAM* as a model, that when an ERV is derepressed in the somatic cells
95 surrounding germ cells, it invades the germline. Furthermore, no initial small RNA response
96 takes place after invasion of the germ cells. Many years after the *flam* partial deletion appeared
97 in the RevI-H2 line, *ZAM* is still expressed in somatic cells that surround germ cells and no new
98 *ZAM* insertion in the *flam* piRNA cluster has been detected. However, once the germline
99 produces piRNAs against this new invader, either from newly inserted dispersed copies or
100 because the invading TE has jumped into a piRNA cluster expressed in the germline, invasion
101 is arrested while the retrotransposon is still activated in the surrounding tissue. Our findings
102 reveal the critical role of the germline piRNA pathway as an effective defense mechanism to
103 counteract invasion by TEs from an adjacent tissue. Our work provides a unique opportunity to
104 better understand the role of each piRNA cluster type in TE regulation at the cellular and tissue
105 scales.
106
107

108 # Results

109 **In an ancient *flam* mutant line, *ZAM* is still expressed and no new *ZAM* copy is inserted**
110 **in the *flam* locus.**

111

112 In the RevI-H2 line, in which a deletion of the proximal part of *flam* eliminats the region in
113 which *ZAM* was inserted, *ZAM* piRNA production by *flam* is completely lost in follicle cells,
114 which are somatic cells of the ovary [23]. RevI-H2i2, RevI-H2i3 and RevI-H2i4 are three isogenic
115 lines derived from the RevI-H2 line that has more than 25 years of laboratory history [25]. *ZAM*
116 actively transposed in the initial RevI-H2 line, but it was not known whether *ZAM* is still
117 transcriptionally active in the RevI-H2-derived lines. We found that these three isogenic RevI-
118 H2 lines have a very similar *ZAM* expression pattern (**Fig 1a, Supplementary Fig. S1a**).
119 Indeed, single-molecule RNA fluorescence *in situ* hybridization (smRNA FISH) experiments
120 revealed that *ZAM* RNA was produced very early during oogenesis in the somatic cells of the
121 germarium. Then, after stage 8, *ZAM* transcripts accumulated in a patch of follicle cells located
122 at the posterior side of the follicles (**Fig. 1a**). We did not detect any *ZAM* staining in germ cells
123 (nurse cells or oocytes). RNase A treatment led to complete loss of *ZAM* staining in the follicle
124 cells (**Supplementary Fig. S1b**). We did not detect any *ZAM* RNA in the Iso1A, w^{1118} and w^{IR6}
125 control ovaries (**Supplementary Fig. S1c**). Then, we used RT-qPCR to confirm that *ZAM* was
126 derepressed in the RevI-H2i2 ovaries. Conversely, other soma-specific TEs that are not
127 included in the deletion, such as *Gypsy*, and germline-specific TEs, such as *Burdock*, were not
128 upregulated (**Fig. 1b**). As *ZAM* is an endogenous retrovirus that encodes Gag, Pol and Env
129 proteins, we showed by immunostaining that Gag and Env proteins accumulated in the posterior
130 follicle cells in late follicles (> stage 9) (**Fig. 1c**). All these results showed that *ZAM* is still
131 expressed in the RevI-H2i2 line. The long reads obtained by Oxford Nanopore Technology
132 (ONT) sequencing of the RevI-H2i2 genome allowed identifying *ZAM* insertions, even in

133 repeated sequences, but no *ZAM* insertion was present in the *flam* piRNA cluster (**Fig. 2a**).
134 Thus, no *ZAM* piRNA is produced from *flam*, the major somatic piRNA source locus,
135 explaining why *ZAM* was still expressed in follicle cells in the RevI-H2i2 line. Although *ZAM*
136 was expressed in follicle cells, the RevIH2i2 line is fertile. Indeed, the ovary morphology and
137 the number of eggs laid were comparable in a RevIH2i2 and a control female (**Fig. 1d, e**).
138 Altogether, these results suggest that *ZAM* expression in follicle cells has no major deleterious
139 effect in the RevI-H2i2 line.
140
141 **RevI-H2i2 is a stable line although *ZAM* is actively expressed in follicle cells.**
142
143 The ONT long-read sequencing data allowed identifying 18 *ZAM* copies in the RevI-H2i2
144 genome (**Fig. 2a**). Only one of these *ZAM* copies was also present in the reference genome
145 (http://flybase.org). The other 17 new insertions were very similar to the *ZAM* reference element
146 identified in the initial RevI-H2 line [25]. Specifically, nine were full-length *ZAM* elements, one
147 harbored a deleted 5'-UTR, three had large internal deletions in the coding regions, and five
148 were what we called "*ZAM*-variants", i.e. full-length insertions with a small deletion at the C-
149 terminal end of the *pol* gene. This small deletion corresponds to positions 5500 to 6126 in the
150 *ZAM* internal sequence [25] (Repbase *ZAM*_I, https://www.girinst.org/repbase/ [26]). This deleted
151 part of the *pol* gene does not correspond to any known protein domain and the fact that the five
152 *ZAM*-variants were at different locations in the RevI-H2i2 line genome indicates that this *ZAM*
153 copy is competent for transposition. The full-length copies and ZAM-variants all potentially
154 encode Gag, Pol and Env proteins. We could clearly localize 14 of the 17 new *ZAM* insertions
155 (**Fig. 2a** and Supplementary **Table S1** for coordinates). In the initial RevI-H2 line, a *ZAM*
156 insertion was found in the germline piRNA cluster *9*, which is close to the X chromosome
157 centromere [12,23]. Here, using ONT sequencing data, we identified three *ZAM* insertions in

158 piRNA cluster *9* (one full-length and two *ZAM*-variants), and also a full-length *ZAM* insertion,
159 which is localized in either piRNA cluster *13* or cluster *56* (uncertainty is due to flanking
160 repeats, piRNA clusters as in [12]).
161 Comparison of the Illumina sequencing data for the initial RevI-H2 line, performed ten years
162 ago [23] , and the ONT sequencing data showed that at least 14 of the new *ZAM* insertions in the
163 RevI-H2i2 line, including all *ZAM* insertions in piRNA clusters, were already present ten years
164 ago. This result indicates that the genomic *ZAM* profile is the same as ten years ago and suggests
165 that no new *ZAM* insertion occurred since then in the germline. Therefore, the RevI-H2 line is
166 stable.
167
168 **Germline *ZAM* piRNAs counteract *ZAM* invasion from adjacent somatic cells.**
169
170 The next step was to determine the mechanisms of the RevI-H2i2 line stability despite *ZAM*
171 expression in follicle cells. It was previously reported [27] that the *ZAM* TE actively transposed
172 in the RevI-H2 line. These transposition events occurred in the germline, as attested by the
173 transmission of the *ZAM* insertions to the offspring despite the fact that *ZAM* is specifically and
174 exclusively expressed in the somatic follicle cells. This indicates that *ZAM* invades the RevI-
175 H2 germline from follicle cells. We showed previously that germ cells of RevI-H2 flies produce
176 high amounts of *ZAM*-derived piRNAs with a ping-pong signature (i.e. an enrichment of 10-
177 nucleotide overlap between sense and antisense *ZAM*-derived piRNAs). One possible source of
178 these germline *ZAM*-targeting piRNAs is a *ZAM* copy inserted in the germline piRNA cluster
179 *9* (piRNA clusters as referenced in [12]).
180 First, we checked whether the RevI-H2i2 line harbors the same characteristics. We found that
181 the RevI-H2i2 line produced sense and antisense *ZAM*-derived-piRNAs with a ping-pong
182 signature and silenced a *ZAM* sensor transgene in germ cells (**Fig. 2b, c, d**). We obtained similar

183 results in the RevI-H2i3 and RevI-H2i4 lines (**Supplementary Fig. S2a, b, c**). Moreover, in all
184 three RevI-H2 isogenic lines the *ZAM* copy was deleted in the *flam* piRNA cluster and *ZAM*
185 was inserted in the germline piRNA cluster 9 (**Supplementary Fig. S2d**).
186 Then, to test whether piRNAs produced in germ cells can thwart *ZAM* invasion from somatic
187 cells to the germline, we abolished the piRNA pathway in the germ cells by germline-specific
188 knock-down (GLKD), in the RevI-H2i2 genetic background, and then followed *ZAM* RNA
189 subcellular localization in ovaries. We first confirmed the loss of *ZAM*-targeting piRNAs. We
190 observed a decrease in the level of antisense piRNAs that map to *ZAM,* on the same extend as
191 intermediate and germ-cell-specific TEs, in ovaries in which *argonaute3* (*ago3*), *zucchini* (*zuc*),
192 or *piwi* (three genes encoding piRNA pathway components) was knocked down in the RevI-
193 H2i2 genetic background (**Fig. 3a**). No decrease was observed for *Gypsy10*, a somatic TE.
194 Furthermore, in *ago3*-GLKD ovaries, the ping-pong signature for *ZAM* and for many other
195 germline-specific TEs was abolished, while it was maintained in control *white*-GLKD and also
196 in *zuc*- and *piwi*-GLKD ovaries (**Fig. 3b, c**).
197
198 We then analyzed the subcellular localization of *ZAM* RNAs by smRNA FISH in the *white*-
199 GLKD RevI-H2i2 control line in which *ZAM*-derived piRNAs are produced in the germline,
200 but not in follicle cells. *ZAM* RNA expression was restricted to follicle cells and almost no *ZAM*
201 RNA signal was detectable in the oocyte (**Fig. 3d**). Conversely, all RevI-H2i2 lines with GLKD
202 of piRNA pathway actors (*vreteno (vret)*, *zuc, ago3*, *armitage (armi*)) clearly showed *ZAM*
203 expression in posterior follicle cells and *ZAM* RNA invasion in oocytes. *ZAM* RNA spread
204 throughout the ooplasm, but was clearly enriched at the posterior pole, adjacent to *ZAM*-
205 expressing follicle cells. Actually, 30-70% of stage 10 follicles showed a strong *ZAM* RNA
206 signal in oocytes (**Fig. 3d, e).** The strongest invasion phenotype was observed for the RevI-
207 H2i2 *ago3*-GLKD condition where we found a drastic decrease in the production of *ZAM*-

208 derived piRNAs (**Fig. 3a**). In this condition, *ZAM* RNA accumulation in the ooplasm was
209 correlated with a strong increase of *ZAM* RNAs in total ovaries (RT-qPCR data) (**Fig. 3f**). These
210 results showed that when *ZAM*-derived piRNAs are absent or downregulated in the germline,
211 *ZAM* RNAs transit from the somatic follicle cells, where they are produced, to the oocyte.
212 Altogether, our data strongly suggest that *ZAM* piRNAs produced by the germline piRNA
213 pathway trigger post-transcriptional silencing of *ZAM* RNAs arriving from follicle cells. In line
214 with this conclusion no *ZAM* RNA was detected in stage 10 ooplasm and *ZAM* expression was
215 not affected when Piwi, which acts thought transcriptional gene silencing, was knocked down
216 in the RevI-H2i2 germline (**Fig. 3e, Supplementary Fig. S3a, b**).
217 We confirmed this finding using a line where no *ZAM*-derived piRNA is produced in the
218 germline (w^{1118} genetic background) and in which we knocked down the piRNA pathway in
219 follicle cells by somatic knock-down (sKD) of *armi*, *vret*, *piwi* or *yb*. sKD of a piRNA pathway
220 actor, such as *vret* or *yb,* led to a strong decrease in *ZAM* piRNA production (**Fig. 4a,**
221 **Supplementary Fig. S4a**). The amount of antisense Piwi-bound piRNAs that map to soma-
222 specific TEs was strongly decreased (**Fig. 4b, Supplementary Fig. S4b**). RT-qPCR analysis
223 confirmed that *ZAM* and other soma-specific TEs, such as *Gypsy*, were derepressed (**Fig. 4c**).
224 Moreover, smRNA FISH results revealed that *ZAM* RNAs were not restricted to the posterior
225 follicle cells. Indeed, like in ovaries harboring GLKD of the piRNA pathway in RevI-H2i2
226 background, *ZAM* RNAs were present in oocytes. Actually, we detected *ZAM* RNAs in the
227 ooplasm in 90% of stage 10 follicles (**Fig. 4d, e, Supplementary Fig. S4c**). This result
228 confirmed that when no *ZAM* piRNA is produced in the germline, *ZAM* RNAs expressed in the
229 somatic follicle cells transit to the oocyte.
230
231

232 ***ZAM* RNAs transcribed in follicle cells transit to the oocyte and are deposited in early**
233 **embryos.**
234
235 Although *ZAM* should be transcribed only in follicle cells [28,29], we wanted to rule out the
236 possibility that *ZAM* RNAs originated from germinal nurse cells. *ZAM* RNAs strongly
237 accumulated at the posterior pole of oocytes, suggesting that these RNAs arrive from the
238 adjacent follicle cells (**Fig. 3e, 4d**). Moreover, we detected *ZAM* RNA staining only in the
239 cytoplasm of the oocyte at late stages of oogenesis (stage 10) upon *ago3*-GLKD in RevI-H2i2
240 ovaries. By contrast to TEs expressed in the germline (eg, Burdock), we never observed *ZAM*
241 RNA staining in the cytoplasm of other germ cells, the nurse cells, at any stage. (**Fig. 5a**). Our
242 experiments indicated that *ZAM* RNAs detected in oocytes originate from the somatic follicle
243 cells and not from nurse cells.
244 Throughout the animal kingdom, the first embryonic development stages are controlled by
245 transcripts and proteins deposited by the mother during oogenesis [30]. In Drosophila, most of the
246 maternal mRNAs are dumped from nurse cells into the oocyte during oogenesis. Although *ZAM*
247 transcripts originate from the somatic follicle cells, when we investigated *ZAM* RNA
248 transmission to the oocyte, we found such transcripts also in early embryos. Indeed, smRNA
249 FISH experiments revealed a strong accumulation of *ZAM* RNAs in early embryos (before the
250 zygotic transition, 0-2 hours after egg laying) from a RevI-H2i2 mother with GLKD of the
251 piRNA pathway, and also from a mother without germinal *ZAM* piRNAs with sKD of the
252 piRNA pathway in follicle cells (**Fig. 5b, Supplementary Fig. S5a**). Furthermore, *ZAM* RNAs
253 accumulated at the posterior pole of early embryos where future germ cells cellularize (**Fig. 5c**,
254 **Supplementary Fig. S5b**). We detected *ZAM* RNA in high quantity until stage 5 and few *ZAM*
255 RNA signal persisting until the cellularization of the blastoderm (~stage 8-9 of embryogenesis)
256 (**Fig. 5d**). In the progeny of the RevI-H2i2 line, we did not detect any *ZAM* RNA, although

257 *ZAM* RNAs were produced in follicle cells of RevI-H2i2 ovaries (**Fig 5b**). This confirmed that
258 *ZAM* invasion of the germline is not allowed in this line. These data revealed an efficient post-
259 transcriptional silencing of *ZAM* RNAs arriving from follicle cells into the oocyte by piRNAs
260 produced in the germline and dumped into the oocyte. This mechanism should limit the
261 transposition of the somatic *ZAM* retrotransposon into the germline genome.
262
263 ***De novo ZAM* reactivation leads to subfertility.**
264
265 In addition to leading to the desilencing of many TEs, the loss of the piRNA pathway in the
266 germline or in the somatic follicle cells results in severe developmental defects during
267 oogenesis [31–34]. We sought to exclude the possibility that these defects could underlie the *ZAM*
268 invasion into the oocyte observed when the piRNA pathway is knocked down. To this end,
269 using CRISPR-Cas9, we *de novo* deleted the *ZAM* copy in the *flam* piRNA cluster in a line
270 carrying the X chromosome of the IsoA1 reference line and in which the piRNA pathway is
271 fully functional. As in this line, the *flam ZAM* copy is at the genomic position
272 X:21,778,810…21,783,994, we used two guides to create a deletion spanning over
273 X:21,777,135…21,784,062 (6926 pb). We named the resulting line *flamΔZAM*. Mapping of
274 unique piRNAs to the *flam* locus highlighted the complete loss of piRNA production at the
275 deleted region compared with the control IsoA1 line (**Fig. 6a**). Conversely, the global
276 production of unique piRNAs mapping upstream and downstream of *ZAM* was not affected.
277 We confirmed that the deletion induced a strong decrease of all piRNAs mapping to internal
278 regions of *ZAM*. However, antisense piRNAs targeting the *ZAM* LTR were still produced (**Fig.**
279 **6b**). In line with these results, PCR amplification and DNA sequencing of the CRISPR-Cas9
280 target locus showed that only the internal regions of *ZAM* were deleted from the *flam* cluster,

