

# Polycomb Group-Dependent, Heterochromatin Protein 1-Independent, Chromatin Structures Silence Retrotransposons in Somatic Tissues Outside Ovaries

J. Dufourt<sup>1</sup>, E. Brassat<sup>2</sup>, S. Desset<sup>1</sup>, P. Pouchin<sup>1</sup>, and C. Vaury<sup>1,\*</sup>

Clermont Université, Université d'Auvergne, BP 10448, F-63000 Clermont-Ferrand, France; Inserm, U 931, F-63001 Clermont-Ferrand, France and CNRS, UMR 6247, F-63001 Clermont-Ferrand, France<sup>1</sup> and Clermont Université, Université d'Auvergne, Institut Universitaire de Technologie, BP 10448, F-63000 Clermont-Ferrand, France; Inserm, U 931, F-63001 Clermont-Ferrand, France and CNRS, UMR 6247, F-63001 Clermont-Ferrand, France<sup>2</sup>

\*To whom correspondence should be addressed. Tel: +33 4-17-81-84. Fax: +33 4-27-61-32.  
E-mail: chantal.vaury@u-clermont1.fr

Edited by Osamu Ohara  
(Received 20 May 2011; accepted 8 August 2011)

## Abstract

**Somatic cells are equipped with different silencing mechanisms that protect the genome against retrotransposons. In *Drosophila melanogaster*, a silencing pathway implicating the argonaute protein PIWI represses retrotransposons in cells surrounding the oocyte, whereas a PIWI-independent pathway is involved in other somatic tissues. Here, we show that these two silencing mechanisms result in distinct chromatin structures. Using sensor transgenes, we found that, in somatic tissues outside of the ovaries, these transgenes adopt a heterochromatic configuration implicating hypermethylation of H3K9 and K27. We identified the Polycomb repressive complexes (PRC1 and 2), but not heterochromatin protein 1 to be necessary factors for silencing. Once established, the compact structure is stably maintained through cell divisions. By contrast, in cells where the silencing is PIWI-dependent, the transgenes display an open and labile chromatin structure. Our data suggest that a post-transcriptional gene silencing (PTGS) mechanism is responsible for the repression in the ovarian somatic cells, whereas a mechanism that couples PTGS to transcriptional gene silencing operates to silence retrotransposons in the other somatic tissues.**

**Key words:** retrotransposons; *Drosophila*; heterochromatin; Polycomb

## 1. Introduction

Due to their ability to transpose to virtually any genomic site, transposable elements (TEs) have the ability to generate deleterious mutations in the host genome. In response, host cells have devised strategies to control TE activity. RNA interference (RNAi), which has a prominent role in this control, is triggered by small RNAs. Several families of small RNAs have now been reported, but it appears that *D. melanogaster* mainly has two RNA silencing mechanisms that repress transposons expression: endo-siRNAs (endogenous small interfering RNAs) and piRNAs. piRNAs (for PIWI-interacting RNAs) are mainly derived from TE antisense strands and are produced from discrete genomic loci.<sup>1</sup>

piRNAs are from 26 to 30 nucleotides (nts) in length and have been reported in the reproductive apparatus of *Drosophila*, mice, rats, and humans.<sup>2,3</sup> Endo-siRNAs are ~21 nts long and associate with the argonaute protein Ago2. Like piRNAs, endo-siRNAs are mainly derived from retrotransposons and other genomic repetitive elements. However, contrasting with piRNAs, endo-siRNAs are likely to be expressed ubiquitously.<sup>4–7</sup> It is currently proposed that piRNAs and endo-siRNAs have a germline-specific function and a soma-specific function, respectively, in the establishment of TE silencing.<sup>4</sup>

Besides these post-transcriptional regulations implicating RNAi silencing pathways, TEs can also be transcriptionally silenced. A range of chromatin modifications suppress their transcription, including histone

modifications and altered chromatin packing.<sup>8</sup> Furthermore, in both plants and mammals, DNA methylation on cytosine residues is another important signal that represses TE transcription and provides a mechanism for inheritance of TE silencing.

We recently reported that PIWI is necessary for silencing of ZAM and *Idefix*, two long terminal repeat retrotransposons from *Drosophila melanogaster*, in the ovarian follicle cells, but is dispensable in other somatic lineages throughout fly development.<sup>9</sup> However, it remains unclear whether a specific chromatin structure is deposited on these elements and in which tissues. Furthermore, since DNA methylation is absent from *D. melanogaster*, nothing is currently known about the mechanisms involved to maintain their silencing through cell divisions. Here, we investigated further the chromatin structures necessary to silence *Idefix*.

Like most transposable families, many copies of *Idefix* are detected within the *Drosophila* genomes. Active copies are generally dispersed on chromosomal arms but many copies corresponding to *Idefix* vestiges also exist and accumulate in the heterochromatic regions. It was then of prime importance to follow the silencing exerted on a copy known to a target of the RNAi pathways. To this aim, we made use of sensor transgenes carrying a GFP (green fluorescent protein) reporter gene followed by an *Idefix* fragment that recapitulates the control exerted on active copies of *Idefix*.<sup>9</sup> Analysis of the chromatin structure deposited on the transgene could then be easily explored over the GFP sequence that is unique in the genome.

We show that in somatic cells in contact with the germline, the silencing needs to constantly target the *Idefix* sequence to be maintained. By contrast, in the other somatic tissues referred to as the soma, the silencing is stably established even when the TE fragment is excised. Our data further indicate that a compact chromatin structure likely associated with Polycomb (PC)-dependent chromatin structures is deposited.

## 2. Materials and methods

### 2.1. *Drosophila* strains and transgenic lines

The following mutant fly stocks were used: ago2<sup>414</sup>,<sup>10</sup> ago2<sup>51B</sup>,<sup>11</sup> dcr2<sup>[L811fsX]</sup>,<sup>12</sup> PIWI<sup>2</sup> and PIWI<sup>3</sup>,<sup>13</sup> r2d2<sup>1</sup>,<sup>14</sup> and loqs<sup>[f00791]</sup>.<sup>15</sup> Bl279 (+;

+/+; MKRS hs-Flip/Tm6) comes from Bloomington stock. pc<sup>1</sup> comes from Cavalli's lab.

About 419 bp homologous to the *Idefix gag* coding region (1003–1422) were inserted either in a sense (tGglds) or anti-sense orientation (tGgldas) according to *gfp* transcription within the pUAST-*gfp* vector. Three and six independent transgenic lines were analyzed for tGglds and tGgldas, respectively (Table 1).

