Production of Small Noncoding RNAs from the *flamenco* Locus Is Regulated by the *gypsy* Retrotransposon of *Drosophila melanogaster*

Vincenzo Guida,*,[1,2] Filippo M. Cernilogar,†,[1] Angela Filograna,*,[3] Roberto De Gregorio,‡ Hirotsugu Ishizu,§ Mikiko C. Siomi,§ Gunnar Schotta,† Gian Carlo Bellenchi,‡ and Davide Andrenacci*,,††,[4]**

*Institute of Biosciences and Bioresources, and ‡Institute of Genetics and Biophysics "Adriano Buzzati Traverso," Consiglio Nazionale delle Ricerche, 80131 Naples, Italy, †Division of Molecular Biology, Biomedical Center and Center for Integrated Protein Science Munich, Ludwig-Maximilians-University, 82152 Planegg-Martinsried, Germany, §Department of Biological Sciences, Graduate School of Science, The University of Tokyo, 113-0032, Japan, **Institute of Molecular Genetics, Consiglio Nazionale delle Ricerche, 40136 Bologna, Italy, and ††SC Laboratory of Musculoskeletal Cell Biology, Rizzoli Orthopedic Institute, 40136 Bologna, Italy

ORCID ID: 0000-0002-4084 (D.A.)

ABSTRACT Protective mechanisms based on RNA silencing directed against the propagation of transposable elements are highly conserved in eukaryotes. The control of transposable elements is mediated by small noncoding RNAs, which derive from transposon-rich heterochromatic regions that function as small RNA-generating loci. These clusters are transcribed and the precursor transcripts are processed to generate Piwi-interacting RNAs (piRNAs) and endogenous small interfering RNAs (endo-siRNAs), which silence transposable elements in gonads and somatic tissues. The *flamenco* locus is a *Drosophila melanogaster* small RNA cluster that controls *gypsy* and other transposable elements, and has played an important role in understanding how small noncoding RNAs repress transposable elements. In this study, we describe a cosuppression mechanism triggered by new euchromatic *gypsy* insertions in genetic backgrounds carrying *flamenco* alleles defective in *gypsy* suppression. We found that the silencing of *gypsy* is accompanied by the silencing of other transposons regulated by *flamenco*, and of specific *flamenco* sequences from which small RNAs against *gypsy* originate. This cosuppression mechanism seems to depend on a post-transcriptional regulation that involves both endo-siRNA and piRNA pathways and is associated with the occurrence of developmental defects. In conclusion, we propose that new *gypsy* euchromatic insertions trigger a post-transcriptional silencing of *gypsy* sense and antisense sequences, which modifies the *flamenco* activity. This cosuppression mechanism interferes with some developmental processes, presumably by influencing the expression of specific genes.

KEYWORDS transposon; small RNA; RNA silencing; primary transcript; ecdysis

EUKARYOTIC genomes consist in part of sequences derived from a wide variety of transposable elements (TEs), some of which can mobilize to new genomic locations (de Koning *et al.* 2011). A genomic consequence of their mobilization is the induction of new mutations and chromosomal rearrangements that may have deleterious effects on fitness. However, they may also provide a fundamental contribution to genetic variation and evolutionary changes (Fedoroff 2012; Warren *et al.* 2015; Elbarbary *et al.* 2016; Mita and Boeke 2016).

TE activation is suppressed by specific silencing mechanisms that act both at the transcriptional level, through chromatin modifications, and at the post-transcriptional level (Buchon and Vaury 2006). Piwi-interacting RNAs (piRNAs), a distinct class of 24- to 30-nt-long RNAs produced by a Dicer-independent biogenesis pathway, are involved in the recognition and selective silencing of transposons during gametogenesis (Sarot *et al.* 2004; Kalmykova *et al.* 2005). In the *Drosophila* ovary germline, the coordinated action of *aubergine* (*aub*), *Argonaute 3*

doi: 10.1534/genetics.116.187922
Manuscript received February 4, 2016; accepted for publication August 18, 2016; published Early Online August 24, 2016.
Supplemental material is available online at www.genetics.org/lookup/suppl/doi:10.1534/genetics.116.187922/-/DC1.
[1]These authors contributed equally to this work.
[2]Present address: Department of Experimental Medicine, Section of Hygiene, Occupational Medicine and Forensic Medicine, School of Medicine, Second University of Naples, 80138 Naples, Italy.
[3]Present address: Institute of Protein Biochemistry, Consiglio Nazionale delle Ricerche, 80131 Naples, Italy.
[4]Corresponding author: Institute of Molecular Genetics, Consiglio Nazionale delle Ricerche, Via di Barbiano 1/10, 40136 Bologna, Italy. E-mail: dandrena@area.bo.cnr.it