281 resulting in the retention of a solo-LTR at the initial *ZAM* insertion site (**Supplementary Fig.**
282 **S6**).
283 In addition, only the production of *ZAM*-derived antisense piRNAs was strongly altered,
284 whereas the production of antisense piRNAs mapping to other TEs was not affected (**Fig. 6c**).
285 This shows that *ZAM* deletion from *flam* impairs *ZAM*-derived piRNA production, but does not
286 affect the global piRNA production in ovaries. Thus, in the *flamΔZAM* line, the piRNA pathway
287 was functional and no *ZAM* piRNA targeting the internal *ZAM* sequences was produced. We
288 then analyzed *ZAM* RNA expression by smRNA FISH in *flamΔZAM* ovaries. Surprisingly, we
289 did not detect any *ZAM* expression (**Fig. 6d- panel 1**). We hypothesized that there was no
290 functional *ZAM* copy in the *flamΔZAM* line. To check this, we exchanged chromosome II or
291 III of the *flamΔZAM* line with chromosome II or III of the RevI-H2i2 line that each contains
292 four or three recently transposed potentially functional *ZAM* copies, respectively. smRNA FISH
293 revealed that *ZAM* RNAs were strongly expressed when these RevI-H2i2 *ZAM* copies were
294 added to the genome of the *flamΔZAM* line (**Fig. 6d – panels 2 and 3**). We also observed a
295 strong invasion of oocytes: approximately 70% of stage 10 follicles presented *ZAM* RNAs in
296 the ooplasm. We obtained similar results when we added the two *ZAM* copies located close to
297 the X chromosome telomeres of the RevI-H2i2 line to the *flamΔZAM* genome (**Fig. 6d – panel**
298 **4**). Overall, our findings show that in the absence of *ZAM*-derived piRNAs in the somatic
299 follicle cells and in the germline, transcripts from functional *ZAM* copies massively invade the
300 oocytes. Our results also demonstrated that the presence of a solo-LTR of *ZAM* (454 bp) in the
301 *flam* piRNA cluster of the *flamΔZAM* line was not sufficient to silence *ZAM* expression in
302 follicle cells **(Supplementary Fig. S6).**
303
304 Moreover, the *flamΔZAM* flies that contain active *ZAM* copies displayed atrophied ovaries and
305 reduced fertility compared with the control *flamΔZAM* flies that lack functional *ZAM* copies or

306 the RevI-H2i2 line (**Fig 6. e, f**). Specifically, 90% of *flamΔZAM* females with active *ZAM*
307 copies did not lay eggs. These findings indicate that the reactivation of one single
308 retrotransposon in a patch of follicle cells is sufficient to induce severe fertility defects. These
309 results suggest that when *ZAM* piRNAs are produced in the germline, they contribute to ensure
310 genome integrity and fertility by protecting oocytes against *ZAM* invasion from adjacent cells.
311

312 **No initial small RNA response takes place after *ZAM* retrotransposon reactivation.**

313 The reactivation of an endogenous retrovirus, such as *ZAM,* in a genetic context where the
314 piRNA pathway is functional provides a unique opportunity to directly examine the genomic
315 immune response to ERV reactivation and germline invasion. piRNAs and siRNAs are the two
316 main classes of small RNAs produced to control the expression of transposable elements and
317 counteract viral infection, respectively. Moreover, invading viral RNAs can be processed to
318 small RNAs to initiate a primary response [35,36]. To characterize the small RNA response upon
319 ERV reactivation, we sequenced small RNAs that are complexed with Argonaute proteins
320 (named here "regulatory" piRNAs) from *flamΔZAM* ovaries containing four or three functional
321 *ZAM* copies (introduced with the RevI-H2i2 chromosome II or III, respectively). Regardless of
322 the *flamΔZAM* line analyzed, *ZAM* piRNAs (23 to 29 nt) were not *de novo* produced upon *ZAM*
323 reactivation and despite germline invasion. Indeed, only few sense and antisense *ZAM* piRNAs
324 were produced in the ovaries of the *flamΔZAM* line that contained functional *ZAM* copies (**Fig.**
325 **7a – Supplementary Fig. S7a**). Moreover, this lack of piRNA production concerned only *ZAM*
326 and the level of regulatory piRNAs targeting other TEs was not reduced (**Fig. 7b-**
327 **Supplementary Fig. S7b,**).
328 We sequenced total small RNAs in *vret-* or *yb*-sKD ovaries where *ZAM* is also reactivated.
329 Surprisingly we observed sense *ZAM*-mapping small RNAs of various sizes, including 23-29
330 nt small RNAs that are not produced in a *white*-sKD control condition (**Supplementary Fig.**

331 **S7c**). However, the sense 23-29 nt *ZAM*-derived small RNAs detected in *yb*- and *vret*-sKD
332 ovaries did not display uridine bias at the 5' end, a feature of mature primary piRNAs
333 (**Supplementary Fig. S7d**). Furthermore, these sense *ZAM*-mapping small RNAs were not
334 immunoprecipitated with Piwi proteins. We detected almost no Piwi-bound sense or antisense
335 *ZAM*-mapping piRNAs in *vret*- or *yb*-sKD ovaries (**Supplementary Fig. S7e**). In addition, no
336 sense regulatory *ZAM*-piRNAs complexed with Argonaute proteins were detected (Fig. 4a).
337 These results suggest that the sense *ZAM*-mapping small RNAs are most likely degradation
338 products of excess *ZAM* sense transcripts and not piRNAs. Thus, in these different conditions,
339 we did not detect any primary piRNA response either in the somatic cells where *ZAM* is
340 expressed or in the oocyte that it invades.
341
342 As *ZAM* is an endogenous retrovirus, we also asked whether a siRNA response was triggered
343 upon *ZAM* reactivation. For this, we analyzed the production of 21 nt small RNAs complexed
344 with Argonaute proteins ("regulatory" siRNAs). Similar levels of *ZAM* 21 nt small RNAs were
345 produced in the *flamΔZAM* lines with functional *ZAM* copies (**Fig. 7c**). Globally, the production
346 of 21 nt small RNAs was comparable for all TEs upon *ZAM* reactivation (**Fig. 7d**). Their
347 production was only slightly increased in the *vret*-sKD condition, compared to the *white*-sKD
348 control line without *ZAM* expression (**Supplementary Fig. S7f**). Overall, our results suggest
349 that no or very few siRNAs that could protect ovaries from *ZAM* activity are produced upon
350 activation of this endogenous retrovirus in follicle cells.
351 Therefore, we propose that no innate small RNA response is triggered in the lines where *ZAM*
352 is reactivated, suggesting that *ZAM* will not be brought under control until the establishment of
353 specific adaptive immunity and immune memory.
354
355

356 # Discussion

357

358 The discovery that piRNAs produced in germ cells not only restrict TE expression in the
359 germline itself but also counteract the invasion of endogenous retroviruses from adjacent cells
360 reveals a novel role of piRNAs produced locally as an effective defense mechanism at the tissue
361 scale. When piRNAs against *ZAM* are produced in the germline, *ZAM* transcripts are limited to
362 a patch of follicle cells and flies are fertile. Conversely, in the absence of *ZAM*-derived piRNAs
363 in the germline, transcripts from functional *ZAM* copies massively invade the oocytes and are
364 even transferred to the embryos. At the molecular level, our study revealed that piRNAs
365 produced in nurse cells and dumped into the oocyte can target RNAs produced by follicle cells
366 and delivered to the oocyte. It is tempting to speculate that the *ZAM* RNAs arriving from
367 somatic cells are targeted by complementary antisense piRNAs and degraded through the ping-
368 pong cycle. This would result in the production of *ZAM* sense piRNAs, thus participating in the
369 amplification of the pool of piRNAs against *ZAM* and strengthening the defense mechanism
370 against the invasion.

371 Transposition in the germ cell genome is crucial for TE propagation in a population because it
372 allows the vertical transmission of new insertions. At the time of invasion (e.g. in the *flamΔZAM*
373 line), no *ZAM*-derived piRNA was produced in somatic or germ cells. Therefore, this condition
374 could be compared to what happens when a TE first enters a new species by horizontal transfer.
375 In this case, according to several studies, an initial transposition burst occurs that leads to TE
376 accumulation in the genome before the induction of an adaptive response by the host to control
377 transposition [37–40]. We found at least 17 new *ZAM* insertions in the RevI-H2i2 genome,
378 testifying that *ZAM* actively transposed. Surprisingly, we identified three *ZAM* insertions in the
379 same germline piRNA cluster on the X chromosome. It has been proposed that two different
380 mechanisms can explain the *de novo* piRNA production required to specifically silence a novel

381 invading TE. Indeed, piRNAs can be produced from a new TE insertion into a pre-existing
382 piRNA cluster, and also by euchromatic TE insertions converted into piRNA-producing loci
383 [14,41–43]. Interestingly, when active ZAM copies are added into a *flamΔZAM* genetic background,
384 these copies do not produce regulatory *ZAM*-derived piRNAs and are not converted into piRNA
385 producing regions. Actually, *ZAM* should be transcriptionally inactive in the germline because
386 *ZAM* expression is regulated by the somatic transcription factor Pointed2 [28]. Therefore, although
387 the contribution of individual copies needs to be investigated, our data strongly suggest that
388 germline *ZAM* piRNAs originate at least in part from piRNA clusters. In line with our findings,
389 computer simulations of the TE invasion dynamics suggested that several insertions in piRNA
390 clusters are likely to be required to stop the invasion in the germline [44]. Conversely, we found
391 that a single insertion in a somatic piRNA cluster, such as *flamenco,* is sufficient for TE
392 silencing in the somatic follicle cells, as previously suggested [24]. Indeed, we showed that in
393 follicle cells, the *flamenco* piRNA cluster acts alone as the principal regulator of transposon
394 activity. A precise deletion of the internal part of the endogenous retrovirus *ZAM*, inserted in
395 *flamenco*, leads to a complete derepression of functional *ZAM* copies. Surprisingly, reactivation
396 of this endogenous retrovirus in only a patch of somatic cells leads to subfertility. However, in
397 the germline the deletion of the three most highly expressed germline piRNA clusters (42AB,
398 38C, and 20A) affected neither fertility nor TE silencing [14]. This could be explained by the fact
399 that in the germline, but not in somatic cells, redundancy in piRNA production among different
400 piRNA clusters is observed, with probably multiple piRNA clusters involved in the silencing
401 of the same TE. Moreover, many TE copies have acquired mutations and are non-functional for
402 transcription or for transposition. In the *flamΔZAM* line, although no *ZAM* piRNA was
403 produced, we did not detect any *ZAM* expression. However, when we introduced functional
404 copies of *ZAM*, *ZAM* was expressed demonstrating that depending on the genetic background,
405 functional copies can be absent or present. In general, the TE content varies considerably among

406 populations [42,45]. Interestingly, we identified *ZAM* copies with a deletion that affects the *pol*
407 gene. The copy numbers of this *ZAM*-variant and of classical *ZAM* elements that have
408 transposed in the RevI-H2 line are similar (5 and 8 copies), suggesting that this *ZAM*-variant
409 can transpose very efficiently. Studies in different organisms have shown that TEs rapidly
410 diversify. For instance, the mariner DNA transposon is present in 68 different versions in grass
411 genomes [46]. Such rapid diversification certainly promotes speciation of new families of active
412 TEs.
413 One interesting question is how *ZAM* RNAs transit into the oocyte and at which stage during
414 Drosophila oogenesis *ZAM* transposes in germ cells. We noticed that *ZAM* RNAs, although
415 transcribed in follicle cells, were deposited and accumulated at the posterior pole of early
416 embryos, where future germ cells cellularize. *ZAM* might have evolved to preferentially
417 mobilize in the dividing primordial germ cells of the offspring instead of the developing oocyte,
418 which is not in a phase favorable to transposition because of its prolonged arrest in prophase I
419 associated with highly compacted chromatin [47].
420 Previous experiments demonstrated that *ZAM* RNAs accumulate in follicle cells when
421 vitellogenesis is defective, suggesting that *ZAM* transmission to the oocytes requires functional
422 vitellogenin trafficking [48]. Moreover, endogenous retroviruses, such as *ZAM* and *gypsy*, can
423 form pseudo-viral particles in follicle cells when they are expressed [29,48–50]. *ZAM* Gag end Env
424 proteins are expressed in RevI-H2i2 follicle cells. Retroviral envelope glycoproteins undergo
425 proteolytic processing by cellular serine endoproteases (furin and proprotein convertases) in
426 order to produce the two functional subunits, a glycosylated hydrophilic polypeptide (SU) and
427 a transmembrane domain (TM). This step is necessary to achieve protein competence to
428 promote membrane fusion and virus infectivity [51,52] . The *ZAM env* gene is similar to the
429 retroviral *env* gene, which is responsible for the infectious properties [25]. Indeed, *ZAM* Env is
430 composed of a SU, a TM and an Arg–X–Lys–Arg conserved domain, which is considered to

431 be a furin consensus proteolytic cleavage site [25,53,54]. A recent study demonstrated that
432 intracellular retrotransposition of the *gypsy* endogenous retrovirus can occur in the absence of
433 the Env protein, while intercellular transmission requires a functional Env [50]. Currently, whether
434 *ZAM* can transpose intracellularly, in follicle cells, and the exact mechanism(s) (Env protein-
435 dependent or independent) used for *ZAM* transmission to the oocyte are unknown.
436 Once a virus enters into the target cells, three models for retrovirus uncoating have been
437 proposed: the capsid (1) is disassembled soon after virus entry, (2) remains intact until it comes
438 into contact with the nuclear pore complex, or (3) is partially disassembled during the
439 cytoplasmic traffic to the nucleus (for review [55]). Capsid disassembly of endogenous
440 retroviruses has not been investigated yet. In our study, we observed the loss of *ZAM* RNA in
441 the oocyte when germline *ZAM*-derived piRNAs are produced. We could hypothesize that *ZAM*
442 capsid is dissociated soon after entry in the oocyte and thus piRNAs can easily target *ZAM*
443 RNAs. Alternatively, *ZAM* RNAs are still encapsidated in the oocyte but the capsid does not
444 protect the *ZAM* RNA from piRNAs and their associated endonuclease. Previous studies on
445 virus-like particles of the yeast retrotransposon Ty suggest that these particles form open
446 structures that leave the RNA at the interior of the capsid accessible to RNase [56,57].
447
448 It has been proposed that in different species, an "innate" response at the piRNA or siRNA level
449 could be initiated in the gonads upon retrovirus invasion (reviewed in [58], [36,59–63]). We found that
450 no clear innate small RNA response is triggered upon *ZAM* reactivation in Drosophila ovaries.
451 Total small-RNA sequencing after *ZAM* reactivation revealed the presence of sense *ZAM*-
452 mapping small RNAs, but they lack characteristics of siRNAs or piRNAs. We hypothesize that
453 these sense small RNAs are degradation products generated when there is an excessive amount
454 of *ZAM* RNAs. Altogether, these results show that no initial small RNA response is mounted
455 upon *ZAM* reactivation, leaving follicle cells, where *ZAM* is expressed, and also oocytes

456 unprotected against *ZAM* transposition. Our findings suggest that the piRNA response is a
457 robust response that, in the case of *ZAM*, must be established either in the follicle cells or in the
458 germ cells to efficiently control the invading TE. However, the time required for developing a
459 robust piRNA response to sustainably control TE transposition remains unknown. Lastly, it is
460 thought that the piRNA pathway has no antiviral role in Drosophila [64]. However, it would be
461 interesting to determine whether this is true also if parts of the viral genome are integrated into
462 a germline piRNA cluster during infection.
463

464 # Online Methods

465 **Fly stocks, transgenic lines, and crosses**

466 Flies were maintained at 25°C under a light/dark cycle and 60% humidity. Between 3 and 6
467 days after hatching, flies were used for experiments. The RNAi lines against components of the
468 piRNA pathway used in this study are listed in **Supplementary Table S2**. Isogenic lines from
469 the RevI-H2 stock were generated by crossing a single female with a single male for five
470 generations. Germline and somatic knock-down have been performed using the nanos-gal4 and
471 the traffic-jam-gal4 driver lines respectively. The Drosophila line carrying the *ZAM* sensor is
472 described in [23]. The sensor expresses the GFP reporter gene under the control of an inducible
473 Upstream Activation Sequence promoter (UASp) and harbors a *ZAM* fragment in its 3′UTR
474 (pGFP-*ZAM*).
475

476 **Single-molecule RNA fluorescence *in situ* hybridization (smRNA FISH) in ovaries and**
477 **embryos.**

478 *ZAM* smRNA FISH was performed using 48 probes that target *ZAM* transcripts in a region that
479 is absent in the *ZAM* inserted in the *flamenco* piRNA cluster to detect only transcripts of active
480 *ZAM* copies (sequences in **Supplementary Table S3**). Ovaries from 3 to 6-day-old flies were

481 dissected in Schneider's Drosophila Medium and fixed in Fixing Buffer (4% formaldehyde,
482 0.3% Triton X-100, 1x PBS) for 20 min at room temperature, rinsed three times in 0.3% Triton
483 X-100, once in PBS, and permeabilized in 70% ethanol at 4°C overnight. Permeabilized ovaries
484 were rehydrated in smRNA FISH wash buffer (10% formamide in 2x SSC) for 10 min. Ovaries
485 were resuspended in 50μL hybridization buffer (10% dextran sulfate, 10% formamide in 2x
486 SSC) supplemented with 1μL of smRNA FISH probes. Hybridization was performed with
487 rotation at 37°C overnight. Ovaries were then washed twice with smRNA FISH wash buffer at
488 37°C for 30 min and twice with 2xSSC solution. Then, DNA was stained with DAPI (1/500
489 dilution in 2x SSC) at room temperature for 20 min. Ovaries were mounted in 30 μL
490 Vectashield mounting medium and imaged on a Zeiss LSM-980 or Zeiss LSM-800 confocal
491 microscope. The resulting images were processed using FIJI/ImageJ. The RNA signal
492 specificity was confirmed by adding 1 mg/mL of RNase A in 2x SSC for 2h before the
493 hybridization step.
494 For embryo staining, flies were caged and fed yeast paste. Embryos (0-2h) were collected,
495 dechorionated in 50% bleach for 4 min and rinsed in water. Eggs were fixed in 4%
496 paraformaldehyde/heptane for 20 min, devitellinized by vigorous shaking in 100% methanol
497 and stored in methanol at -20°C. Embryos were rehydrated with 1:1 methanol in 1xPBS, 0.1%
498 Tween-20 for 5 min and twice in 1xPBS, 0.1% Tween-20. Embryos were resuspended in
499 smRNA FISH wash buffer (10% formamide in 2x SSC) for 10 min and then processed for
500 smRNA FISH as described for ovaries. Immunostaining combined with smRNA FISH was
501 performed by adding primary antibodies to the smRNA FISH probes in the hybridization buffer
502 and incubating at 37°C overnight. Embryos were washed twice with smRNA FISH wash buffer
503 and incubated with the secondary antibody (1/1000 diluted in 2x SSC) for 90 min. After two
504 washes in 2x SSC, embryos were mounted in 30 μL Vectashield mounting medium.
505

506 **Immunofluorescence analysis of ovaries and embryos.**

507 Ovaries from 3-5-day-old flies were dissected in supplemented Schneider's medium, ovarioles
508 were separated, and the muscle sheath was removed before fixation to obtain undistorted
509 follicles. Then, ovaries were fixed in 4% formaldehyde/1x PBS/2% Tween-20 for 15 min,
510 rinsed three times with 1x PBS/2% Tween-20, and incubated 2 hours in PBT (1x PBS, 0.1%
511 Triton X-100, 1% BSA) for blocking. Ovaries were incubated with primary antibodies in PBT
512 at 4°C overnight (antibodies are described in **Supplementary Table S4**). After three washes in
513 PBT, ovaries were incubated with the corresponding secondary antibodies (1/1000 diluted in
514 PBT), coupled to Alexa-488 or Cy3 for 90 min. After two washes in 1x PBS, DNA was stained
515 with DAPI (1/500 dilution in 1x PBS) for 20 min. Ovaries were mounted in 30 μL Vectashield
516 mounting medium and imaged on a Zeiss LSM-980 or Zeiss LSM-800 confocal microscope.
517 The resulting images were processed using FIJI/ImageJ.