### 2.2. Chromatin immunoprecipitation assays

Chromatin immunoprecipitation (ChIP) was performed using the standard procedures, which was modified as follows: for each immunoprecipitation, three OD<sub>260</sub> of chromatin in ChIP buffer was pre-incubated in the presence of 80 μl of Protein A-Agarose (PAA) beads (Millipore) for 30 min at 4°C. PAA were removed, antibodies were added (a control in the absence of antibody named Input was included), and samples were incubated overnight at 4°C in a rotating wheel. Then, 60 μl of PAA were added and incubation was continued for 1 h at 4°C. Samples were washed in saline concentration buffer. Chromatin was eluted from PAA in 500 μl of elution buffer (sodium dodecyl sulfate and NaHCO<sub>3</sub>) at room temperature for 15 min. The eluate was incubated overnight at 65°C to reverse cross-links and treated by proteinase K for 1 h at 50°C. Samples were phenol–chloroform extracted and ethanol precipitated. DNA was resuspended in 40 μl of 10 mM Tris (pH 7.5).

qPCR (quantitative polymerase chain reaction) (ABI 7300) from at least two biological-independent extractions and three technical replicate was performed using SYBR Green (Sigma) and Platinum Taq DNA Polymerase (Invitrogen). Enrichment was calculated relative to *RpL32* and values were normalized to input measurements; the error is indicated by the standard error of mean (SEM). Antibodies used include: H3K9-di-methylation #07-441, H3K9-tri-methylation #07-523, H3K9/K14-acetylation #06-599, H3K27-di-methylation #06-421, H3K27-tri-methylation #06-449, H4K20-mono-methylation #07-440, H2A-ubi #05-678 from Millipore, heterochromatin protein 1a (HP1a) (C1A9) from DSHB, and PC and PH antibodies were provided by G. Cavalli Lab. The following primers were used uptssF: CCAAGCTTTGCGTACTCG and uptssR: CCGTGGGGTTTGAATTAAC for the upstream *gfp* TSS; gfpF: ACCATTACCTGTCCACACAA and gfpR:

**Table 1.** Characteristics of tGglds and tGgldas lines in somatic tissues and their homologous piRNA and endo-siRNAs (esiRNAs) reported in the databases<sup>4,31</sup>

	Number of lines	Silencing in follicle cells	Silencing in larvae	Number of sense piRNAs	Number of antisense piRNAs	Number of sense esiRNAs	Number of antisense esiRNAs
tGgld sense	3	On	Off	0	194	0	0
tGgld antisense	6	On	On	34	0	3	0

CCAGCAGCTGTTACAACTC for the *gfp* gene; and Rpl32F: CCGCTTCAAGGGACAGTATC and Rpl32R: GACAATCTCCTTGCGCTTCT for the control gene *Rpl32*.

### 2.3. Fluorescent staining and microscopy

Ovaries were dissected and fixed in 5% formaldehyde in phosphate buffer solution for 15 min each hour after the heat shock. GFP was viewed in whole-mount ovaries using the 488 nm filter set of the ZEISS LSM 510 confocal microscope.

### 2.4. In situ hybridization

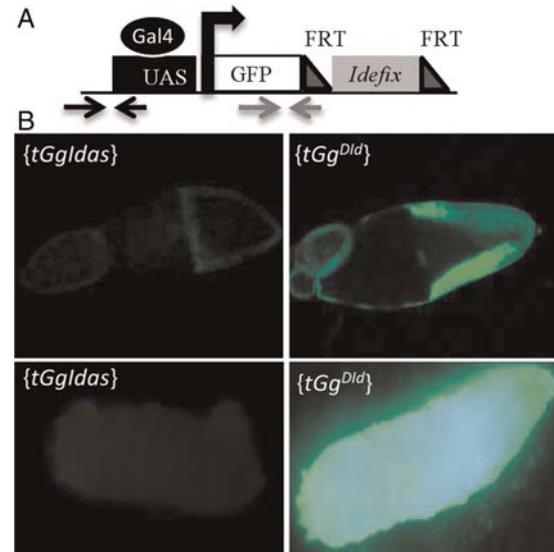
FISH (fluorescence in situ hybridization) was performed on larval imaginal discs, as described previously.<sup>16</sup> Detailed coordinates of the PCR fragments used to produce the probes can be provided upon request. Three-dimensional images were acquired on a Leica SP5 confocal microscope using a 63X objective. The colocalization was searched for automatically by using a script written on Imaris XT software (Supplementary Fig. S2). Three hundred and five and 179 nuclei were examined from seven tGgldas and seven tGg<sup>Did</sup> larval discs, respectively.

## 3. Results

### 3.1. *Idefix* sequence is continuously required for its silencing in the follicle cells but not in other somatic tissues

To characterize further the silencing targeting *Idefix* in the somatic tissues, we used a sensor transgene called tGgldas. In this transgene, the *gfp* reporter gene driven by the UAS promoter was linked to a fragment of the *gag* gene of *Idefix* (Fig. 1A).<sup>9</sup> When tGgldas is driven by a ubiquitous driver, Actine-Gal4, a GFP signal is detected neither in the follicle cells nor in other somatic tissues throughout development: larvae, pupae, and adults (Fig. 1B). When *Idefix* is excised and a new line established with the sensor transgene lacking *Idefix* sequence, then GFP is fully recovered (see tGg<sup>Did</sup> line, Fig. 1B). Since tGgldas is repressed in all the somatic tissues within and outside of the ovaries because of the *Idefix* sequence, we used it to test whether its silencing might switch to an active state as soon as the targeted fragment of *Idefix* has been excised.

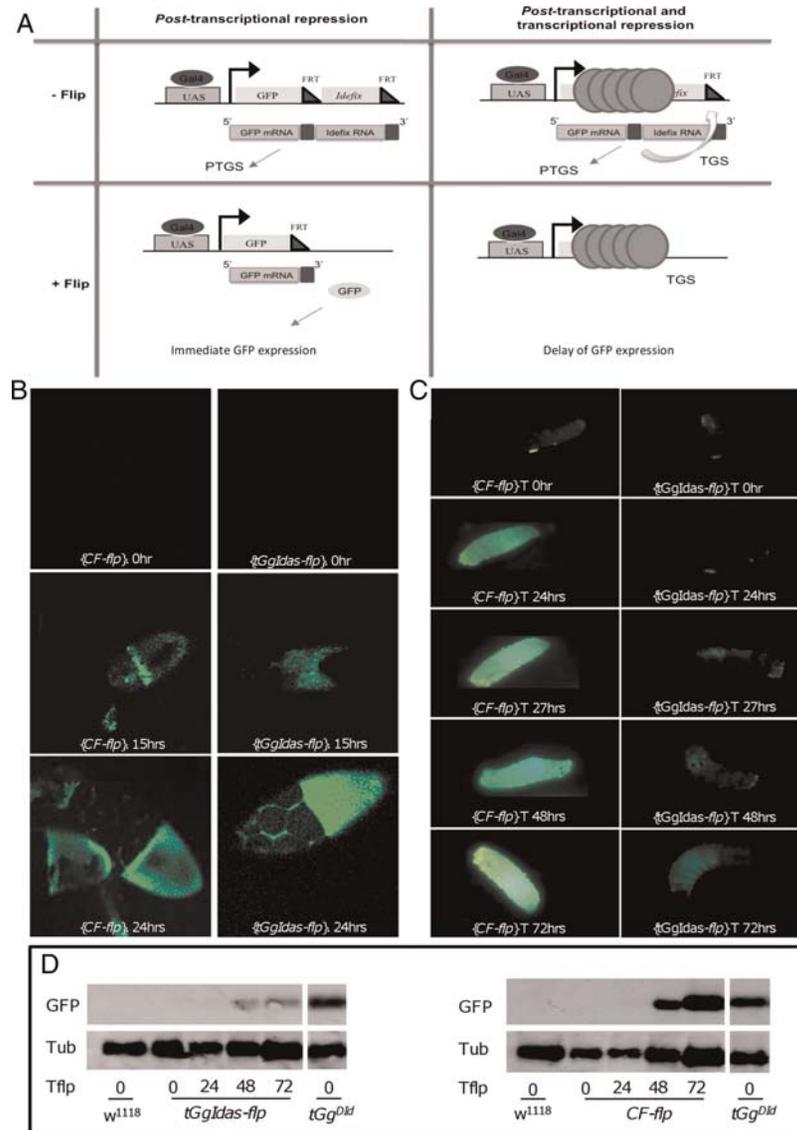
We made use of the *Flp*/FRT (Flip/FLP recognition target) system and its selective induction by the heat shock *Flp* (HS-*Flp*) driver. Heat shocks were performed for 1 h on third instar larvae. If tGgldas silencing is stably established in somatic tissues, GFP repression should persist through cell divisions even after the *Idefix* fragment has been flipped out. Alternatively, if the silencing is labile and depends