(*AGO3*), and *piwi* suppresses activity of a broad group of TEs through the formation of piRNAs involving both a primary processing and a secondary "ping-pong" amplification loop (Brennecke *et al.* 2007; Gunawardane *et al.* 2007; Li *et al.* 2009; Malone *et al.* 2009), while in the follicular epithelium a different mechanism exists with piRNAs generated solely from primary transcript derived mainly from the *flamenco* locus. Somatic activation of TEs is mainly controlled by endogenous small interfering RNAs (endo-siRNAs), generated by *Dicer-2* (*Dcr-2*) and *Argonaute 2* (*AGO2*) genes (Czech *et al.* 2008; Ghildiyal *et al.* 2008). Endo-siRNAs can arise from genomic loci from which piRNAs also originate.

The *flamenco* locus, also known as the COM locus, is a major regulator of the *gypsy* retrotransposon and other TEs, such as *ZAM* and *Idefix* (Pélisson *et al.* 1994; Desset *et al.* 1999; Desset *et al.* 2003; Goriaux *et al.* 2014b). *gypsy* is repressed in the follicular epithelium of *flamenco* restrictive females, while it is active and able to propagate within the germline in permissive backgrounds (Pélisson *et al.* 1994; Bucheton 1995; Song *et al.* 1997; Chalvet *et al.* 1999). It is reported that natural populations of *Drosophila melanogaster* exhibit a restrictive/permissive polymorphism for *flamenco* (Pélisson *et al.* 1997). The *flamenco* locus is a uni-strand cluster comprising a region of ~180 kb, located in the pericentromeric heterochromatin of the *X* chromosome at 20A1–A3. The sequence of *flamenco* mostly consists of a high density of antisense transposon fragments, the majority of which belong to the LTR group of retrotransposons (Brennecke *et al.* 2007; Mevel-Ninio *et al.* 2007). The *flamenco* locus can be considered a small RNA cluster since in the ovary it acts as a source of piRNAs, while in somatic tissues outside of the ovary it acts as a source of endo-siRNAs (Brennecke *et al.* 2007; Ghildiyal *et al.* 2008). The role of *flamenco* is to prevent the expression of *gypsy* and other TEs having complementary sequences in this locus, and for this reason it is considered part of the "adaptive immunity" for transposon control (Bergman *et al.* 2006). The ability to control TE transposition is variable and seems to be associated with the high structural dynamics of the *flamenco* locus, which depends on chromosomal rearrangements and loss or gain of TEs inside the locus (Zanni *et al.* 2013).

Important information regarding the mechanism by which small noncoding RNAs repress TEs derive from the study of the *flamenco* locus. However, still little is known about *flamenco* regulation. It is known that unidirectional transcription of *flamenco*, which is transcribed from an RNA polymerase II promoter located ~2 kb upstream of the *DIP1* gene and activated by the Cubitus interruptus transcription factor, produces a long precursor transcript (Goriaux *et al.* 2014a). The *flamenco* precursor transcript is then subjected to alternative splicing to produce RNA precursors that are then processed to piRNAs. It is also known that small RNA clusters like *flamenco* are subjected to heterochromatin silencing and there is evidence that their expression can be modified by mutations affecting heterochromatin formation (Moshkovich and Lei 2010).

In this study, we analyzed the effect of some euchromatic *gypsy* mutations in *Drosophila* strains carrying *flamenco* alleles permissive for *gypsy* mobilization. Besides the known phenotypes due to the *gypsy* insertion in the specific locus, we found additional phenotypes depending on a genetic interaction between the specific euchromatic *gypsy* insertion and the *flamenco* permissive allele. To characterize this genetic interaction, the experiments were mainly performed using *gypsy*-induced mutations directly isolated from a *flamenco* permissive strain, using the original strain as a control. In this way, we detected variations in phenotypes and in expression levels in flies with a common genetic background. This minimized the possibility of detecting background effects. We found that these phenotypes are accompanied by cosuppression of both *gypsy* retrotransposons and *gypsy* antisense fragments that are part of the *flamenco* primary transcript. The cosuppression mechanism seems to be triggered by the *gypsy* euchromatic insertion site, which shows specific characteristics of new small RNA clusters. Cosuppression of *gypsy* and *flamenco* depends on an RNA silencing mechanism which is not accompanied by an increase of the heterochromatic marks H3K9me3 and H3K27me3 at the *flamenco* locus, at *gypsy* sequences, as well as at the other TEs analyzed. These data suggest that the described cosuppression mechanism occurs at post-transcriptional level.