518

519 **RT-qPCR analysis of transposon expression**

520 10-20 pairs of dissected ovaries were homogenized in TRIzol reagent (Ambion). Following
521 DNase I treatment, cDNA was prepared from 1 μg RNA by random priming of total RNA using
522 Superscript IV Reverse Transcriptase (ThermoFisher Scientific). Quantitative PCR was
523 performed with Roche FastStart SYBR Green Master and the Lightcycler® 480 Instrument.
524 cDNA was used as template for RT-qPCR quantification of transposon mRNA levels (primer
525 sequences in **Supplementary Table S5**). All experiments were done with three biological
526 replicates and with technical triplicates. Steady-state RNA levels were calculated from the
527 threshold cycle for amplification with the $2^{-\Delta\Delta CT}$ method; *rp49* was used for normalization.

528

529

530

531 **Immunoprecipitation of Piwi-ribonucleoprotein complexes**

532 For each genotype, 50 pairs of ovaries from 3-5-day-old flies were dissected, lysed in 1 ml lysis
533 buffer (20 mM HEPES-NaOH at pH 7.0, 150 mM NaCl, 2.5 mM $MgCl_2$, 250 mM sucrose,
534 0.05% NP40, 0.5% Triton X-100, 1x Roche-Complete). Samples were cleared by centrifugation
535 at 10000 × g at 4°C for 10 min. Extracts were incubated with rotation with rabbit polyclonal
536 anti-Piwi antibodies (4 μg per sample) at 4°C for 4h followed by overnight incubation with
537 Dynabeads Protein A (50 μl, Invitrogen, 10002D) at 4°C with rotation. Before incubation,
538 beads were equilibrated in NT2 buffer (50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 1 mM $MgCl_2$,
539 0.05% NP40). Beads were washed twice with ice-cold NT2 and twice with NT2 in which NaCl
540 concentration was adjusted to 300 mM. Nucleic acids that co-immunoprecipitated with Piwi
541 were isolated by treating beads with 0.7 mg/ml proteinase K in 0.3ml proteinase K buffer (0.5%
542 SDS, 10mM Tris-HCl pH 7.4, 50mM NaCl, 5mM EDTA), followed by phenol/chloroform
543 extraction (phenol at neutral pH) and ethanol precipitation.

544

545 **DNA Isolation, Oxford Nanopore Technology sequencing and genome analysis**

546 DNA was extracted from 200 RevI-H2i2 females using the Qiagen Blood & Cell Culture DNA
547 Midi kit. The genomic DNA quality and quantity were evaluated using a Femto Pulse (Agilent)
548 and a Qbit 3.0 (Invitrogen) respectively. Oxford Nanopore Technology (ONT) sequencing was
549 performed by I2BC (Gif-sur-Yvette, France) using five micrograms of genomic DNA. Adapter-
550 trimmed ONT reads were analyzed using NCBI Blastn
551 (https://blast.ncbi.nlm.nih.gov/Blast.cgi) or command-line Blastn (Blastn 2.8.1). The RevI-
552 H2i2 ONT reads containing *ZAM* insertions were detected by Blastn with the reference *ZAM*
553 [65], Repbase *ZAM*_I and *ZAM*_LTR sequences (https://www.girinst.org/repbase/, [26]). The *ZAM*
554 containing reads were then recovered with bedtools getfasta
555 (https://bedtools.readthedocs.io/en/latest/content/tools/getfasta.html, [66]). The *ZAM* insertion

556 sites where determined using Blastn of the *ZAM* containing reads to the *D. melanogaster*
557 Release 6 genome (http://flybase.org). The absence of empty sites, without a given *ZAM*
558 insertion, was verified by Blastn of the respective empty insertion sites, recovered from the *D.*
559 *melanogaster* Release 6 genome, to all ONT reads.
560

561 **Small RNA sequencing**

562 Regulatory small RNA extraction was performed as described in [67]. Briefly, 50 pairs of ovaries
563 from 3-5-day-old flies were lysed and Argonaute-small RNA complexes were isolated using
564 TraPR ion exchange spin columns. Small RNAs were subsequently purified using TRIzol
565 (Ambion). Total RNA also was extracted for some samples (80-100 ovaries for each sample)
566 using the classical TRIzol extraction, followed by 2S depletion and size selection before
567 sequencing. Sequencing was performed by Fasteris SA (Geneva, Switzerland) on an Illumina
568 NextSeq550 instrument (13-15 million reads per sample).
569 Illumina small RNA sequencing reads were loaded in the small RNA pipeline sRNAPipe [68] for
570 mapping to the various genomic sequence categories of the *D. melanogaster* genome (Release
571 6). For the analysis, 23-29nt genome mappers were selected as piRNAs, and 21nt genome
572 mappers were selected as siRNAs. "Unique" piRNAs were defined as 23-29 reads that mapped
573 uniquely across the reference genome. piRNAs or unique piRNAs were mapped to TEs (0-3
574 mismatches) or to piRNA clusters, as defined in [10] (no mismatch allowed). The window size
575 was of 91nt for *ZAM* and 140nt for the *flamenco* region at X:21,769,891…21,789,891 to
576 establish the piRNA density profile. For comparison between samples, all read counts were
577 normalized by the number of reads used as input (genome-mapped piRNAs, unique genome-
578 mapped piRNAs, or genome-mapped siRNAs). The ping-pong signature was assessed by
579 calculating the percentage of sense piRNAs with 10nt 5'-overlap with antisense piRNAs, based
580 on piRNA mapping to *ZAM* (0 to 3 mismatches). The percentages of 1 to 23nt-long overlaps

581 were determined, and the percentage of 10nt overlaps was defined as the ping-pong signature.
582 The Z-score for 10nt overlaps was determined using the percentages of 1 to 23nt-long overlaps
583 as a background and was considered significant when it was >1.96. Scatter plots were done
584 with RStudio with antisense piRNAs or siRNAs mapping to individual internal sequence of
585 TEs (0-3 mismatches for piRNAs and 0-1 mismatches for siRNAs). All small RNA sequencing
586 done for this study is available in **Supplementary Table S6.**
587

588 **Sterility tests**

589 For the sterility tests, RevI-H2i2 homozygous females were compared to w^{1118} females, and
590 homozygous *flamΔZAM* females that carry two functional *ZAM* copies from the RevI-H2i2 X
591 chromosome were compared to *flamΔZAM* females without functional *ZAM* copies in their
592 genome. Thirty 3-5-day-old virgin females of each genotype were individually mated with three
593 w^{1118} males, and eggs were collected at 25°C for 24 hours. The number of eggs laid by each
594 female was determined. Then, eggs were kept at 25°C for another 24 hours before determining
595 the egg hatching rate. The experiment was done twice for a total of approximatively 60
596 replicates for each condition.
597

598 **Generation of the *flamΔZAM* line using the CRISPR-Cas9 technology**

599 The *ZAM* copy in the *flamenco* piRNA cluster was deleted using a pCFD6 plasmid that
600 expresses two single guide RNAs (sgRNAs) under the Gal4/UAS system. The sgRNAs were
601 designed to target genome-unique sequences located upstream and downstream of the *flamenco*
602 *ZAM* copy: TTGTAGCGCTCTTCTTCTCT (sgRNA *flamΔZAM_1*) and
603 AGCGCAACCACGTACAGCGA (sgRNA *flamΔZAM_2*). The pCFD6 plasmid harboring
604 these sgRNAs was injected by the Bestgene company into embryos from the BL#9736 stock
605 (Bloomington stock center) for integration of the plasmid into the genomic site 53B2. The

606 obtained flies were crossed with nanos-gal4; UAS-Cas9 flies (BL#54593, Bloomington stock
607 center). Prior to crossing, the X chromosomes of both Drosophila lines were replaced by the X
608 chromosome of the IsoA1 Drosophila line. After establishment of 100 lines, homozygote
609 females were screened by PCR using primers specific for the flamenco *ZAM* copy. Lines
610 without any amplification of these *flamenco ZAM* sequences were selected for PCR
611 amplification of the target locus with primers that framed the *flamenco ZAM* copy and the target
612 sequences of the sgRNAs. The obtained amplicons were sequenced.
613

614 ACKNOWLEDGEMENTS

615 We thank the *Bloomington Stock Center* for providing stocks. We thank Julius Brennecke for
616 providing antibodies. We thank flybase.org for providing databases. We thank Séverine
617 Chambeyron and Kirsten-André Senti for helpful discussions and the members of the Brasset
618 laboratory for comments on the manuscript. We thank Yoan Renaud and Pierre Pouchin for
619 helpful advices and development of the GReD bioinfomatic platform. We thank Gentyane
620 Facility for the Quality Control of gDNA.
621

622 FUNDING

623

624 This work was supported by grants from the Agence Nationale pour la Recherche (ANR-
625 CHApiTRE (AAPG2020 – CE12), ANR-BiopiC (AAPG2021 – PRC – CE12), and ANR-
626 EpiTET projects (17-CE12-0030-03)), the Association de Recherche Cancer
627 (PJA20171206129). M.Y. was supported by the Ministère de l'Enseignement Supérieur et de
628 la Recherche (MESR) and the Fondation pour la Recherche Medical (FRM)
629 (FDT202106012950). This research is supported by the French government Emergence-IDEX-
630 ISITE initiative 16-IDEX-0001 (CAP20-25). The funders had no role in study design, data
631 collection and analysis, decision to publish, or preparation of the manuscript.
632

633 COMPETING INTERESTS

634 The authors declare that no competing interest exists.
635
636

637 **REFERENCES**

638 1. Byrd, D. T. & Kimble, J. Scratching the niche that controls Caenorhabditis elegans
639 germline stem cells. *Semin. Cell Dev. Biol.* **20**, 1107–1113 (2009).
640 2. Kidder, G. M. & Vanderhyden, B. C. Bidirectional communication between oocytes
641 and follicle cells: ensuring oocyte developmental competence. *Can. J. Physiol.*
642 *Pharmacol.* **88**, 399–413 (2010).
643 3. Saunders, P. T. K. Germ cell-somatic cell interactions during spermatogenesis. *Reprod.*
644 *(Cambridge, England). Suppl.* **61**, 91–101 (2003).
645 4. Jemc, J. C. Somatic gonadal cells: the supporting cast for the germline. *Genesis* **49**,
646 753–775 (2011).
647 5. Hall, D. H. *et al.* Ultrastructural features of the adult hermaphrodite gonad of
648 Caenorhabditis elegans: relations between the germ line and soma. *Dev. Biol.* **212**,
649 101–123 (1999).
650 6. McClintock, B. The origin and behavior of mutable loci in maize. *Proc. Natl. Acad.*
651 *Sci. U. S. A.* **36**, 344–355 (1950).
652 7. Aravin, A. *et al.* A novel class of small RNAs bind to MILI protein in mouse testes.
653 **442**, 203–207 (2006).
654 8. Grivna, S. T., Beyret, E., Wang, Z. & Lin, H. A novel class of small RNAs in mouse
655 spermatogenic cells. *Genes Dev.* **20**, 1709–1714 (2006).
656 9. Lau, N. C. *et al.* Characterization of the piRNA complex from rat testes. *Science (80-.*
657 *).* **313**, 363–367 (2006).
658 10. Brennecke, J. *et al.* Discrete Small RNA-Generating Loci as Master Regulators of
659 Transposon Activity in Drosophila. *Cell* **128**, 1089–1103 (2007).
660 11. Aravin, A. a, Hannon, G. J. & Brennecke, J. The Piwi-piRNA Pathway Provides an
661 Adaptive Defense in the Transposon Arms Race. *Science (80-.).* **318**, 761–764 (2007).

662 12. George, P. *et al*. Increased production of piRNAs from euchromatic clusters and genes
663 in Anopheles gambiae compared with Drosophila melanogaster. *Epigenetics and*
664 *Chromatin* **8**, (2015).
665 13. Malone, C. D. *et al*. Specialized piRNA Pathways Act in Germline and Somatic
666 Tissues of the Drosophila Ovary. *Cell* **137**, 522–535 (2009).
667 14. Gebert, D. *et al*. Large Drosophila germline piRNA clusters are evolutionarily labile
668 and dispensable for transposon regulation. *Mol. Cell* 1–14 (2021)
669 doi:10.1016/j.molcel.2021.07.011.
670 15. Handler, D. *et al*. The genetic makeup of the drosophila piRNA pathway. *Mol. Cell* **50**,
671 762–777 (2013).
672 16. Olivieri, D., Sykora, M. M., Sachidanandam, R., Mechtler, K. & Brennecke, J. An in
673 vivo RNAi assay identifies major genetic and cellular requirements for primary piRNA
674 biogenesis in Drosophila. *EMBO J*. **29**, 3301–3317 (2010).
675 17. Desset, S., Meignin, C., Dastugue, B. & Vaury, C. COM, a heterochromatic locus
676 governing the control of independent endogenous retroviruses from Drosophila
677 melanogaster. *Genetics* **164**, 501–509 (2003).
678 18. Pélisson, A. *et al*. Gypsy transposition correlates with the production of a retroviral
679 envelope-like protein under the tissue-specific control of the Drosophila flamenco
680 gene. *EMBO J*. **13**, 4401–4411 (1994).
681 19. Bucheton, A. The relationship between the flamenco gene and gypsy in Drosophila :
682 how to tame a retrovirus. **11**, 349–353 (1995).
683 20. Robert, V., Prud'homme, N., Kim, A., Bucheton, A. & Pélisson, A. Characterization of
684 the flamenco region of the Drosophila melanogaster genome. *Genetics* **158**, 701–713
685 (2001).
686 21. Mével-Ninio, M., Pelisson, A., Kinder, J., Campos, A. R. & Bucheton, A. The

687 flamenco Locus Controls the gypsy and ZAM Retroviruses and Is Required for
688 Drosophila Oogenesis ´. **1624**, 1615–1624 (2007).
689 22. Lécher, P., Bucheton, A. & Pélisson, A. Expression of the Drosophila retrovirus gypsy
690 as ultrastructurally detectable particles in the ovaries of flies carrying a permissive
691 flamenco allele. *J. Gen. Virol.* **78**, 2379–2388 (1997).
692 23. Duc, C. *et al.* Trapping a somatic endogenous retrovirus into a germline piRNA cluster
693 immunizes the germline against further invasion. *Genome Biol.* **20**, 1–14 (2019).
694 24. Zanni, V. *et al.* Distribution, evolution, and diversity of retrotransposons at the
695 flamenco locus reflect the regulatory properties of piRNA clusters. *Proc. Natl. Acad.*
696 *Sci.* **110**, 19842–19847 (2013).
697 25. Leblanc, P., Desset, S., Dastugue, B. & Vaury, C. Invertebrate retroviruses: ZAM a
698 new candidate in D. melanogaster. *EMBO J.* **16**, 7521–7531 (1997).
699 26. Jurka, J. *et al.* Repbase Update, a database of eukaryotic repetitive elements.
700 *Cytogenet. Genome Res.* **110**, 462–467 (2005).
701 27. Desset, S. *et al.* Mobilization of two retroelements, ZAM and Idefix, in a novel
702 unstable line of Drosophila melanogaster. *Mol. Biol. Evol.* **16**, 54–66 (1999).
703 28. Meignin, C., Dastugue, B. & Vaury, C. Intercellular communication between germ line
704 and somatic line is utilized to control the transcription of ZAM, an endogenous
705 retrovirus from Drosophila melanogaster. *Nucleic Acids Res.* **32**, 3799–3806 (2004).
706 29. Leblanc, P. *et al.* Life Cycle of an Endogenous Retrovirus, ZAM, in Drosophila
707 melanogaster. *J. Virol.* **74**, 10658–10669 (2000).
708 30. Atallah, J. & Lott, S. E. Evolution of maternal and zygotic mRNA complements in the
709 early Drosophila embryo. *PLoS Genet.* **14**, e1007838 (2018).
710 31. Aravin, A. A. & Hannon, G. J. Small RNA silencing pathways in germ and stem cells.
711 *Cold Spring Harb. Symp. Quant. Biol.* **73**, 283–290 (2008).