**Figure 1.** Structure and silencing of tGgldas. (A) Structure of tGgldas: the *gfp* reporter gene is depicted by a white rectangle and its TSS by an arrow. The *Idefix* portion (grey) are inserted between two FRT sites (dark grey triangles). Position of the qPCR primer sets taken in the *gfp* sequence and upstream of the TSS are indicated by arrows. (B) An example of GFP silencing targeting tGgldas in the follicle cells (above) and in larvae (below) is presented on the left panels; expression of the reference line carrying a sensor transgene, tGg<sup>Did</sup>, in which the *Idefix* fragment has been excised, is presented on the right.

on a regulatory pathway that needs to constantly target *Idefix* sequence to establish the repression, then the silencing should be lost and the GFP expression recovered during the life of the fly as soon as the *Idefix* sequence is flipped out (Fig. 2A).

In a first set of experiments, we determined the time necessary for the FLP recombinase to excise a DNA fragment flanked by FRT sequences after induction by heat shock. We used a triple transgenic line designated as CF (for control *Flp*). A first transgene is the FLP recombinase under the control of the heat shock promoter (HS-*Flp*). A second transgene (Actin>FRT-*yellow*-FRT>Gal4) carries the ubiquitous actin promoter upstream of a *yellow* gene flanked by FRT sequences and a Gal4 reporter gene placed downstream. When *yellow* is present, Gal4 is not expressed. The third transgene of the CF line carries the UAS promoter placed upstream of the *gfp* reporter gene. If a heat shock is performed in this triple transgenic line, the *yellow* gene flanked by FRT sites is excised. The Gal4 protein is then expressed, and in turn, activates the UAS-*gfp* transgene giving rise to fluorescence. A kinetic of GFP appearance in both the follicle and larval cells at time intervals of 1 h after heat shock was performed, and the time necessary to recover GFP expression was determined in both tissues. We estimated that 15 and 24 h was the beginning of expression of the UAS-*gfp* reporter gene in the follicle cells and the larval tissues, respectively, and a complete expression was recovered after



**Figure 2.** The silencing exerted on tGldas is lost in the follicle cells as soon as the targeted retrotransposon sequences are excised, but remains in the soma. (A) Tested hypothesis: (left) if tGldas silencing is labile and depends on a regulatory pathway that needs to constantly target Idefix sequence to establish the repression, then the silencing should be lost and the GFP expression recovered even during the life of the fly as soon as the Idefix sequences is flipped out; (right) if the silencing is stably established in tissues due to a compact heterochromatic structure able to maintain a cellular memory, GFP repression should persist through cell divisions even after the Idefix fragment has been flipped out. (B) (left column) When the CF line [*HS-Flp*, *Actin*>*FRT*>*yellow*>*FRT*>*Gal4*, *UAS-gfp*] is submitted to a heat shock treatment, GFP expression is detected in the follicle cells as soon as 15 h after treatment (second panel) and a full expression at 24 h (third panel). (right column) When the tGldas line is submitted to the heat shock, GFP is detected as soon as 15 h after treatment and a full expression at 24 h compared with tGg<sup>Did</sup> (first panel). (C) (left column) When CF larvae are submitted to a heat shock treatment, GFP expression is detected 24 h after treatment; GFP is fully recovered after 48 h (fifth panel). (right column) After heat shock, GFP expression remains faint in tGldas larval tissues even 72 h after heat shock. (D) Western blot analysis for GFP expression in larvae submitted to heat shock kinetics. Western blotting was performed with an antibody that specifically detects GFP. The w<sup>1118</sup> line with no *gfp* transgene is used as a negative control and tGg<sup>Did</sup> is used as a positive control for GFP expression. The transgenic lines tGldas (left panel) and the test line CF (right panel) are analyzed at time intervals presented below the figure: 0, 24, 48 and 72 h. Tub is a loading control.

24 h in the follicle cells and 48 h in larvae (Fig. 2B and C, left columns).

Then, we determined how long the silencing of the tGldas transgene persists after *Idefix* has been flipped out. A heat shock was performed on flies or larvae with the genetic backgrounds [*Actin-Gal4/+*; tGldas/*HS-Flp*] or [*Actin-Gal4/+*; tGldas/TM6].

Ovaries were dissected every hour from the heat-shocked flies and GFP expression examined in the follicle cells. We found that fluorescence was detected in some patches 15 h after the heat shock (Fig. 2B, right column, middle panel) and a strong green signal in the entire follicular epithelium was recovered after 24 h (Fig. 2B, right

column, lower panel). These data indicate that the repression exerted in the follicle cells is reversible as soon as the targeted retrotransposon sequence is excised.

When examined in larvae, GFP expression remained undetectable 24 s after heat shock induction in flies [Actin-Gal4/+; tGgldas/HS-Flp]. Fluorescence was hardly detected 48 h after treatment and remained very faint even after 72 h, indicating that a full expression was not recovered (Fig. 2C, right column).

As controls, we verified first that the *Idefix* fragment has been indeed excised after the heat shock treatment (PCR amplifications not shown), and second, that GFP expression is not due to the heat shock treatment, but indeed to *Idefix* excision since no GFP was detected in cells with the [Actin-Gal4/+; tGgldas/TM6] genetic background after heat shock.