Materials and Methods

Drosophila stocks

Drosophila stocks were maintained at 25° on standard cornmeal/yeast medium. $y\ w^{67c23}$ and Canton-S have been used as wild-type strains. $AstC\text{-}R2^{f03116}$, ct^{6}, ct^{n}, $Su(Hw)^{8}$, Bx^{2}, $flam^{GB02658}$, $AGO3^{t2}$, $Su(var)3\text{-}7^{DG10405}$, $Su(var)3\text{-}9^{1}$, $Su(var)3\text{-}9^{2}$, $Su(var)205^{5}$, f^{1}, and *Df(1)l11/FM7c*, $flam^{FM7c}$; *P{gypsy-lacZ.p12}* were obtained from Bloomington Stock Center. $flam^{1}$ and $flam^{su(f)}$ were kindly provided by Alain Pélisson. Since $AstC\text{-}R2^{f03116}$ flies were positive for *Wolbachia*, they were treated for two generations with high-dose tetracycline (0.25 mg/ml) to remove the pathogen. To eliminate viruses and other horizontally transmitted pathogens, $AstC\text{-}R2^{f03116}$ eggs were dechlorinated in 50% bleach for 2 min. Flies carrying the *gypsy-lacZ* transgene were obtained by crossing *Df(1)l11/FM7c*, $flam^{FM7c}$; *P{gypsy-lacZ.p12}* females with $w\ flam^{A}$, $w\ ct^{A}\ flam^{A}$, or $flam^{su(f)}$ males. $w\ flam^{A}/flam^{1}$; *P{gypsy-lacZ.p12}*/+ flies were obtained by crossing $w\ flam^{A}$; *P{gypsy-lacZ.p12}* females with $flam^{1}$ males. $ct6\ flam^{1}$, $ct^{n}\ flam^{1}$, and $Bx^{2}\ flam^{1}$ flies were obtained by recombination, by selecting *ct* and *Bx* phenotypes and the eye phenotype induced by the mal^{1} mutation which maps very close to the $flam^{1}$ mutation. Different independent lines were analyzed to confirm the phenotype. The $ct^{A}\ flam^{BG}$ line was obtained by recombination, by selecting the *ct* phenotype and variegated eye color due to the presence of the *mini-white* gene carried by the *P* element inserted in the *flamenco* locus. Genetic combinations between $w\ flam^{A}$ or $w\ ct^{A}\ flam^{A}$ and $Su(Hw)^{8}$, $Su(var)3\text{-}7^{DG10405}$, $Su(var)3\text{-}9^{1}$, $Su(var)3\text{-}9^{2}$,

AGO2^{414}, and *AGO3*^{t2} were obtained by crossing *w flam*^{A} *or w ct*^{A} *flam*^{A} females carrying the *TM6b/TM3 Sb* balancer combination to males carrying the specific balanced mutations. Genetic combinations between *w flam*^{A} or *w ct*^{A} *flam*^{A} and *Su(var)205*^{5}, *Dcr-2*^{L811fsX}, and *aub*^{HN2} were obtained by crossing *w ct*^{A} *flam*^{A}; *CyO/Sp* females with males carrying the specific balanced mutations. All experiments with *flamenco* permissive alleles, except for *flam*^{GB}, were performed using females homozygous for the *X* chromosome carrying the *flamenco* allele.

Mortality determination

For calculation of mortality during adult ecdysis, parent flies were allowed to lay eggs in standard vials for 3–4 days. Eclosed adults and adults that died within the puparium were then counted, and the percentage of mortality was calculated. Three separate experiments were carried out counting at least 100 individuals per experiment, and the average and SD were calculated. *P* value was calculated using the two-tailed unpaired Student's *t*-test (** $P < 0.01$).