712 32. Khurana, J. S. & Theurkauf, W. piRNAs, transposon silencing, and Drosophila
713 germline development. *J. Cell Biol.* **191**, 905–913 (2010).
714 33. Saito, K. & Siomi, M. C. Small RNA-mediated quiescence of transposable elements in
715 animals. *Dev. Cell* **19**, 687–697 (2010).
716 34. Klattenhoff, C. *et al.* Drosophila rasiRNA pathway mutations disrupt embryonic axis
717 specification through activation of an ATR/Chk2 DNA damage response. *Dev. Cell*
718 **12**, 45–55 (2007).
719 35. Parameswaran, P. *et al.* Six RNA viruses and forty-one hosts: viral small RNAs and
720 modulation of small RNA repertoires in vertebrate and invertebrate systems. *PLoS*
721 *Pathog.* **6**, e1000764 (2010).
722 36. Yu, T. *et al.* The PiRNA Response to Retroviral Invasion of the Koala Genome. *SSRN*
723 *Electron. J.* 1–12 (2019) doi:10.2139/ssrn.3391942.
724 37. Le Rouzic, A. & Capy, P. The first steps of transposable elements invasion: Parasitic
725 strategy vs. genetic drift. *Genetics* **169**, 1033–1043 (2005).
726 38. Kofler, R., Senti, K. A., Nolte, V., Tobler, R. & Schlötterer, C. Molecular dissection of
727 a natural transposable element invasion. *Genome Res.* **28**, 824–835 (2018).
728 39. Robillard, É., Le Rouzic, A., Zhang, Z., Capy, P. & Hua-Van, A. Experimental
729 evolution reveals hyperparasitic interactions among transposable elements. *Proc. Natl.*
730 *Acad. Sci. U. S. A.* **113**, 14763–14768 (2016).
731 40. Schaack, S., Gilbert, C. & Feschotte, C. Promiscuous DNA: Horizontal transfer of
732 transposable elements and why it matters for eukaryotic evolution. *Trends Ecol. Evol.*
733 **25**, 537–546 (2010).
734 41. Mohn, F., Sienski, G., Handler, D. & Brennecke, J. The Rhino-Deadlock-Cutoff
735 complex licenses noncanonical transcription of dual-strand piRNA clusters in
736 Drosophila. *Cell* **157**, 1364–1379 (2014).

737 42. Luo, S. *et al*. The evolutionary arms race between transposable elements and piRNAs
738 in Drosophila melanogaster. *BMC Evol. Biol*. **20**, 1–18 (2020).
739 43. Shpiz, S., Ryazansky, S., Olovnikov, I., Abramov, Y. & Kalmykova, A. Euchromatic
740 Transposon Insertions Trigger Production of Novel Pi- and Endo-siRNAs at the Target
741 Sites in the Drosophila Germline. *PLoS Genet*. **10**, (2014).
742 44. Kofler, R. Dynamics of transposable element invasions with piRNA clusters. *Mol. Biol.*
743 *Evol*. **36**, 1457–1472 (2019).
744 45. Lerat, E. *et al*. Population-specific dynamics and selection patterns of transposable
745 element insertions in European natural populations. *Mol. Ecol*. **28**, 1506–1522 (2019).
746 46. Feschotte, C. & Wessler, S. R. Mariner-like transposases are widespread and diverse in
747 flowering plants. *Proc. Natl. Acad. Sci. U. S. A*. **99**, 280–285 (2002).
748 47. Navarro-Costa, P. *et al*. Early programming of the oocyte epigenome temporally
749 controls late prophase I transcription and chromatin remodelling. *Nat. Commun*. **7**,
750 12331 (2016).
751 48. Brasset, E. *et al*. Viral particles of the endogenous retrovirus ZAM from Drosophila
752 melanogaster use a pre-existing endosome/exosome pathway for transfer to the oocyte.
753 *Retrovirology* **3**, 1–9 (2006).
754 49. Chalvet, F. *et al*. Proviral amplification of the Gypsy endogenous retrovirus of
755 Drosophila melanogaster involves env-independent invasion of the female germline.
756 *EMBO J*. **18**, 2659–2669 (1999).
757 50. Keegan, R. M., Talbot, L. R., Chang, Y. H., Metzger, M. J. & Dubnau, J. Intercellular
758 viral spread and intracellular transposition of Drosophila gypsy. *PLoS Genet*. **17**, 1–22
759 (2021).
760 51. Seidah, N. G. & Prat, A. Precursor convertases in the secretory pathway, cytosol and
761 extracellular milieu. *Essays Biochem*. **38**, 79–94 (2002).

762 52. Bergeron, F., Leduc, R. & Day, R. Subtilase-like pro-protein convertases: from
763 molecular specificity to therapeutic applications. *J. Mol. Endocrinol.* **24**, 1–22 (2000).
764 53. Klenk, H. D. & Garten, W. Host cell proteases controlling virus pathogenicity. *Trends*
765 *Microbiol.* **2**, 39–43 (1994).
766 54. Hosaka, M. *et al.* Arg-X-Lys/Arg-Arg motif as a signal for precursor cleavage
767 catalyzed by furin within the constitutive secretory pathway. *J. Biol. Chem.* **266**,
768 12127–12130 (1991).
769 55. Yamashita, M. & Engelman, A. N. Capsid-Dependent Host Factors in HIV-1 Infection.
770 *Trends Microbiol.* **25**, 741–755 (2017).
771 56. Palmer, K. J. *et al.* Cryo-electron microscopy structure of yeast Ty retrotransposon
772 virus-like particles. *J. Virol.* **71**, 6863–6868 (1997).
773 57. Burns, N. R. *et al.* Symmetry, flexibility and permeability in the structure of yeast
774 retrotransposon virus-like particles. *EMBO J.* **11**, 1155–1164 (1992).
775 58. Kolliopoulou, A. *et al.* PIWI pathway against viruses in insects. *Wiley Interdiscip. Rev.*
776 *RNA* **10**, e1555 (2019).
777 59. Rozhkov, N. V. *et al.* Small RNA-based silencing strategies for transposons in the
778 process of invading Drosophila species. *Rna* **16**, 1634–1645 (2010).
779 60. Rozhkov, N. V. *et al.* Evolution and dynamics of small RNA response to a
780 retroelement invasion in drosophila. *Mol. Biol. Evol.* **30**, 397–408 (2013).
781 61. Kawamura, Y. *et al.* Drosophila endogenous small RNAs bind to Argonaute 2 in
782 somatic cells. *Nature* **453**, 793–797 (2008).
783 62. Rehwinkel, J. *et al.* Genome-Wide Analysis of mRNAs Regulated by Drosha and
784 Argonaute Proteins in Drosophila melanogaster . *Mol. Cell. Biol.* **26**, 2965–2975
785 (2006).
786 63. Katsuma, S. *et al.* Transcriptome profiling reveals infection strategy of an insect

787 maculavirus. *DNA Res. an Int. J. rapid Publ. reports genes genomes* **25**, 277–286
788 (2018).
789 64. Petit, M. *et al*. piRNA pathway is not required for antiviral defense in Drosophila
790 melanogaster . *Proc. Natl. Acad. Sci.* **113**, E4218–E4227 (2016).
791 65. Leblanc, P., Desset, S., Dastugue, B. & Vaury, C. Invertebrate retroviruses: ZAM a
792 new candidate in D. melanogaster. *EMBO J.* **16**, 7521–7531 (1997).
793 66. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing
794 genomic features. *Bioinformatics* **26**, 841–842 (2010).
795 67. Grentzinger, T. *et al*. A universal method for the rapid isolation of all known classes of
796 functional silencing small RNAs. *Nucleic Acids Res.* **48**, (2020).
797 68. Pogorelcnik, R., Vaury, C., Pouchin, P., Jensen, S. & Brasset, E. SRNAPipe: A
798 Galaxy-based pipeline for bioinformatic in-depth exploration of small RNAseq data.
799 *Mob. DNA* **9**, 4–9 (2018).
800
801
802
803
804
805
806
807
808
809
810
811

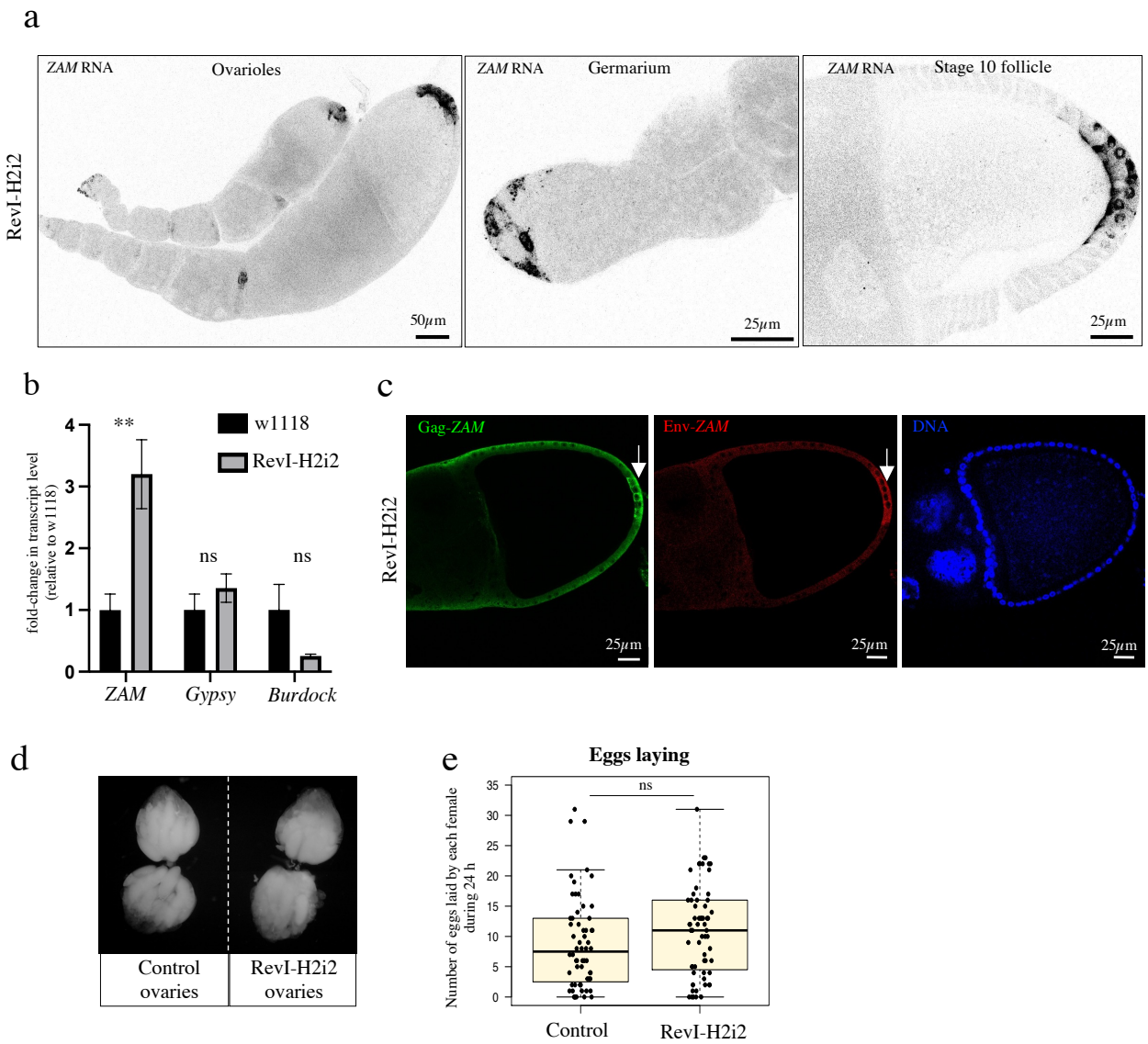


Figure 1: *ZAM* is expressed in follicle cells in the RevI-H2i2 isogenic line.
(a) Color-inverted projection of confocal images of ovarioles, germarium, and stage 10 follicles showing *ZAM* expression by smRNA FISH (in black) in ovaries from the RevI-H2i2 line (Z-projection of 20 stacks). (b) Fold-change in the steady-state *ZAM*, *Gypsy* and *Burdock* RNA levels for RevI-H2i2 ovaries compared with w^{1118} ovaries (control), quantified by RT-qPCR (primer sequences in Supplementary Table S5). At least three biological replicates and three technical replicates were used. ***p-value* <0.01; ns, not-significant (*p-value* >0.05) (Mann-Whitney test). Error bars indicate the standard deviation (SD). (c) Confocal sections of stage 10 RevI-H2i2 egg chambers showing *ZAM*-encoded Gag (green) and Env (red) proteins. DNA was stained with DAPI (blue). (d) Morphology of control and RevI-H2i2 adult ovaries. (e) Box plot displaying the number of eggs laid per fly per day by control and RevI-H2i2 females. Each dot represents an individual female. The midline indicates the median value, and the box shows the first and third quartile. Error bars indicate the SD. ns, not-significant (*p-value* >0.05) (Mann-Whitney test).

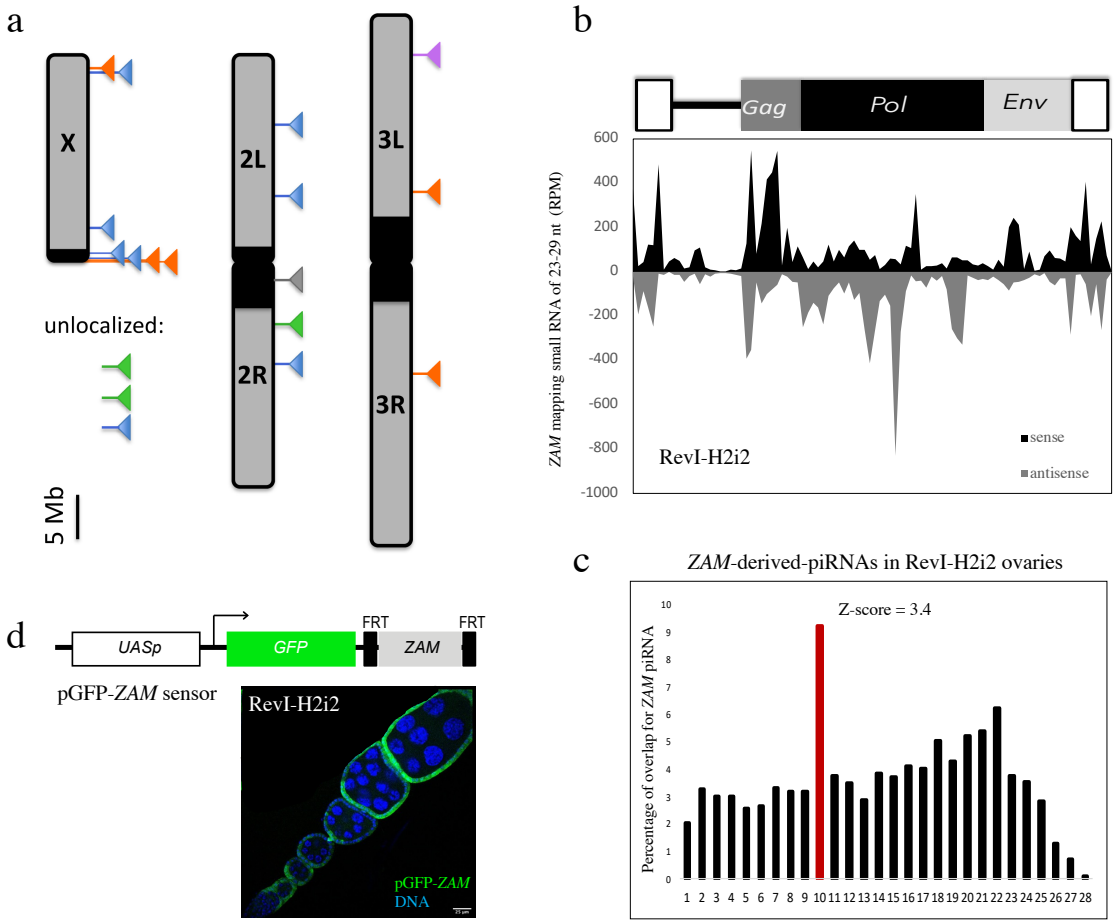


Figure 2: *ZAM* elements in the RevI-H2i2 genome.
(**a**) Each triangle represents a *ZAM* insertion. Blue: full-length *ZAM* elements; orange: *ZAM*-variants, in which positions 5500-6126 of the *ZAM* internal sequences (*ZAM*_I in RepBase) are deleted; violet: *ZAM* full-length with deleted 5'-UTR; green: *ZAM* with internal deletions other than the *ZAM*-variant; gray: *ZAM* copy also present in the reference Release 6 genome (http://flybase.org). Four of the X chromosome *ZAM* copies are in X pericentromeric heterochromatin (in black), all the others are in euchromatic regions. No *ZAM* insertion was found in chromosome 4 (1.35 Mb, not shown). For coordinates and details see Supplementary Table S1. **(b)** Density plot of *ZAM*-mapping regulatory piRNAs along the *ZAM* sequence in RevI-H2i2 ovaries (up to 3 mismatches). **(c)** Histogram showing the percentage of 5'-overlaps between sense and antisense *ZAM*-derived piRNAs (23–29 nt) in RevI-H2i2 ovaries. The proportion of 10 nt 5'-overlap pairs is in red and the corresponding Z-score is indicated. **(d)** Confocal images of ovarioles with GFP (green) and DNA (blue) staining. Ovarioles of the progeny of a cross between RevI-H2i2 females and males carrying the pGFP-*ZAM* sensor transgene driven by actin-Gal4.