Then, proteins were extracted from the treated larvae at different time intervals from 0 to 72 h and western blot experiments were performed with an antibody raised against GFP. The reference line established with the tGg<sup>Did</sup> transgene with no *Idefix* sequence and the w<sup>1118</sup> line with no *gfp* transgene were used as positive and negative controls of GFP expression, respectively. In accordance with results described above, western blots revealed that in [Actin/+; tGgldas/hs-flp] larval tissues (denoted tGgldas-flp on Fig. 2D), the level of GFP was

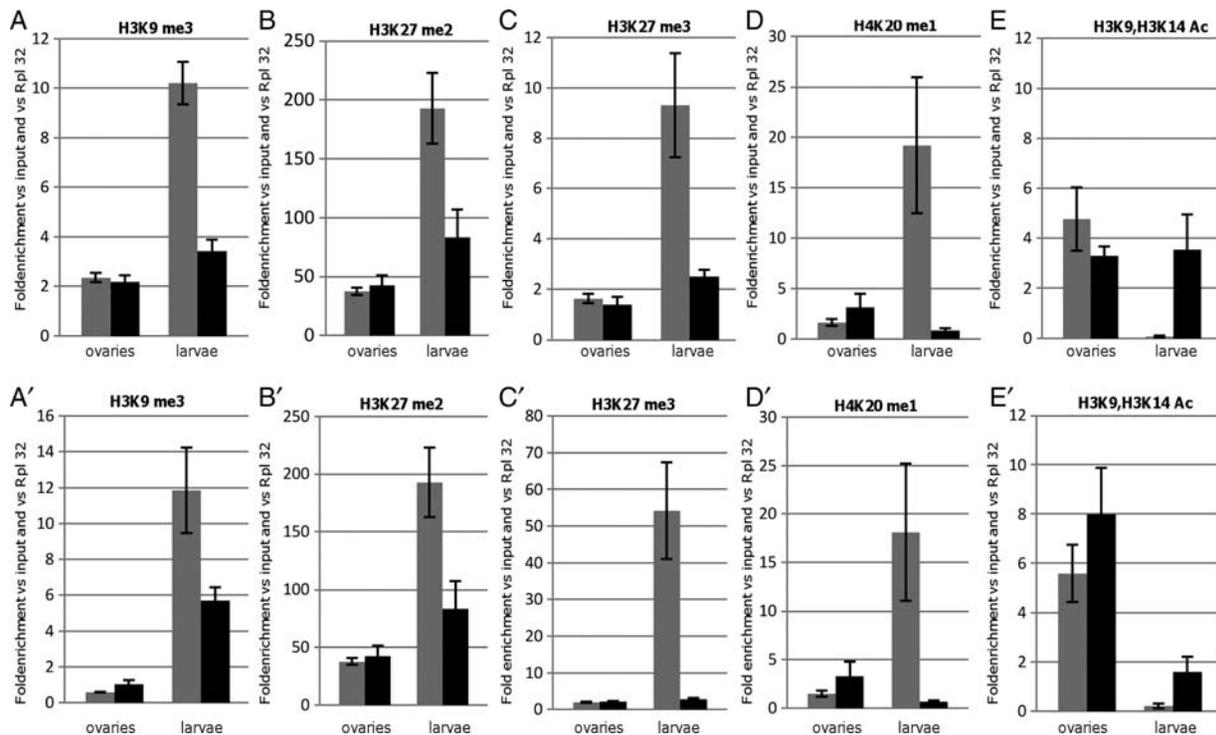
hardly detected by 48 or 72 h as opposed to the level of GFP expressed from the CF line 48 or 72 h after heat shock (Fig. 2D, left and right panels, respectively).

Overall, experiments performed on larval tissues indicate that in these somatic tissues, the silencing persists even after the *Idefix* sequence has been excised what contrasts with the silencing exerted in follicle cells.

### 3.2. Histone methylation marks cover silenced transgenes in somatic tissues outside ovaries

We asked whether specific chromatin structures could mark tGgldas within and/or outside of the ovaries. In a first set of experiments, post-translational histone modifications including lysine acetylation and methylation were analyzed. Whereas acetylation is generally linked with gene activity, methylation of histone H3 on Lys9 (H3K9) and Lys27 (H3K27) correlates with gene silencing. We performed ChIP assays using antibodies against H3K9/K14 acetylation, H3K9me3, H3K27me2, H3K27me3, and H4K20me1, followed by a real-time PCR performed with a primer set targeting *gfp* from both tGgldas and tGg<sup>Did</sup> transgenes (Fig. 1A).

In the ovaries, the H3K9me3, H3K27me2, or H3K27me3 marks were found absent from tGgldas



**Figure 3.** Histone post-translational modifications associated with the silenced sensor transgene in ovaries and larvae: the histograms show relative amounts of H3K9me3 (A and A'), H3K27me2 (B and B'), H3K27me3 (C and C'), H4K20me1 (D and D'), and H3K9/H3K14Ac (E and E'). Each panel displays the results obtained from ovaries (left) and larvae (right) of transgenic lines tGgldas (gray bar) and tGg<sup>Did</sup> (black bar). ChIP was performed on the *gfp* reporter gene (A–D) and upstream of its TTS (A'–D') (see primer sets in Materials and methods). Measurements were normalized to the *Rpl32* gene. Error bars, SEM.

and tGg<sup>Dld</sup> transgenes (Fig. 3A–C). As expected, when the repressive marks are absent, both of these transgenes displayed an enrichment of H3K9 acetylation (Fig. 3E).

By contrast, a very strong enrichment of the methylated marks was detected on tGgldas in larval tissues (Fig. 3A–C). These repressive chromatin hallmarks depend on the presence of *Idefix* because they are enriched >3–10-fold when the *Idefix* sequence is present compared with tGg<sup>Dld</sup>. They also depend on tGgldas transcription since they are found absent when it is not activated by the Gal4 driver (data not shown). In contrast, histone marks such as H3K9 acetylation were found completely absent from tGgldas compared with tGg<sup>Dld</sup> (Fig. 3E).

We further verified that the H4K20me1 mark, generally associated with transcriptional repression and chromosomal condensation, is also enriched on tGgldas when compared with tGg<sup>Dld</sup> in larval tissues, whereas it is absent in ovaries (Fig. 3D). Finally, we also analyzed the H3K9me2 modification and found no enrichment on tGgldas compared with tGgDld (Supplementary Fig. S1).

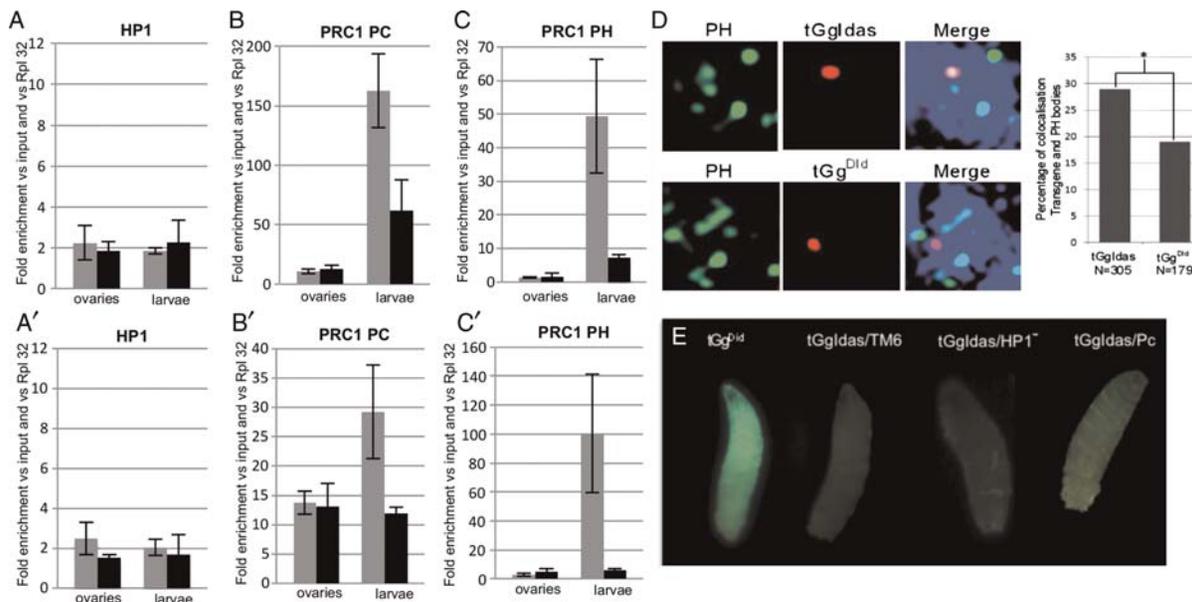
Given the evidence that heterochromatin structures can spread along the chromatin fiber, we wondered whether histone marks associated with the sensor transgenes in the larval tissues could be detected

outside of the transcription start site (TSS). We performed ChIP experiments using primers spanning a region from –440 to –330 bp upstream of the TSS (Fig. 1A). We observed similar results as those obtained with primer targeting *gfp*. On one hand, there is an enrichment of H3K9me3, H3K27me2, H3K27me3, or H4K20me1 on tGgldas in the larval tissues. On the other hand, these repressive marks were found absent from the sensor transgene in the ovaries, whereas an H3K9 acetylation similar to tGg<sup>Dld</sup> was observed (Fig. 4A'–E').

We conclude that distinct chromatin structures cover tGgldas in larval versus ovarian tissues. Post-translational histone modifications generally linked to repression are only observed in non ovarian somatic tissues.

### 3.3. PC binding, and not HP1, correlates with histone methylation at silenced transgenes

HP1 has been implicated in TE silencing in ovaries.<sup>17</sup> Thus, we were interested to test whether HP1 could be detected on the silenced transgene either in ovaries or larvae. ChIP experiments revealed that tGgldas could not be immunoprecipitated with an anti-HP1 antibody at levels above background (Fig. 4A and A'). ChIP experiments performed on a region targeted by HP1 (the



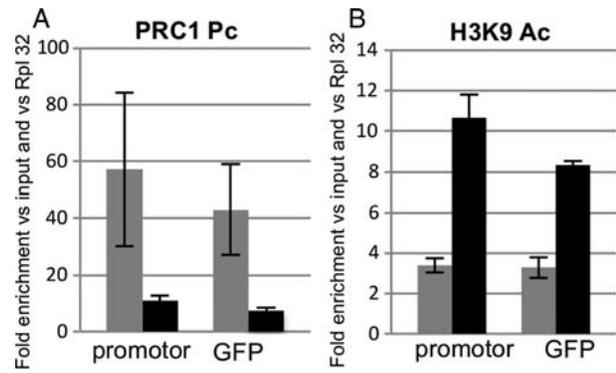
**Figure 4.** PC and PH, not HP1, are required to silence the sensor transgene in larvae: the histograms show relative amounts of HP1 in A and A', the PC protein in B and B', and PH protein in C and C'. Each panel displays the results obtained from ovaries (left) and larvae (right) of transgenic lines tGgldas (gray bar) and tGg<sup>Dld</sup> (black bar). ChIP was performed on the *gfp* reporter gene (A–C) and upstream of its TSS (A'–C'). Measurements were normalized to the *RpL32* gene. Error bars, SEM. (D) The silenced tGgldas transgenes colocalize with Pc-G bodies: characteristic individual nuclei analyzed by FISH-I in larval imaginal discs: Hoechst staining, the Pc-G bodies, the transgene insertion site and the merge of the three channels are shown. The scale bar represents 1  $\mu$ m. On the right, quantification of the percentage of nuclei in which the transgene, either tGgldas or tGg<sup>Dld</sup>, is present in a Pc-G body. Genotype and the total number of nuclei analyzed are indicated below each bar. Asterisk denotes that the colocalization with PH was statistically different between tGg<sup>Dld</sup> and tGgldas ( $P < 0.02$ ). (E) GFP expression from tGgldas in an heterozygous *Su(var)205<sup>5</sup>* (noted HP1<sup>-</sup>) and *pc<sup>1</sup>* mutant background (last two larvae on the right) compared with a non-silenced transgene tGg<sup>Dld</sup> when *Idefix* is excised from tGgldas (first larvae on the left) and to a strong silenced transgene tGgldas in a Pc and HP1 wild-type background (second larvae from the left).

heterochromatic *flam/COM* locus) was used as a positive control for the ChIP experiment (Supplementary Fig. S1). These results indicate that HP1 does not cover tGgldas in somatic cells.

Evidence of H3K27 methylation in larval tissues suggested that Polycomb group (Pc-G) proteins might be implicated in the transgene silencing. Indeed, the methylation of H3K27 is a hallmark of the silencing mediated by PRC2, since it is carried out by the Histone Methyl Transferase Enhancer of Zeste, E(z), the catalytic subunit of PRC2. The chromo-domain of the *Drosophila* PC protein from the PC repressive complex 1 (PRC1) binds to the H3 tail peptide trimethylated at K9 and more strongly at K27me2/3. This prompted us to investigate Pc-G proteins occupancy. ChIP assays were performed with antibodies against PC and Polyhomeotic (PH). Again, two distinct patterns of enrichment were observed. A very strong enrichment of both these proteins was detected on the *gfp* gene in larval tissues, whereas neither PC nor PH proteins could be detected on the *gfp* gene in the ovarian follicle cells (Fig. 4B and C).

Given the evidence that heterochromatin structures can spread along the chromatin fiber, ChIP assays were performed using primers taken upstream of the TSS of tGgldas (Fig. 1). Results indicated that an enrichment of PC and PH was also detected in larval tissues outside of the TSS (Fig. 4B' and C'). A comparably low amount of PC and PH marks was detected on tGgldas and tGg<sup>Dld</sup> in the follicle cells (Fig. 4B' and C').