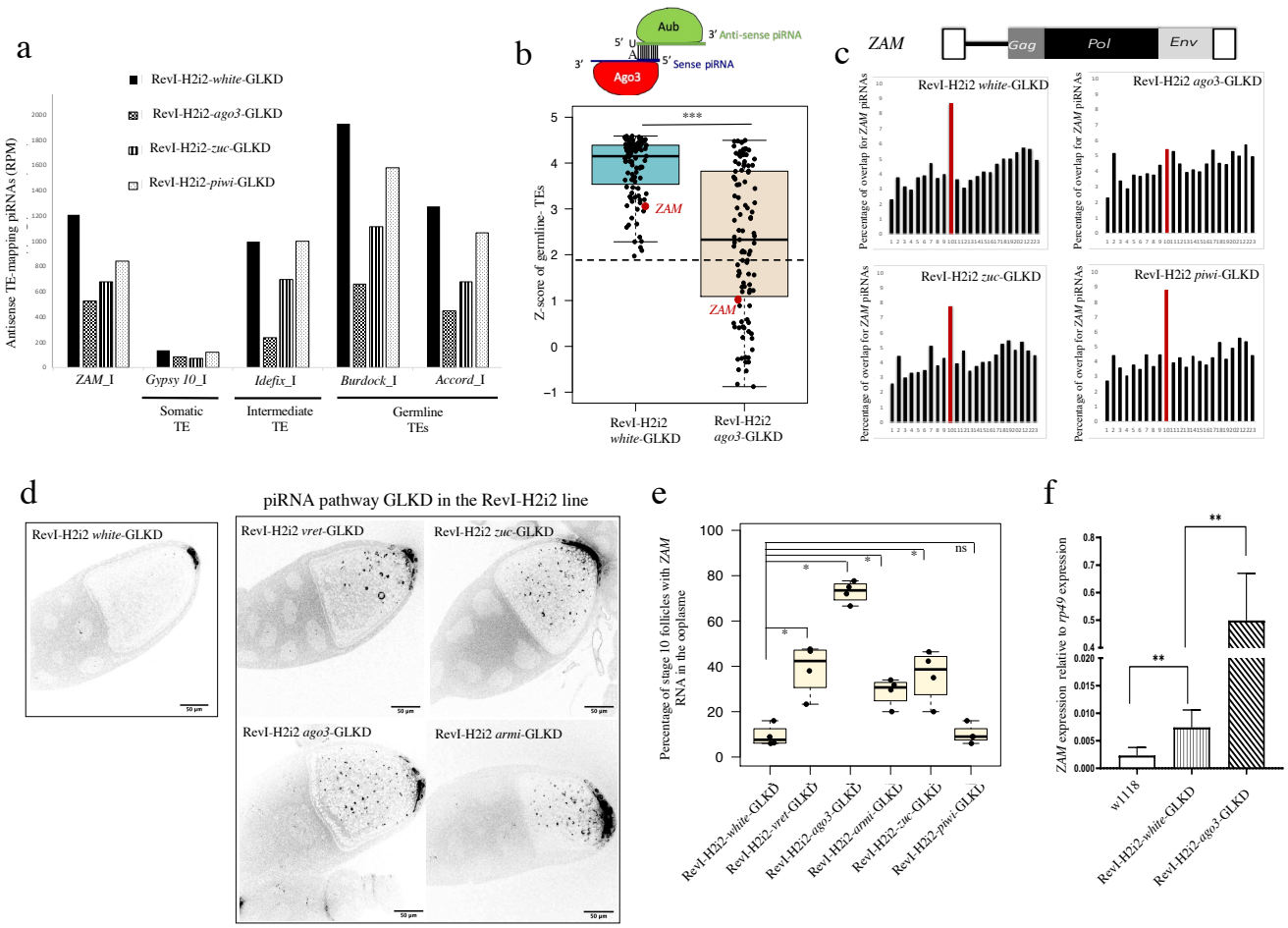


Figure 3: Germline *ZAM* piRNAs produced in RevI-H2i2 germ cells counteract *ZAM* invasion.

(a) Antisense piRNAs mapping to TE internal sequences (0-3 mismatches) in the RevI-H2i2 line upon *white*- (control), *ago3*-, *zuc*-, or *piwi*-GLKD. Normalized per million of genome-mapping piRNAs. **(b)** Box plot displaying the Z-score distribution of 10 nt 5'-overlaps for TE-mapping piRNAs in the *white*-GLKD and *ago3*-GLKD RevI-H2i2 lines. Each dot represents a TE with a Z-score >1.96 (dotted line) in the control condition (*white*-GLKD). *ZAM* is highlighted in red. The midline indicates the median value, and the box shows the first and third quartile. Error bars indicate the SD. ****p-value* <0.001 (Mann-Whitney test). **(c)** Histogram showing the percentage of 5'-overlaps between sense and antisense *ZAM*-mapping piRNAs (up to 3 mismatches) in *white*-, *ago3*-, *zuc*- or *piwi*-GLKD RevI-H2i2 ovaries. The percentage of 10 nt overlapping pairs is in red. **(d)** Color-inverted confocal images of stage 10 egg chambers showing *ZAM* smRNA FISH signal in ovaries of the indicated genotypes. **(e)** Bar plot showing the percentage of stage 10 follicles with *ZAM* RNA detected in the ooplasm by smRNA FISH of RevI-H2i2 ovaries with the indicated GLKD. This experiment was done three times and approximately 50 follicles per conditions were observed for each experiment. Error bars indicate the SD. **(f)** Bar plot showing the steady-state *ZAM* RNA levels in total RNA from ovaries with the indicated genotypes (normalized to *rp49*). At least three biological replicates and two technical replicates were used. ***p-value* <0.01 (Mann-Whitney tst); error bars indicate the SD.

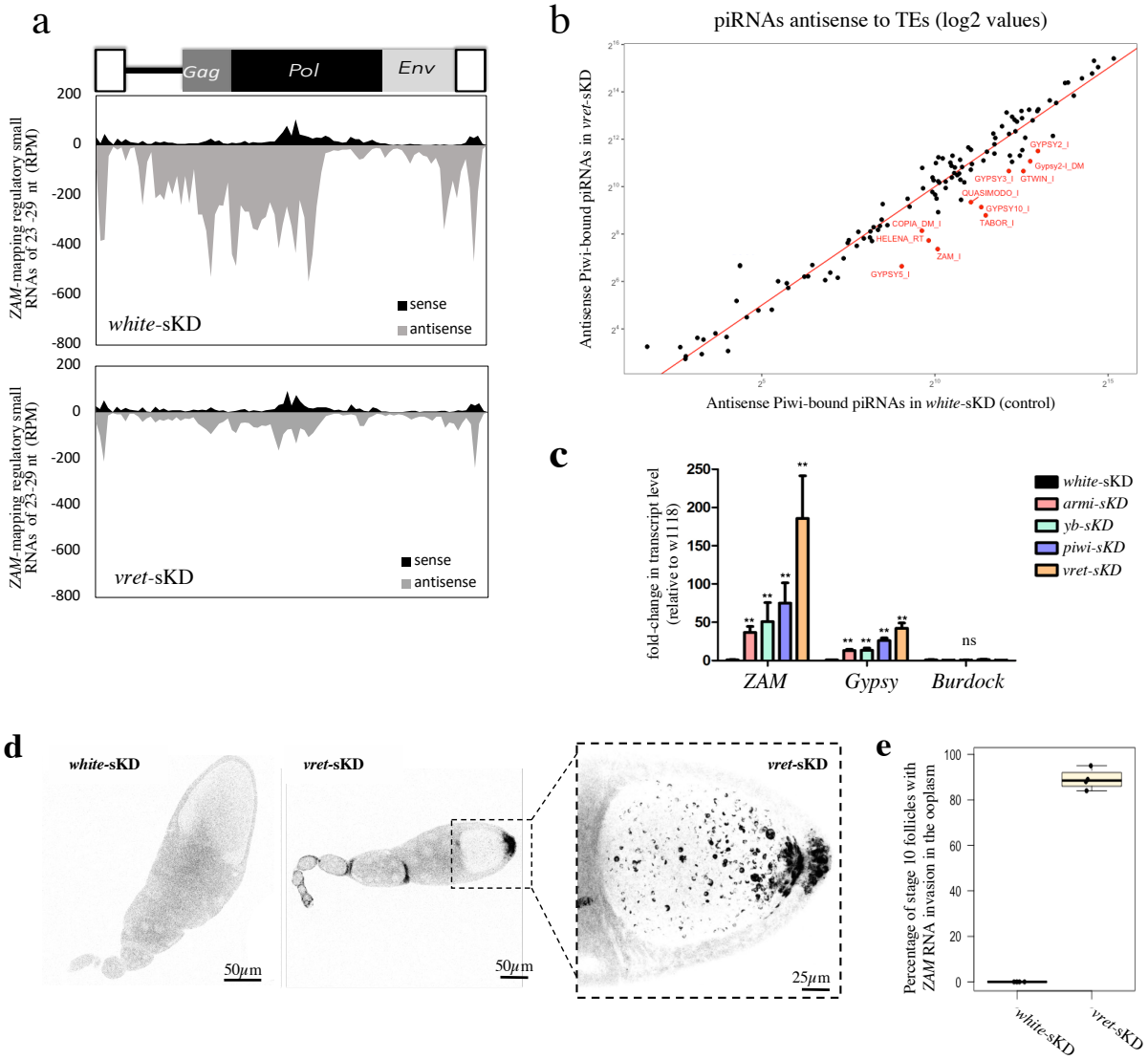


Figure 4: *ZAM* invades the oocyte when no *ZAM*-derived piRNA is produced in the germ cells.

(a) Density plot along the *ZAM* sequence of *ZAM*-mapping regulatory piRNAs produced in *white-* and *vret*-sKD ovaries (up to 3 mismatches). **(b)** Scatter plot showing the log2 of the normalized counts [log2(RPM)] of antisense Piwi-bound piRNAs mapping to individual internal TEs sequences in control ovaries (*white*-sKD) and *vret*-sKD ovaries. Antisense piRNA counts, mapped allowing up to 3 mismatches, were normalized per million of genome-mapping piRNAs (RPM). *TEs in red have a *vret*-sKD/*white*-sKD ratio <0.3. **(c)** Fold-change in the steady-state *ZAM, Gypsy* and *Burdock* RNA level for the indicated sKD ovaries, compared with *white*-sKD ovaries, measured by RT-qPCR (primer sequences in Supplementary Table S5). At least three biological replicates and two technical replicates were used. ***p-value* <0.01 (Mann-Whitney test); error bars indicate the SD. **(d)** Color-inverted confocal images of ovarioles or stage 10 egg chambers showing *ZAM* smRNA FISH signal in *white-* and v*ret*-sKD ovaries. *ZAM* RNA is observed in follicle cells and invaded the ooplasm in *vret*-sKD ovaries (zoomed inset). **(e)** Bar plot showing the percentage of stage 10 follicles with *ZAM* RNA invasion of the ooplasm, assessed by *ZAM* smRNA FISH, of *white-* (control) and *vret*-sKD ovaries. This experiment was done twice and ~50 follicles per condition were observed for each experiment. Error bars indicate the SD

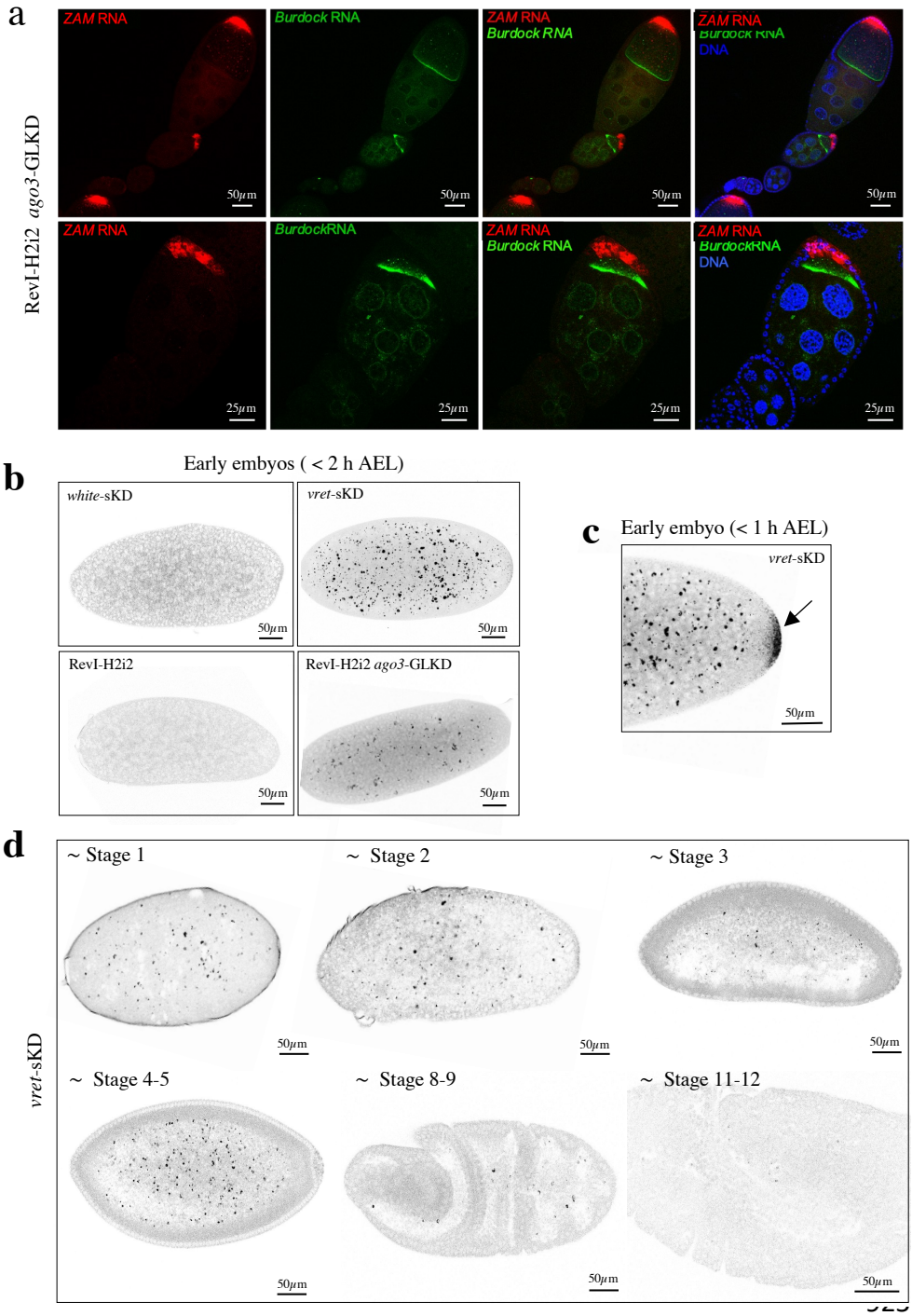


Figure 5: *ZAM* RNA transcribed in the follicle cells transits to the oocyte and is deposited in early embryos.
(a) Confocal images of ovarioles (upper panels) and of an isolate stage 8 egg chamber (lower panels) in the *ago3*-GLKD RevI-H2i2 line. *Burdock* (green) and *ZAM* mRNA (red) were detected by smRNA FISH. DNA was stained with DAPI (blue). **(b)** Color-inverted confocal images showing *ZAM* smRNA FISH signal in early embryos collected 0-2 hours after egg laying (AEL). Mothers were from the indicated genotype. **(c)** Zoom on the posterior part of a *vret*-sKD early embryo collected at 0-1 hours AEL showing *ZAM* RNA accumulation at the posterior pole. **(d)** Color-inverted confocal images showing *ZAM* smRNA FISH signal in embryos laid by *vret*-sKD females at the indicated developmental stages.