To further verify that the binding of Pc-G protein is indeed a hallmark of the chromatin structure associated with the repressed transgene in the somatic tissues outside of the ovaries, we performed three additional tests. First, we investigated whether a colocalization of Pc-G bodies and the targeted transgene might be visualized in larval imaginal discs. A combination of 3-D FISH and immunostaining technique (FISH-I) was used, allowing the detection of the relative localization of the transgene with Pc-G bodies visualized using an anti-PH antibody. The tGgldas transgene used in this study is integrated on chromosome 3R at the cytological position 25591560, 937 bp upstream of the *Kay* gene. This site is not a Pc-G binding site reported by Negre *et al.*<sup>18</sup> We found that the colocalizations of tGg<sup>Dld</sup> on one side, and tGgldas on the other, with Pc-G bodies are statistically different, 19 versus 29%, respectively ( $P < 0.02$ ; Fig. 4D). Thus, insertion of tGgldas results in a higher incidence of this locus to colocalize with Pc-G bodies than the same transgene without the *Idefix* sequence. Second, we performed similar ChIP assays on tissues extracted from adult carcasses and verified that indeed the same enrichment in Pc-G proteins is observed in these somatic tissues (Fig. 5A) compared with a low H3K9 acetylation (Fig. 5B). Third, we



**Figure 5.** H3K9 acetylation and PC association with the silenced sensor transgene in adults: the histograms show relative amounts of PC (A) and H3K9ac (B) from carcasses of the transgenic lines tGgldas (gray bar) and tGgDld (black bar). Each panel displays the results obtained when ChIPs were performed with primer sets taken at the GFP reporter gene (right) and upstream of the TSS (left) (see primer sets in Materials and methods). Measurements were normalized to the *Rpl32* gene. Error bars, SEM.

verified that H2A ubiquitination of lysine 119, which is performed by the PRC1 E3 ubiquitin ligase Sce/RING, is also a histone mark linked to PC silencing of tGgldas<sup>19,20</sup> (Supplementary Fig. S1).

Finally, to definitively implicate Pc-G protein in the silencing, we aim at analyzing the effect of PC mutant on the silencing exerted on tGgldas. However, described mutants of PC are recessive lethals. Moreover, a clonal analysis was not possible due to the FRT sequences present in the transgene. Nevertheless, since *Pc* alleles are found regularly but with low expressivity as dominant phenotypes, transgenic flies tGgldas were crossed to *pc*<sup>1</sup>/TM6 mutant backgrounds. GFP expression was examined in the larval progeny [tGgldas/+; *pc*<sup>1</sup>/+] and compared with [tGg<sup>Dld</sup>/+; +/+] larvae as a control for GFP expression when the *Idefix* sequences has been flipped out and to [tGgldas/+; +/+] as a positive control for GFP silencing. As exemplified Fig. 4E, a faint but significant expression of GFP was recovered in the *Pc* heterozygous mutant background (Fig. 4E, right). Finally, to verify that HP1 is not implicated in the silencing of tGgldas, we examined GFP expression in larvae heterozygous for an HP1 mutant allele, [tGgldas/+; Su(var)205<sup>5</sup>/+]. No GFP expression was recovered in these larvae (Fig. 4E).

Overall, the data indicate that a repressive structure implicating Pc-G protein is associated with the sensor transgene tGgldas in somatic tissues outside of the ovaries.

### 3.4. Which signal is required in the soma to target Pc-G proteins to the sensor transgene?

We wondered whether the silencing observed in the somatic cells has the capacity to discriminate one

strand of the sensor transgene from the other. To this end, we established new transgenic lines carrying a transgene, tGglds, similar to tGgldas except that the same *Idefix* fragment has been inserted in a sense orientation. Three independent transgenic lines were established. When the expression of tGglds and tGgldas was examined in the follicle cells where the Piwi pathway has been clearly implicated, we found that both of them were repressed in this tissue (Fig. 6A). However, we found that, unlike tGgldas, tGglds is not targeted to silencing in the soma where a full GFP expression is observed (Fig. 6A) and it is not enriched in repressive chromatin hallmarks such as H3K27me3 as illustrated in Supplementary Fig. S1.

This result indicated that the silencing mechanism involved in the soma is likely to be an RNA silencing pathway able to discriminate between the two mRNA strands produced by the transgenes. Since the silencing of TEs is presumed to be directed by the endo-siRNA pathway in the soma, we investigated whether proteins involved in the endo-siRNA pathway were necessary actors for the silencing exerted on tGgldas. Therefore, we compared GFP expression in homozygous mutant and wild-type flies for *dcr2*, *ago2*, *loq*, and *R2D2*. We found no release of GFP silencing in either of these mutant backgrounds (Fig. 6B).

#### 4. Discussion

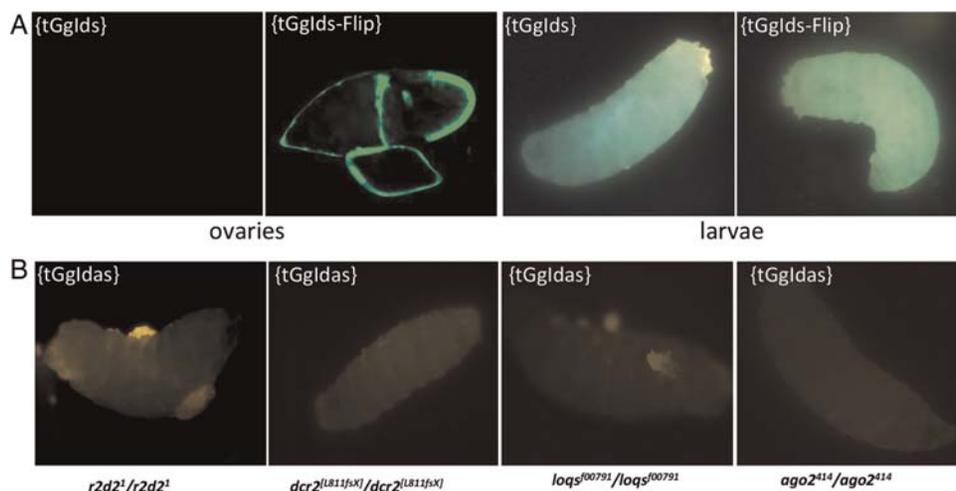
In this study, we tested whether the silencing targeting a TE sequence in somatic lineages constantly

needs the presence of the target or alternatively, once established, the silencing is maintained even though the targeted sequence is lost. We further correlated differences observed in the ovarian follicle cells versus other somatic tissues with differences in the chromatin structures established on the target.

We found that, *in vivo*, the *Idefix* sequence is continuously required for its ongoing repression in the follicle cells. Its repression is not heritably maintained through cell divisions. Accordingly, neither the hallmarks of histone post-translational modifications that generally correlate with compact chromatin structure, nor HP1 or Pc-G proteins are linked to the silenced sensor transgenes. This result argues against the current model in which piRNA–PIWI protein complexes silence transposons in ovarian tissues by directing assembly of heterochromatin-like domains.<sup>17</sup> Interestingly, using ovarian somatic stem cells transfected with double-strand RNAs corresponding to Piwi, Haase *et al.*<sup>21</sup> reported recently that the integrity of the piRNA pathway is essential for the ongoing repression of mobile elements. Consistent with our results, they further provided evidence that, once set by the action of Piwi proteins on chromatin, a transposon silencing cannot be autonomously maintained.

Considered together, these studies promote a model in which, in the follicle cells, TEs are submitted to a post-transcriptional gene silencing (PTGS) which is not coupled to an epigenetic silencing state implicating an HP1-dependent heterochromatin structure.

By contrast, the assembly of a compact chromatin structure marked by the methylation of H3K9 and



**Figure 6.** Characteristics of the silencing targeting tGglds (A) and tGgldas (B) transgenes. (A) The silencing targeting the sensor transgene tGglds is active in the ovaries and inactive in the larvae (first and third panels from the left, respectively). Expression of tGglds is driven by the ubiquitous actin-Gal4 driver. Expression of the transgene after flp-recombinase action (*tGglds-Flip*) is presented as a positive expression of the GFP reporter gene in the second and last panels. (B) Mutations of genes involved in the endo-siRNA pathway do not release the silencing exerted on the tGgldas sensor transgene. GFP expression is never recovered in larvae from homozygous mutants for *r2d2*, *dcr2*, *loq*, and *ago2*.

H3K27 targets TEs in the somatic cells outside of the ovaries. Furthermore, our study identifies Pc-G proteins as major actors. The two PRCs, PRC1 and PRC2, are involved in this structure seeing that H3K27me3 mediated by the histone methyl transferase E(z), belonging to PRC2, is detected as well as PC and PH proteins which belong to PRC1. We further demonstrate that this compact chromatin structure has two main consequences on the targeted locus. First, it can spread along the chromatin fiber and invade the genome upstream of the TSS of the sensor transgene. Second, once established, repression of the sensor transgene is then stably maintained throughout cell divisions and does not need the TE sequence to constantly direct the silencing.

Exploiting the mechanism of PC-mediated silencing to repress endogenous retroelements has also been described by several studies performed on mice.<sup>22,23</sup> Furthermore, in a recent study, the E(z) homolog, EZH, in the green algae *Chlamydomonas reinhardtii*, has been implicated in the repression of both (hetero)euchromatic transgenes and dispersed retrotransposons.<sup>24</sup> Therefore, consistent with the present study, all these data point PRC1 and PRC2 as evolutionarily conserved epigenetic factors employed to silence TE and convey the silencing from one generation of cells to the next.

Interestingly, our study further points out that HP1 is not a partner of the chromatin structure associated with tGldas either in the ovaries or in the soma. Although HP1 has been frequently reported as associated with TE,<sup>17,25</sup> our data are in line with recent data published by the *Drosophila* modENCODE project.<sup>26</sup> This project has generated data sets that profile histone modifications, chromosomal proteins, transcription factors, and nucleosome properties across a developmental time course and in multiple cell lines. Their data show that only TE located in heterochromatin are marked by HP1. By contrast, TE inserted in euchromatic regions are not covered by HP1 as found for tGldas. Thus, one can assume that HP1 might be a signature of TE vestiges because of their heterochromatic localization, whereas HP1-independent silencing pathways might act within euchromatic regions to specifically silence full-length TE and prevent their mobilization.

Overall, if our data implicate PRC1 and PRC2 as epigenetic factors employed to silence TE, they fail to explain how the TE sequence is recognized. In our effort to identify the effector mechanism directing Pc-G proteins on the targeted sensor transgene, we found that the silencing machinery active in the soma has the capacity to discriminate between both orientations of the *Idefix* sequence present within the sensor transgene. Indeed, tGldas is targeted to silencing, whereas tGldis is not. This result is

consistent with a recognition occurring at the RNA, and not the DNA, level.

However, PRC1 and PRC2 employed in this silencing do not have specificity for only one strand and not the other. Therefore, one has to assume that their binding occurs under the presence of a primary signal which brings the sequence specificity. In our study, we were unable to identify this signal. Kanhere *et al.*<sup>27</sup> reported that short RNA from 50 to 200 nts transcribed from the 5' ends of PC target genes play a role in the association of PRC2. Such short RNAs complementary to *Idefix* sequences could potentially explain the targeting. Another likely possibility is that the recognition might occur via homologous siRNAs produced by the cell. Epigenetic complexes made of small RNAs together with PRC1 or PRC2 might potentially explain their recruitment to homologous genomic sequences.<sup>28–30</sup>

Given the evidence that TEs are mainly silenced in the soma via the endo-siRNA pathway, genetic mutants affecting these pathways should have resulted in a loss of the sensor transgene repression. However, no variation in the GFP expression could be detected in these mutant backgrounds. A plausible explanation is that, once established on the transgene, the silent epigenetic state cannot be released even though the targeted fragment is excised or the silencing pathway necessary for the target recognition mutated. Nevertheless, our study pointed out that, like in the follicle cells where piRNAs are involved (Table 1),<sup>9,31</sup> the primary effectors acting in the soma are likely to be single-stranded molecules since only one strand of the mRNA produced by the sensor transgene is targeted. This is comforted by the list of endo-siRNAs reported in the databases (Table 1). Endo-siRNAs able to target and silence tGldas are reported in the databases, whereas none is complementary to the *Idefix* fragment inserted in tGldis.

Overall, in the soma, we propose that the silencing of TE is initiated by a PTGS mechanism which is successively coupled to transcriptional gene silencing. This latter displays all the molecular imprint underlying cell memory and epigenetic inheritance. This epigenetic state, absent in the ovarian follicle cells, accounts for the molecular mechanism controlling maintenance of TE silencing during development.

**Acknowledgements:** We are grateful to F. Bantignies for helpful discussions and PC and PH antibodies. We thank S. Ronsseray for mutant flies. We thank P. George and C. Fritsch for critical review of the manuscript. We thank F. Pellissier for technical assistance and Flyfacility for transgenic lines ([www.Fly-facility.com](http://www.Fly-facility.com)).

**Supplementary data:** Supplementary data are available at [www.dnaresearch.oxfordjournals.org](http://www.dnaresearch.oxfordjournals.org).

## Funding

This work was supported by grants from the Association pour la Recherche contre le Cancer (ARC 1139). J.D. received a graduate grant from the Ministère de l'Enseignement Supérieur et de la Recherche (MESR). E.B. and P.P. received a grant from the Région Auvergne.