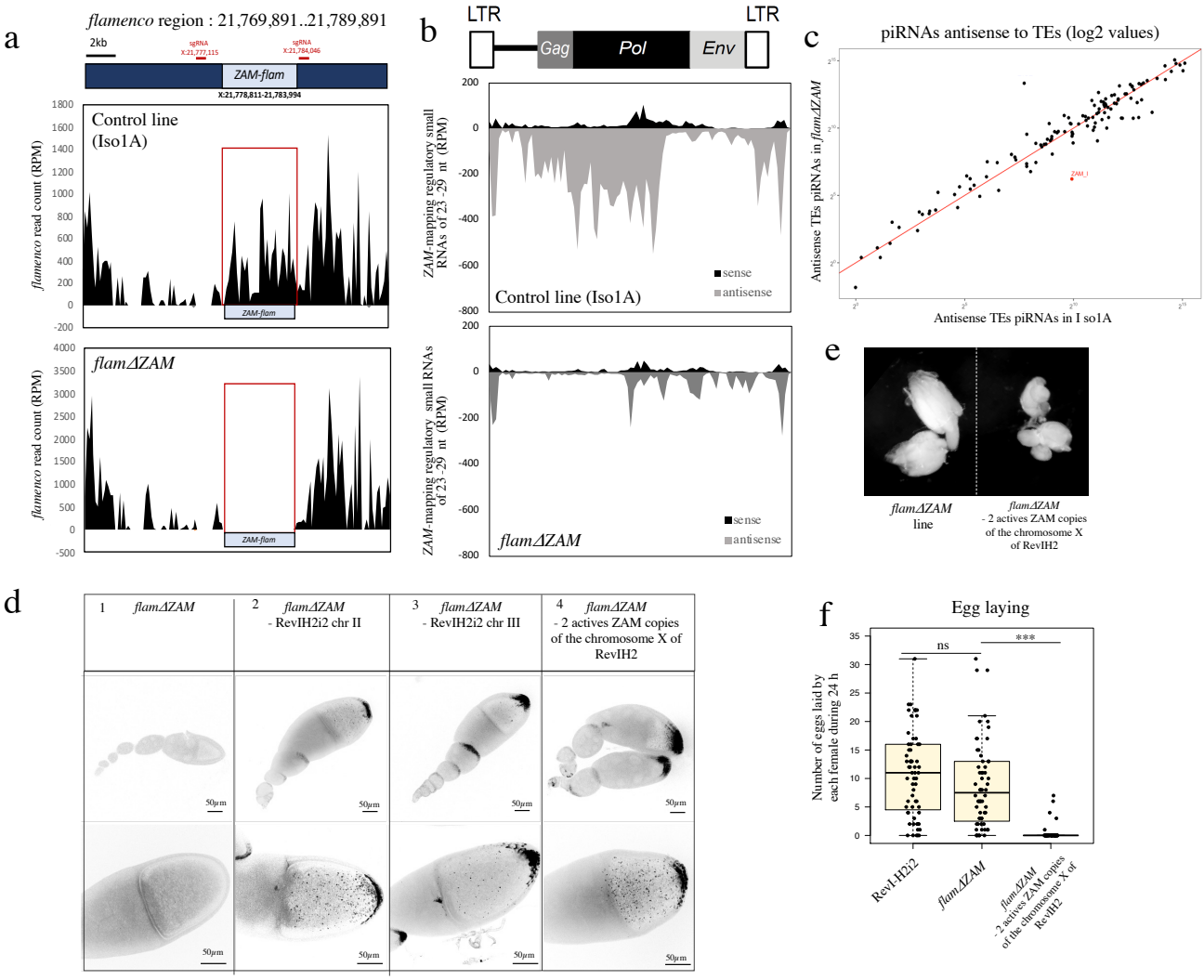
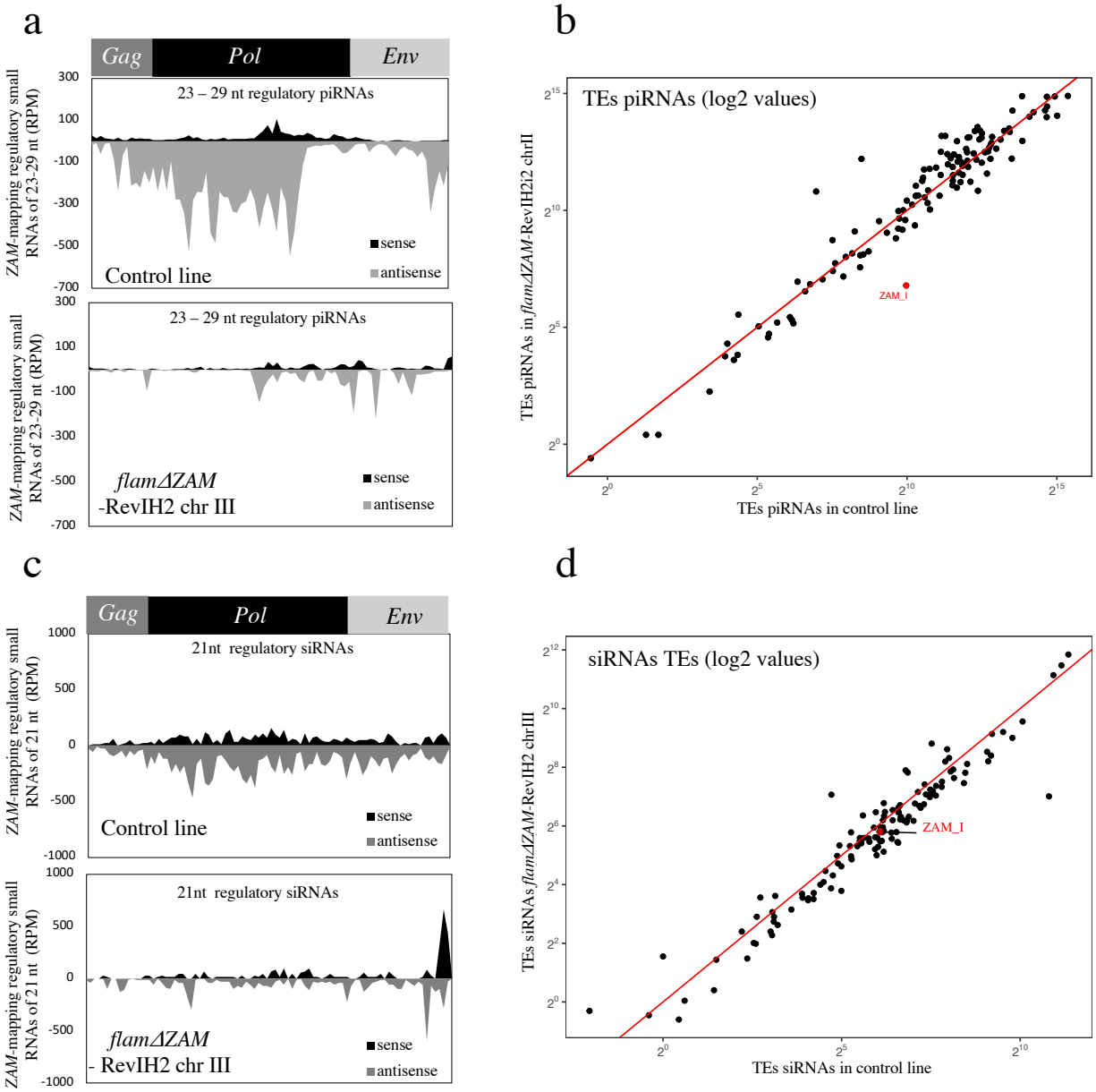


Figure 6: *ZAM* deletion from the *flamenco* piRNA cluster leads to *ZAM* reactivation and subfertility.

(a) Density plots showing genome-unique piRNAs mapping over 10 kb of the *flamenco* piRNA cluster (without mismatch) where the *ZAM* insertion is located (Release 6: X:21,769,891..21,789,891) in the control (Iso1A) and *flamΔZAM* lines. The position of the *flamenco ZAM* copy *(ZAM-flam)* is indicated. The positions of the sgRNAs used for *ZAM-flam* deletion by CRISPR-Cas9 are shown in red. **(b)** Density plot of *ZAM*-mapping regulatory piRNAs along the *ZAM* sequence in control and *flamΔZAM* ovaries (up to 3 mismatches). **(c)** Scatter plot showing the log2 of normalized counts [log2(RPM)] of antisense regulatory piRNA mapping to individual internal TE sequences in control ovaries (Iso1A) versus *flamΔZAM* ovaries. Antisense piRNA counts, mapped allowing up to 3 mismatches, were normalized per million of genome-mapping piRNAs (RPM). *TEs in red have a *flamΔZAM/Iso1A* ratio <0.3. **(d)** Color-inverted confocal images of ovarioles (upper panels) and stage 10 egg chambers (lower panels) from the indicated genotypes showing *ZAM* smRNA FISH signal. **(e)** Ovaries of the *flamΔZAM* line and of the *flamΔZAM* line with two active *ZAM* copies introduced by genetic crossing with the RevI-H2i2 line (the two euchromatic copies of the X chromosome). **(f)** Box plot displaying the number of eggs laid per fly per day by RevI-H2i2, *flamΔZAM and flamΔZAM* with 2 active *ZAM* copies females. Each dot represents an individual female. The midline indicates the median value, and the box shows the first and third quartile. Error bars indicate the SD. ****p-value* <0.001; ns, not-significant (*p-value* >0.05) (Mann-Whitney test).

957



958
959 **Figure 7: No initial piRNA and siRNA response after *ZAM* retrotransposon reactivation**
960 **(a)** Density plot of regulatory piRNAs along the *ZAM* internal sequence in ovaries from a
961 control line (*white*-sKD) and from the *flamΔZAM*-RevI-H2i2 chr III line (up to 3 mismatches).
962 **(b)** Scatter plots showing the log2 of normalized counts [log2(RPM)] of regulatory piRNAs
963 mapping to individual internal TE sequences in control ovaries (*white-sKD*) versus *flamΔZAM*-
964 RevIH2 chr III ovaries. piRNA counts, mapped allowing up to 3 mismatches, were normalized
965 per million of genome-mapping piRNAs (RPM). (**c**) Density plot of *ZAM*-mapping regulatory
966 siRNAs (21nt small-RNAs complexed with Argonaute proteins) along the *ZAM* internal
967 sequence produced in control line (*white*-sKD) and *flamΔZAM*-RevI-H2i2 chr III ovaries (0-1
968 mismatch). (**d**) Scatter plot showing the log2 of normalized counts [log2(RPM)] of antisense
969 regulatory siRNAs mapping to individual internal TE sequences in control ovaries (*white-sKD*)
970 versus *flamΔZAM*-RevIH2 chr III ovaries. Antisense siRNA counts, mapped allowing up to 1
971 mismatch, were normalized per million of genome-mapping siRNAs (RPM).

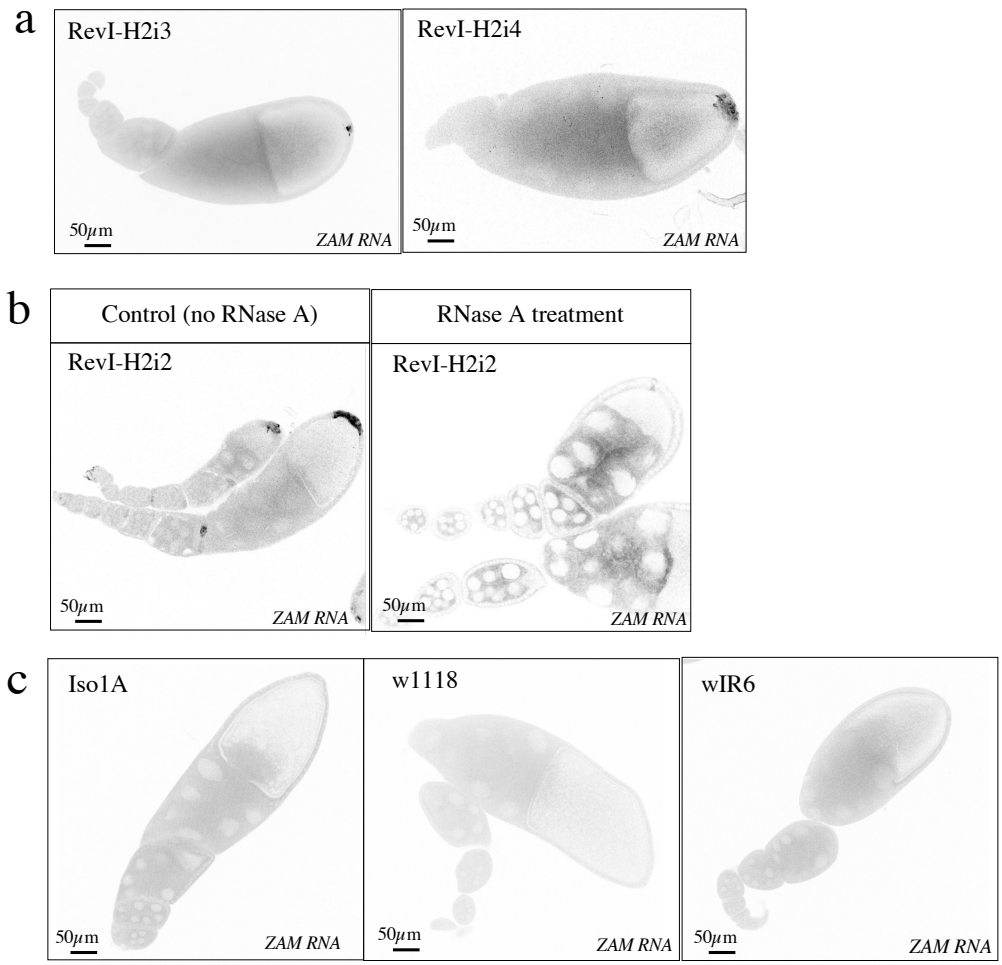


Figure S1 (related to Figure 1)

(a) Color-inverted confocal images of ovarioles showing *ZAM* RNA signal in ovaries (smRNA FISH) from the indicated RevI-H2 isogenic lines. **(b)** Color-inverted confocal images of ovarioles from RevI-H2i2 ovaries showing *ZAM* RNA signal without (left panel) and after (right panel) RNase A treatment. **(c)** Color-inverted confocal images of ovarioles showing that *ZAM* RNA is completely absent in Iso1A, w^{1118}, and w^{IR6} ovaries.

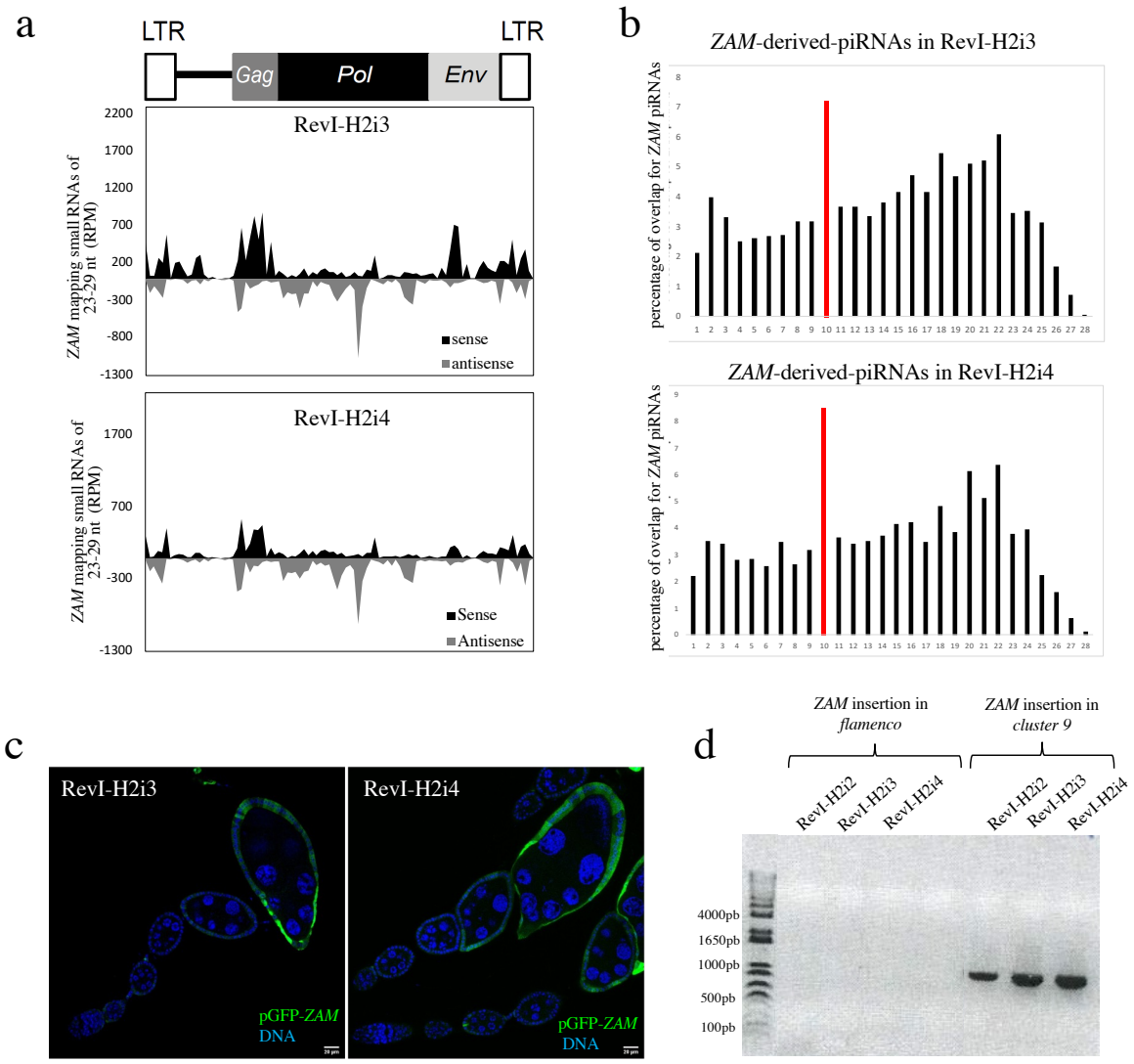


Figure S2 (related to Figure 2)

(a) Density plot of regulatory piRNAs along the *ZAM* sequence in ovaries of the indicated RevI-H2 isogenic lines (up to 3 mismatches). **(b)** Histogram showing the percentage of 5'-overlaps between sense and antisense *ZAM*-derived piRNAs in ovaries of the indicated RevI-H2 isogenic lines. The percentage of 10-nt 5'-overlaps is in red. **(c)** Confocal images of ovarioles with the GFP sensor transgene (GFP in green) and DNA (blue) staining. Ovarioles were from the progeny of a cross between RevI-H2i3 or RevI-H2i4 females and males carrying the pGFP-*ZAM* sensor transgene driven by actin-Gal4. **(d)** Genomic DNA from the RevI-H2 isogenic lines (RevIH-2i2, RevI-H2i3 and RevI-H2i4) was PCR amplified using primers in order to detect the presence of the *ZAM* insertion in *flamenco* or a *ZAM* insertion in piRNA cluster *9* (specific of the RevI-H2 line). Amplification can only occur if *ZAM* is present at the target locus (primer sequences in Table S5). In all tested RevI-H2 isogenic lines, no amplification was observed for *ZAM* in *flamenco*, while a PCR product of ~1000 pb was observed for the *ZAM* insertion in piRNA cluster *9*.