## References

- Brennecke, J., Aravin, A.A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., et al. 2007, Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*, *Cell*, **128**, 1089–103.
- Aravin, A., Gaidatzis, D., Pfeffer, S., Lagos-Quintana, M., Landgraf, P., Iovino, N., et al. 2006, A novel class of small RNAs bind to MILI protein in mouse testes, *Nature*, **442**, 203–7.
- Girard, A., Sachidanandam, R., Hannon, G.J. and Carmell, M.A. 2006, A germline-specific class of small RNAs binds mammalian Piwi proteins, *Nature*, **442**, 199–202.
- Ghildiyal, M., Seitz, H., Horwich, M.D., Li, C., Du, T., Lee, S., et al. 2008, Endogenous siRNAs derived from transposons and mRNAs in *Drosophila* somatic cells, *Science*, **320**, 1077–81.
- Chung, W.J., Okamura, K., Martin, R. and Lai, E.C. 2008, Endogenous RNA interference provides a somatic defense against *Drosophila* transposons, *Curr. Biol.*, **18**, 795–802.
- Czech, B., Malone, C.D., Zhou, R., Stark, A., Schlingeheyde, C., Dus, M., et al. 2008, An endogenous small interfering RNA pathway in *Drosophila*, *Nature*, **453**, 798–802.
- Kawamura, Y., Saito, K., Kin, T., Ono, Y., Asai, K., Sunohara, T., et al. 2008, *Drosophila* endogenous small RNAs bind to argonaute 2 in somatic cells, *Nature*, **453**, 793–97.
- Slotkin, R.K. and Martienssen, R. 2007, Transposable elements and the epigenetic regulation of the genome, *Nat. Rev. Genet.*, **8**, 272–85.
- Desset, S., Buchon, N., Meignin, C., Coiffet, M. and Vauray, C. 2008, In *Drosophila melanogaster* the COM locus directs the somatic silencing of two retrotransposons through both Piwi-dependent and -independent pathways, *PLoS One*, **3**, e1526.
- Saito, K., Nishida, K.M., Mori, T., Kawamura, Y., Miyoshi, K., Nagami, T., et al. 2006, Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the *Drosophila* genome, *Genes Dev.*, **20**, 2214–22.
- Xu, K., Bogert, B.A., Li, W., Su, K., Lee, A. and Gao, F.B. 2004, The fragile X-related gene affects the crawling behavior of *Drosophila* larvae by regulating the mRNA level of the DEG/ENaC protein pickpocket1, *Curr. Biol.*, **14**, 1025–34.
- Lee, Y.S., Nakahara, K., Pham, J.W., Kim, K., He, Z., Sontheimer, E.J., et al. 2004, Distinct roles for *Drosophila* dicer-1 and dicer-2 in the siRNA/miRNA silencing pathways, *Cell*, **117**, 69–81.
- Lin, H. and Spradling, A.C. 1997, A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the *Drosophila* ovary, *Development*, **124**, 2463–76.
- Liu, Q., Rand, T.A., Kalidas, S., Du, F., Kim, H.E., Smith, D.P., et al. 2003, R2D2, a bridge between the initiation and effector steps of the *Drosophila* RNAi pathway, *Science*, **301**, 1921–25.
- Forstemann, K., Tomari, Y., Du, T., Vagin, V.V., Denli, A.M., Bratu, D.P., et al. 2005, Normal microRNA maturation and germ-line stem cell maintenance requires loquacious, a double-stranded RNA-binding domain protein, *PLoS Biol.*, **3**, e236.
- Brasset, E., Bantignies, F., Court, F., Cheresiz, S., Conte, C. and Vauray, C. 2007, Idefix insulator activity can be modulated by nearby regulatory elements, *Nucleic Acids Res.*, **35**, 2661–70.
- Klenov, M.S., Lavrov, S.A., Stolyarenko, A.D., Ryazansky, S.S., Aravin, A.A., Tuschl, T., et al. 2007, Repeat-associated siRNAs cause chromatin silencing of retrotransposons in the *Drosophila melanogaster* germline, *Nucleic Acids Res.*, **35**, 5430–38.
- Negre, N., Hennetin, J., Sun, L.V., Lavrov, S., Bellis, M., White, K.P., et al. 2006, Chromosomal distribution of PcG proteins during *Drosophila* development, *PLoS Biol.*, **4**, e170.
- Fritsch, C., Beuchle, D. and Muller, J. 2003, Molecular and genetic analysis of the Polycomb group gene *sex combs extra/ring* in *Drosophila*, *Mech. Dev.*, **120**, 949–54.
- Wang, H., Wang, L., Erdjument-Bromage, H., Vidal, M., Tempst, P., Jones, R.S., et al. 2004, Role of histone H2A ubiquitination in Polycomb silencing, *Nature*, **431**, 873–78.
- Haase, A.D., Fenoglio, S., Muerdter, F., Guzzardo, P.M., Czech, B., Pappin, D.J., et al. 2010, Probing the initiation and effector phases of the somatic piRNA pathway in *Drosophila*, *Genes Dev.*, **24**, 2499–2504.
- Golding, M.C., Zhang, L. and Mann, M.R. 2010, Multiple epigenetic modifiers induce aggressive viral extinction in extraembryonic endoderm stem cells, *Cell Stem Cell*, **6**, 457–67.
- Leeb, M., Pasini, D., Novatchkova, M., Jaritz, M., Helin, K. and Wutz, A. 2010, Polycomb complexes act redundantly to repress genomic repeats and genes, *Genes Dev.*, **24**, 265–76.
- Shaver, S., Casas-Mollano, J.A., Cerny, R.L. and Cerutti, H. 2010, Origin of the polycomb repressive complex 2 and gene silencing by an E(z) homolog in the unicellular alga *Chlamydomonas*, *Epigenetics*, **5**, 301–12.
- Haynes, K.A., Caudy, A.A., Collins, L. and Elgin, S.C. 2006, Element 1360 and RNAi components contribute to HP1-dependent silencing of a pericentric reporter, *Curr. Biol.*, **16**, 2222–27.

26. Riddle, N.C., Minoda, A., Kharchenko, P.V., Alekseyenko, A.A., Schwartz, Y.B., Tolstorukov, M.Y., et al. 2011, Plasticity in patterns of histone modifications and chromosomal proteins in *Drosophila heterochromatin*, *Genome Res.*, **21**, 147–63.
27. Kanhere, A., Viiri, K., Araujo, C.C., Rasaiyaah, J., Bouwman, R.D., Whyte, W.A., et al. 2010, Short RNAs are transcribed from repressed polycomb target genes and interact with polycomb repressive complex-2, *Mol. Cell*, **38**, 675–88.
28. Grimaud, C., Negre, N. and Cavalli, G. 2006, From genetics to epigenetics: the tale of Polycomb group and trithorax group genes, *Chromosome Res.*, **14**, 363–75.
29. Zhang, H., Christoforou, A., Aravind, L., Emmons, S.W., van den Heuvel, S. and Haber, D.A. 2004, The *C. elegans* Polycomb gene SOP-2 encodes an RNA binding protein, *Mol. Cell*, **14**, 841–47.
30. Zhao, J., Sun, B.K., Erwin, J.A., Song, J.J. and Lee, J.T. 2008, Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome, *Science*, **322**, 750–56.
31. Brennecke, J., Malone, C.D., Aravin, A.A., Sachidanandam, R., Stark, A. and Hannon, G.J. 2008, An epigenetic role for maternally inherited piRNAs in transposon silencing, *Science*, **322**, 1387–92.