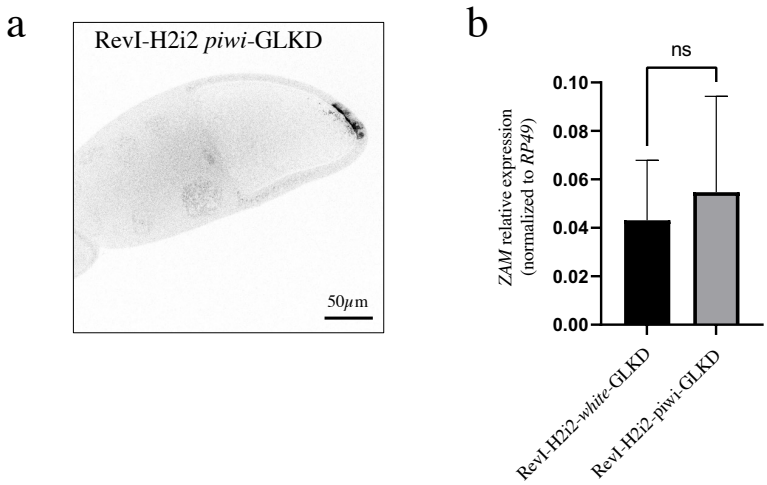


Figure S3 (related to Figure 3)

(a) Color-inverted confocal images of stage 10 egg chambers from RevI-H2i2 *piwi*-GLKD ovaries showing *ZAM* smRNA FISH signal. **(b)** Bar plot showing the fold-changes in steady-state *ZAM* RNA levels in total ovarian RNA from the indicated genotypes (normalized to *rp49*). At least three biological replicates and two technical replicates were used. ns: *p*-value >0.05 (Mann-Whitney test). Error bars indicate the SD.

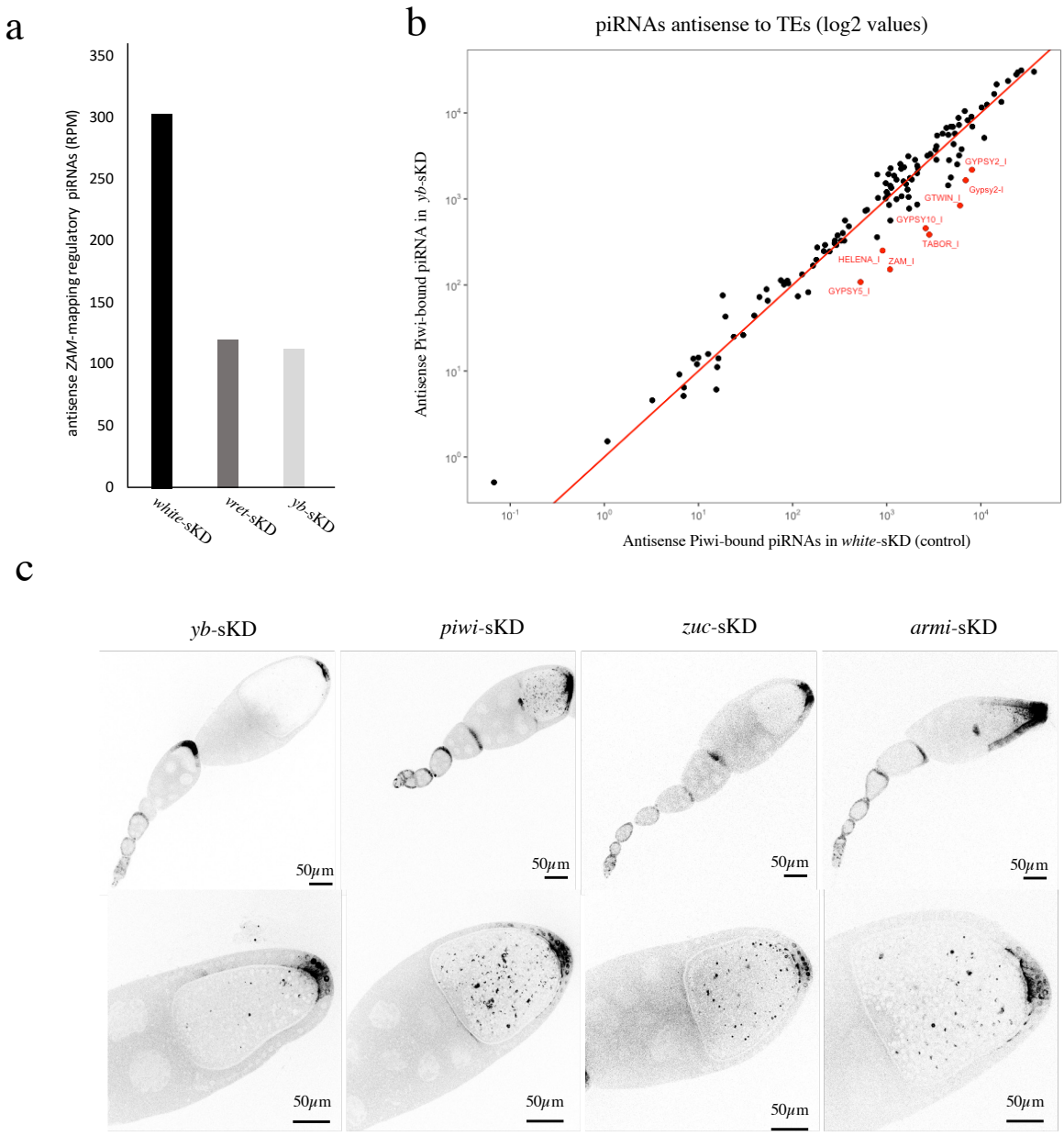


Figure S4 (related to Figure 4)

(a) Amount of antisense piRNAs mapping to *ZAM* (up to 3 mismatches) in *white*-, *vret*-, and *yb*-sKD ovaries (normalized per millions of genome-mapping piRNAs). **(b)** Scatter plot showing the log2 of normalized counts [log2(RPM)] of antisense Piwi-bound piRNAs mapping to individual internal TEs sequence in control ovaries (*white*-sKD) versus *yb*-sKD ovaries. Antisense piRNA counts, mapped allowing up to 3 mismatches, were normalized per million of genome-mapping piRNAs (RPM). *TEs in blue have a *yb*-sKD/*white*-sKD ratio >3; TEs in red have a ratio <0.3. **(c)** Color-inverted confocal images of ovarioles (upper panels) and stage 10 egg chambers (lower panels) from the indicated sKD lines showing *ZAM* RNA expression in follicle cells and in the ooplasm in all conditions.

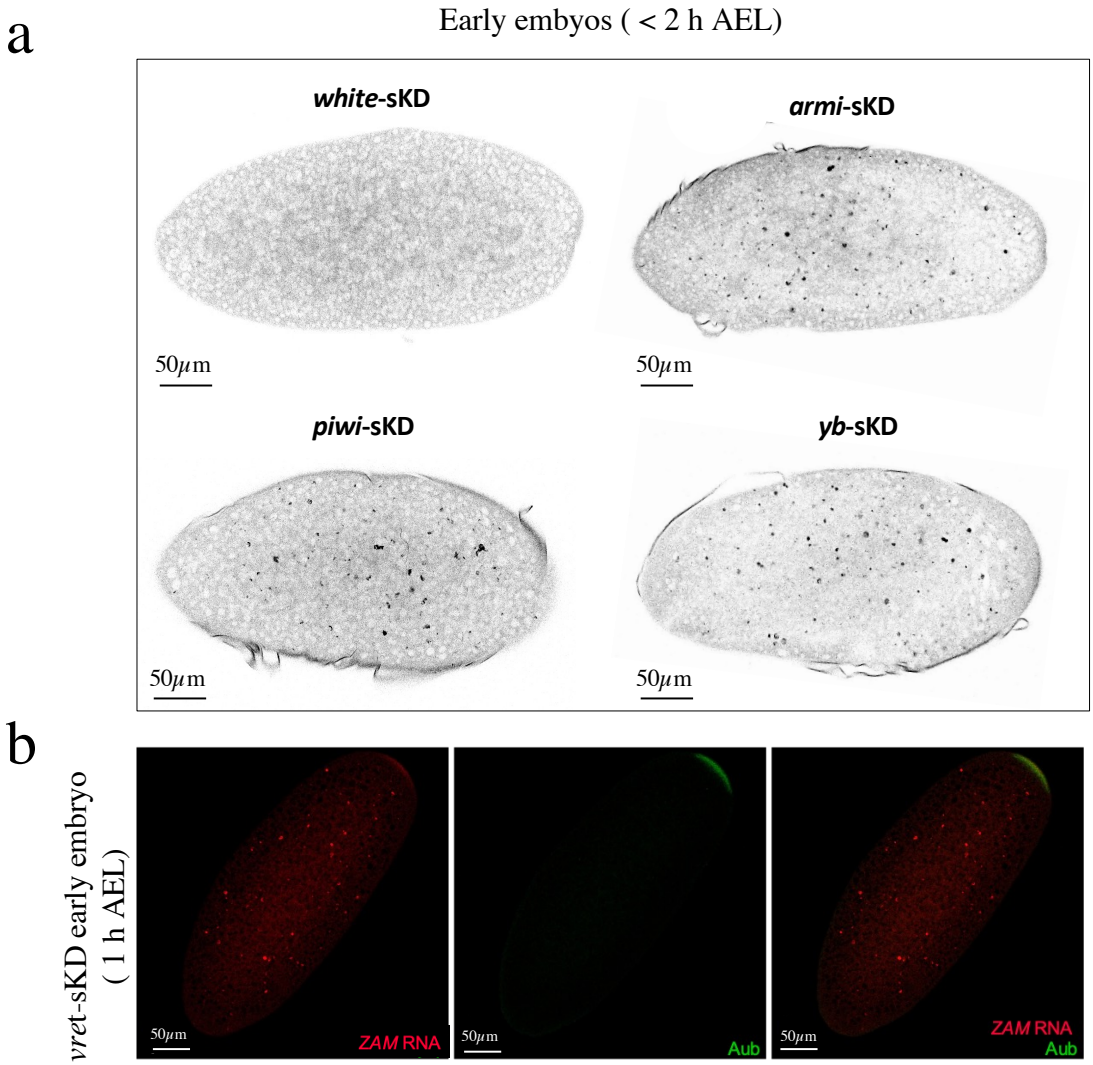


Figure S5 (related to Figure 5):
(a) Color-inverted confocal images of early embryos collected 0-2 hours after egg laying (AEL). Mothers harbor the indicated sKD of the piRNA pathway. *ZAM* RNA signal was detected by smRNA FISH. **(b)** Projection of confocal images of stage 1 embryos (0-1 h AEL) laid by *vret*-sKD females showing *ZAM* RNA (smRNA FISH) and Aub protein immunostaining. Both accumulated at the posterior pole of the embryo where future germ cells cellularize (Z-projection on 25 stacks).

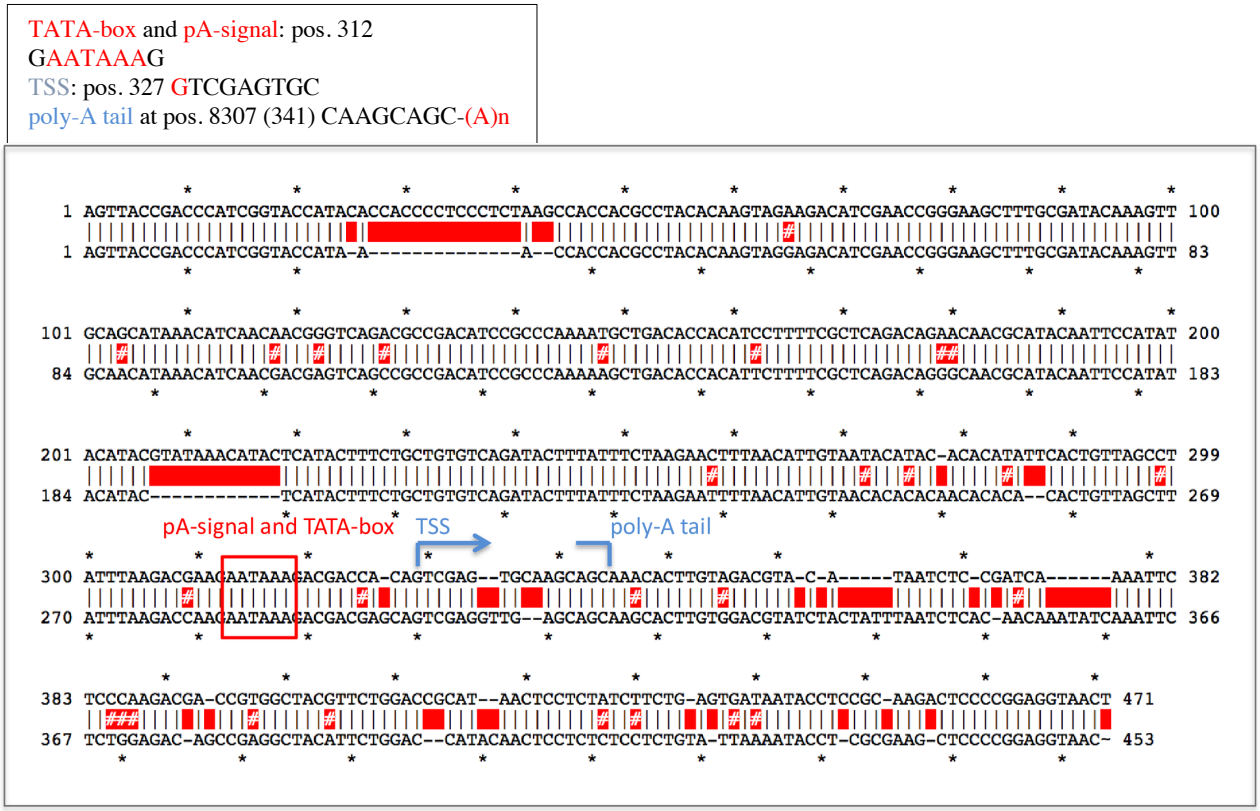


Figure S6 (related to Figure 6):
Sequence alignment of the RevI-H2 consensus *ZAM* LTR and the *ZAM* solo-LTR, which remains after CRISPR-Cas9 deletion of the internal *ZAM* sequence in *flamenco* in the *flamΔZAM* line. The image shows the RevI-H2 consensus *ZAM* LTR (upper sequence) and the *ZAM* solo-LTR (lower sequence) with their TATA-box and polyadenylation signal (pA-signal) at position 312 of the *ZAM* consensus (upper sequence), the transcription start site (TSS, at position 327), the start of the poly-A tail at position 341 (CAAGCAGC-(A)n), according to [25].

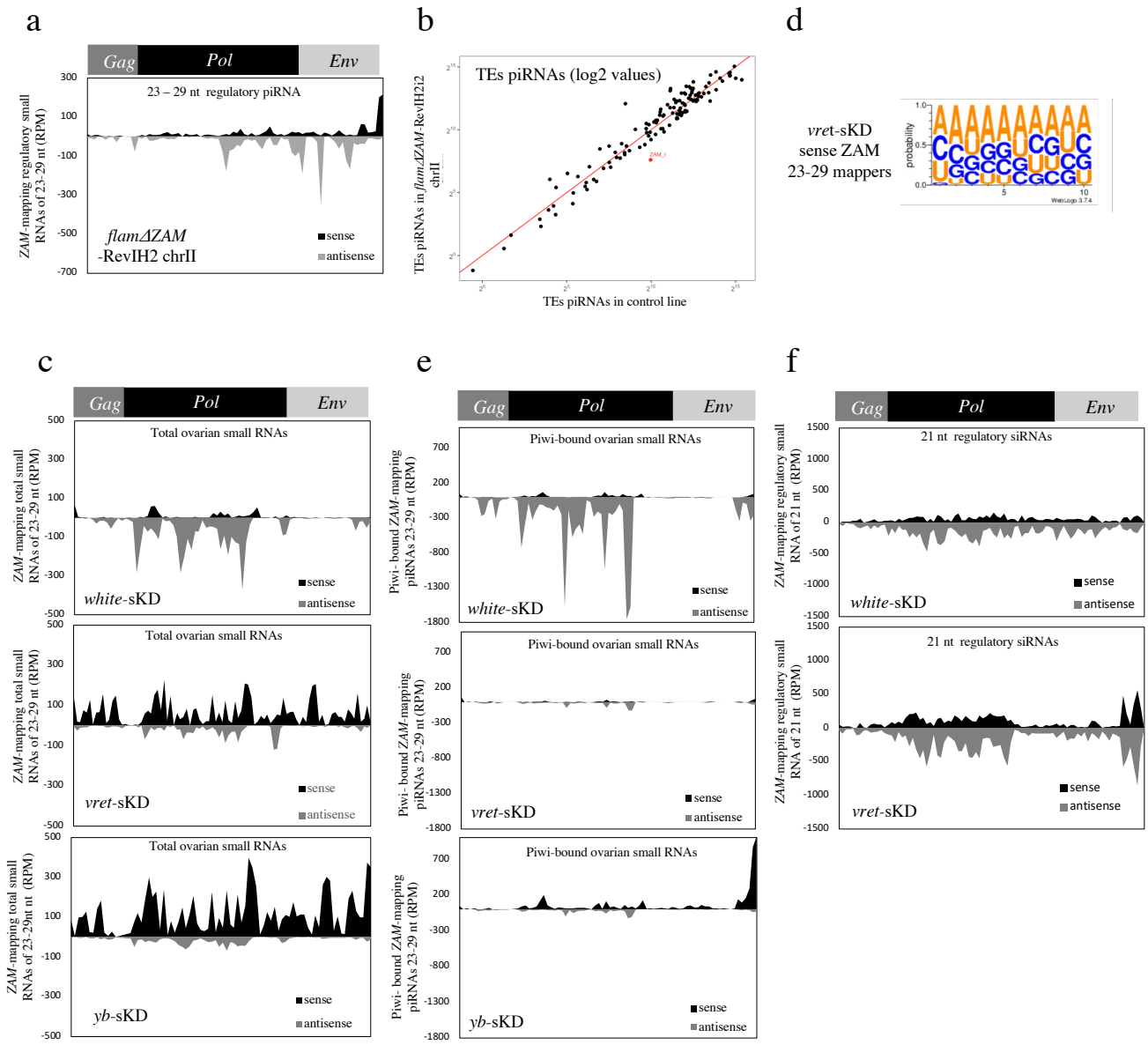


Figure S7 (Related to Figure 7)

(a) Density plot of *ZAM*-mapping regulatory piRNAs (23-29 nt small-RNAs complexed with Argonaute proteins) along the *ZAM* internal sequence in *flamΔZAM*- RevIH2 chr II ovaries (up to 3 mismatches). **(b)** Scatter plots showing the log2 of normalized counts [log2(RPM)] of regulatory piRNAs mapping to individual internal TE sequences in control ovaries (*white*-sKD) versus *flamΔZAM*-RevIH2 chr II ovaries. Antisense piRNA counts, mapped allowing up to 3 mismatches, were normalized per million of genome-mapping piRNAs (RPM). **(c)** Density plot of total small RNAs of 23 to 29 nt along the *ZAM* sequence in *white*-sKD, *vret*-sKD and *yb*-sKD ovaries (up to 3 mismatches). **(d)** Sequence logo for the first ten positions in ZAM-mapping sense small RNAs of 23 to 29nt (from total ovarian small RNA) in *vret*-sKD ovaries. The nucleotide height represents its relative frequency at that position. **(e)** Density plot of *ZAM*-mapping Piwi-bound piRNAs along the *ZAM* internal sequence in *white*-sKD, *vret*-sKD and *yb*-sKD ovaries (up to 3 mismatches). **(f)** Density plot of *ZAM*-mapping regulatory siRNAs (21nt small-RNAs complexed with Argonaute proteins) along the *ZAM* internal sequence produced in *white*-sKD and *vret*-sKD ovaries (0-1 mismatch).

Supplementary Table S1 – ZAM insertions in the RevI-H2i2 genome

number	chromosome	insertion site (Release 6 genome)	orientation	type	locus	comment
1	X	X:1602129-1602133	minus	ZAM-variant	intron of the overlapping Mur-2B and broad genes	ZAM deleted for internal positions 5500-6126 (Repbase ZAM_I)
2	X	X:1980714-1980717	minus	ZAM full-length	intron of gene Hr4	
3	X	X:19249375-19249375	minus	ZAM full-length	intron of gene Hs3st-B	inserted in a Jockey TE which is absent in the reference genome
4	X	X:22257753 or X:22373661	minus	ZAM full-length	piRNA cluster 56 or 13 (George, Jensen et al., 2015)	
5	X	X:23212829	plus	ZAM full-length	piRNA cluster 9 (X:23150001-23545000, George, Jensen et al., 2015)	
6	X	X:23368838-23368841	minus	ZAM-variant	piRNA cluster 9 (X:23150001-23545000, George, Jensen et al., 2015)	ZAM deleted for internal positions 5500-6126 (Repbase ZAM_I)
7	X	X:23513106-23513113	minus	ZAM-variant	piRNA cluster 9 (X:23150001-23545000, George, Jensen et al., 2015)	ZAM deleted for internal positions 5500-6126 (Repbase ZAM_I)
8	2L	2L:8158963-8158966	plus	ZAM full-length	intron of gene PAPLA1	inserted in a NOMAD TE which is absent in the reference genome
9	2L	2L:15773891-15773894	plus	ZAM full-length	intergenic region	
10	2R	2R:1808663-1817084	plus	ZAM full-length	intergenic region	ZAM copy also present in the reference genome Release 6
11	2R	2R:6793620-6793623	plus	ZAM with an internal deletion	intron of gene Epac	ZAM deleted for internal positions 2332-5754 (Repbase ZAM_I)
12	2R	2R:11601808-11601811	minus	ZAM full-length	intron of gene toutatis	
13	3L	3L:4299131-4299137	minus	ZAM full-length	intron of gene dyl	ZAM deleted within the 5'-UTR, internal positions 60-1218 (Repbase ZAM_I)
14	3L	3L:19980882-19980883	minus	ZAM-variant with additional internal deletion	intergenic region	ZAM deleted for internal positions 13-2993 and 5500-6126 (Repbase ZAM_I)
15	3R	3R:12079947-12079950	plus	ZAM-variant	intron of gene ATP8B	ZAM deleted for internal positions 5500-6126 (Repbase ZAM_I)
16	unknown	NW_007931121.1:45136-45142	plus	ZAM full-length	rDNA	
17	unknown	possibly close to X:23239427	(plus)	ZAM with internal deletions	possibly piRNA cluster 9 (X:23150001-23545000, George, Jensen et al., 2015)	ZAM deleted for internal positions 2332-5754 and 6116-7265 (Repbase ZAM_I)
18	unknown	NW_001845412.1:1075-1078	minus	ZAM with an internal deletion	unknown	ZAM deleted for internal positions 2929-5615 (Repbase ZAM_I)

1075

Supplementary Table S2 – Drosophila lines

Category	*Gene name*	*CG_Number*	*Stock number*	*Genotype*	*Source*	*Notes*
RNAi line	white	CG2759	BDSC_35573	y[1] sc[*] v[1] sev[21]; P{y[+t7.7] v[+t1.8]=TRiP.GL00094}attP2	Bloomington	
RNAi line	armitage	CG11513	BDSC_34789	y[1] sc[*] v[1] sev[21]; P{y[+t7.7] v[+t1.8]=TRiP.HMS00098}attP2	Bloomington	
RNAi line	yb	CG6272	VDRC_110056	P{KK105851}VIE-260B	VDRC	
RNAi line	Piwi	CG6122	BDSC_37483	y[1] sc[*] v[1] sev[21]; P{y[+t7.7] v[+t1.8]=TRiP.GL00626}attP40	Bloomington	
RNAi line	vreteno	CG4771	VDRC_34897	w1118; P{GD11424}v34897	VDRC	
RNAi line	zucchini	CG12314	BDSC_35227	y[1] sc[*] v[1] sev[21]; P{y[+t7.7] v[+t1.8]=TRiP.GL00111}attP2	Bloomington	
RNAi line	Argonaute 3	CG40300	BDSC_35232	y[1] sc[*] v[1] sev[21]; P{y[+t7.7] v[+t1.8]=TRiP.GL00117}attP2	Bloomington	
RNAi line	Aubergine	CG6137	BDSC_35201	y[1] sc[*] v[1] sev[21]; P{y[+t7.7] v[+t1.8]=TRiP.GL00076}attP2	Bloomington	
Wildtype	white (w1118)		BDSC_3605	w1118;;;	Bloomington	
Nos-Gal4 driver			BDSC_4937	w[1118]; P{w[+mC]=GAL4::VP16-nos.UTR}CG6325[MVD1]	Bloomington	
Traffic jam gal4 driver				w, Tjgal4/cyO; ;	Unknown	
RevI-H2i1					Produced in the lab	Isogenic line derived from the RevI-H2 stock
RevI-H2i2					Produced in the lab	Isogenic line derived from the RevI-H2 stock
RevI-H2i3					Produced in the lab	Isogenic line derived from the RevI-H2 stock
RevI-H2i4					Produced in the lab	Isogenic line derived from the RevI-H2 stock

1076

Supplementary Table S3 – smRNA FISH probes

Oligo number	ZAM (this paper) 5'- 3'	Burdock (from Senti and al,.2015) 5' - 3'
1	GAAAAAGAAATGTTGGCAAT	AGGTTATCACTGTCACTCAT
2	AATATGGGCAGTAAAATACT	GCTTGACACTAGGTTGTCTA
3	AGACCATATATTTATGGCGT	TGGTGCGCTGACCATTTATT
4	TACAGATCACAAGCCACTAA	TTTTTTCTGCACTGTTTTGC
5	TGAATTTCAAAGAACCCAAC	CCACCAATTATCTGATGACT
6	AAAAATAATTCGTTGGAGAC	CGCTAGTTGTCTTTGTTTTT
7	TCCAACTCATGGAATACAAT	AACTGAGCTTCAAATTCCAT
8	TCACAAGAAAGGTTCACAAA	AACAGCACTCTCTACGATCG
9	TGTAATTGCAGACGCCTTAA	TGATTGTTTTTGGACTGCGA
10	GTAGAGCGGACCCAAATTTA	TCATTCATCTCCTGCAATTG
11	AACTACAACGAAACACTGAC	GGTGTTCACTGTTAATTTCC
12	TGTTAAGCCTTGCCCCACAT	CTACATAAGTTTCCACCTCT
13	CCGAAAAACCTATTAACGAA	TATATCTAGAGGCTCGCTAC
14	AACACGCAACTCATACTAGA	GCCATCAAAATCTGGCAGAG
15	TCTTACCAAACTACAACACC	CTTCCCTCGTAATCTTTGAA
16	AGGAAAAAATATTCACAGCC	CCAAGAGCTTGGTAAAATGT
17	AAAATCTTGAAAGGAACCCT	AAGCCAATACTGTATTCGCT
18	AAAACCTAACAGGATTTGCG	GCTTTGAAATGTAACGGAGT
19	CATTCTTGGCGGACGATAAT	TGTGAAATCAAGACGGCTGA
20	TCCGCATTAATCGAAAAAGC	TAGATAGATCGGCCTTTTGT
21	TTCTCCACGAAATCGTAGGA	GTCGCAAAGTTGATAGCTCT
22	ACCCCGAACAAAACAAATTC	TCAGTAAGAGTCATGTCTCC
23	GGAATATCACACTAACAGCA	TCTTGTTGGTAAGTAGGGTC
24	GAGGCATAGACGAAACATTC	GCCAAGGCACTATCAAATGT
25	TTCACCTCAAACGAGAAACC	GTCCTGTACTTTTCATTCAG
26	ACTTCCCCAATATGAAAAAC	CTCCGGTTACAAATACACGT
27	AATTAGGAATTGCGAAACCT	ATGTCGCTCAACGATTTCTT
28	GTCTAAAACTCAAATACGAC	ATCTTTTGGACCTTTTGCAA
29	AGGCAACCACAAAATATAGT	TGGTTCGACTCGACCTCTTG
30	ATTTGAAACCCCAGAAACCC	AAGGGCGAATTGGTAACGCT
31	GAAACCCCTCGACATAATAC	CCCTGTCTCCAATATTTTTA
32	ACCTGACAATCATAGACAAA	TGCCTTTGTTCGATTTTCTG
33	TCTCAAAATTCGCAGCTGTC	TGAGTTTCTATCCTTATCGC
34	TACCCCATCCCAAATAGAAA	TTTTTAGTTTGCATGGGCAT
35	CGGCATCAATTGCATCAAAG	CACCTGACGCTTGCTAAAAT
36	TTCAGTCAATTCGGACTACC	GGATTTGTTAACTGGACTGG
37	CAAAAAACTAATACACGACC	TTCTTCTAGATCGCATGGAT
38	AGGAGTAGAATTTTGCAACG	GGAGTTTGTCCAAAATTTCC
39	CCATGTCACATCCTTCCAGC	AATATTTCCCTGAGTCGGAA
40	AGTAATTCTCCAGTAGAACG	GGTTGCGATGCGAATTGATT
41	ACACTCCTCTTTGACAGAGA	CTGACTTGTGTTTTGCTGTT
42	AATACTAGACACACGGAAAA	TTGCGCATACCCAAAATGTC
43	ACACAAATTACCTACAGACC	CACTCGTTGGTCTTTTTGAT
44	ACGAAGAAATAATGTCAGAA	TTTTGGCCCTGTAAACCTTG
45	ATAACAACGCAATCCACTCC	GGTAAGTAGTTGATCCTCTG
46	CCACCAAACACACCCCTTTT	CTCACATTGACCTTTCTCAT
47	TGGTAGGACCCATTTATTCG	CCTTTTGATAACCGTCTGTA
48	ATACCCAATAATGAGCATGA	ATAATCATCAACCTCCGCTT

1077

Supplementary Table S4 – Antibodies used

Antibody	Source	Dilution	Fluorochrome	Species	Reference
Anticorps primaires					
Goat anti-GFP	abcam	1/1000			ab5450
Rabbit anti-Aubergine	Homemade (Eurogentec) - polyclonal antibody	1/1000			
Rabbit anti-Gag-ZAM	Home-made	1/100			
Rabbit anti Env-ZAM	Home-made	1/1000			
Secondary antibody					
Alexa Fluor 488 donkey anti-goat IgG (H+L)	Invitrogen (Fisher)	1/1000	Alexa 488	goat	A11055
Donkey anti-rabbit Alexa 488 IgG (H+L) 2mg/ml	life tech	1/1000	Alexa 488	rabbit	A21206
Goat anti-rat Cy3	Jackson Immunoresearch	1/1000	Cy3	rat	112-165-143
Fluorescent dyes					
DAPI	Sigma	1/500			D8417-1MG

1078

Supplementary Table S5 - Primers used

	Name	Sequence	Used for
Rp49	rp49_R	AAACGCGGTTCTGCATGAG	RT-qPCR
	rp49_F	GACGCTTCAAGGGACAGTATCTG	
Burdock	Burdock_F	GGACAAACTCCATTTCCGACTC	RT-qPCR
	Burdock_R	TCCCTGAGCCTGACTTGTGT	
Gypsy	Gypsy-F	CGCTCGAAGGTTACCAGGTAGGTTC	RT-qPCR
	Gypsy_R	CTTCACGTTCTGCGAGCGGTCT	
ZAM	Zam_env_F	CTCCCAATCATCTCCTCCAA	RT-qPCR
	Zam_env_R	AGAAGTCGGGGTGCTCTTCT	
ZAM-flam	ZAM-FLAM_F	TCGCTTAGGAGGGTAGCTTG	CRISPR SCREEN
	ZAM-FLAM_R	CTTAACACCTACTGTCAGTGC	
ZAM insertion in *flamenco*	flam_R	ctagtgtctttggtaccatgg	Screen
	flam_F	tatcaactatcacgactatcact	
ZAM insertion in *cluster 9*	ZAM insertion_R	CTCACCATTTCCTCCTTGAC	Screen
	ZAM insertion_F	CTCCCAATCATCTCCTCCAA	

1079

Supplementary Table S6 – small-RNA seq

Small RNA seq
- sRNA from RevI-H2i1
- sRNA from RevI-H2i2
- sRNA from RevI-H2i3
- sRNA from RevI-H2i4
- sRNA- from RevI-H2i2- nos-Gal4- expressed dsRNA-white
- sRNA- from RevI-H2i2- nos-Gal4- expressed dsRNA Ago3
- sRNA- from RevI-H2i2- nos-Gal4- expressed dsRNA Piwi
- sRNA- from RevI-H2i2- nos-Gal4- expressed dsRNA Zuc
- sRNA from tj-Gal4-expressed dsRNA-Vret
- sRNA from tj-Gal4-expressed dsRNA-white
- sRNA from CRISPR flamΔZAM
- Total RNA from tj-Gal4-expressed dsRNA-Vret
- Total RNA from tj-Gal4-expressed dsRNA-white
- Total RNA from tj-Gal4-expressed dsRNA-yb

1080