β-Galactosidase staining

Ovaries were dissected in PBS, transferred and fixed in PBS plus Triton X-100 (PBT) (PBS with 0.1% Triton X-100) containing 0.1% glutaraldehyde for 5 min, and rinsed three times with PBT. Each ovary sample was then incubated with 0.2% X-gal in staining solution {10 mM phosphate buffer, pH 7.2, 1 mM $MgCl_2$, 5 mM $K_4[FeII(CN)_6]$, 5 mM $K_3[Fe(III)(CN)_6]$, 0.1% Triton X-100} for 45 min. After staining, ovary samples were rinsed three times with PBT, mounted in PBS containing 50% glycerol, and examined by bright-field microscopy.

Microscopic analysis

Living pupae, pharate adults extracted from cocoons, and adult eyes were analyzed and photographed by bright-field stereoscopic microscopy. Adult wings were dissected and dehydrated in ethanol, mounted in a solution of lactic acid: ethanol (6:5), and examined by bright-field microscopy.

Quantitative RT- PCR

Total RNA was extracted by crushing 15–20 pupae (48 and 96 hr), 20 ovaries (3 days old), and 50–100 heads (0- to 24-hr-old females) in TRI Reagent (Sigma-Aldrich) and subsequently treated with TURBO DNase (Ambion). Complementary DNAs (cDNAs) were prepared using M-MLV Reverse Transcriptase (Ambion) and random primers, according to the manufacture's protocol. To analyze strand-specific expression, cDNAs were produced using strand-specific primers. Real-time PCRs were performed in Fast 96-well Reaction Plates (Applied Biosystems, Foster City, CA) using an Applied Biosystems 7900 HT Fast Real-Time PCR System according to the manufacturer's suggested procedure. After the PCR run, dissociation curve analysis was performed to determine target specificity. Three replicates were used for each experimental sample. Messenger RNA (mRNA) levels were normalized to the internal standard gene *Rp49*. To measure the fold change of expression levels between control and experiment(s), the ΔΔCt method was used. Three biological replicates were used and the average and SD were calculated. Differences between experiments and controls were tested with the one-sample Student's *t*-test (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.005$). For quantitative PCR (qPCR) primers see Supplemental Material, Table S1.

Chromatin immunoprecipitation

Chromatin preparation and immunoprecipitation was carried out as previously described (Schauer *et al.* 2013), with some modifications. Heads of frozen flies were separated using 630 and 400 μm sieves. Then, 800–1000 frozen fly heads were ground to a fine powder with a mortar and pestle (chilled on dry ice), resuspended in 25 ml ice-cold homogenization buffer [350 mM sucrose, 15 mM HEPES, pH 7.6, 10 mM KCl, 5 mM $MgCl_2$, 0.5 mM EGTA, 0.1 mM EDTA, 0.1% Tween, with 1 mM DTT and Hoffman La Roche (Nutley, NJ) Protease Inhibitor Cocktail added immediately prior to use], incubated for 5 min on ice, and then further ground in a mechanical homogenizer (rotation 2000 rpm, 20 ups and downs, slowly). After centrifugation (5 min, 2000 × *g*, 4°), the homogenate was resuspended in 25 ml room-temperature homogenization buffer, fixed using 1% formaldehyde for 10 min at room temperature, and then quenched for 5 min with 0.125 M glycine. The tissue debris was removed by filtration with a 70-μm nylon cell strainer (BD Falcon) and nuclei were collected by centrifugation (5 min, 2000 × *g*, 4°). Nuclei were resuspended in 1 ml ice-cold radio immunoprecipitation assay (RIPA) buffer [150 mM NaCl, 25 mM HEPES, pH 7.6, 1 mM EDTA, 1% Triton X-100, 0.1% SDS, 0.1% sodium deoxycholate (DOC), with protease inhibitors added prior to use], transferred to a 1.5-ml low-binding tube, and washed twice with 1 ml ice-cold RIPA. Nuclei were collected by centrifugation (1 min, 3500 × *g*). Finally, the nuclei were resuspended in 1 ml Shearing Buffer (150 mM NaCl, 25 mM HEPES, pH 7.6, 1 mM EDTA, 0.3% SDS, 0.1% DOC, with protease inhibitors added prior to use) and transferred to 1 ml Covaris Adaptive Focused Acoustics tubes (catalogue number: 520081). Chromatin was fragmented to an average size of 100–400 bp in a Covaris S220 device (30 min; 4°; peak incident power, 120; duty factor, 20; cycles per burst, 200). After shearing, the debris was collected by centrifugation (10 min, 16,000 × *g*, 4°), the chromatin was transferred to a low-binding tube, and stored at −80°. Fragment size was checked after cross-link reversal on a 1% agarose gel. A 150 μl volume of chromatin (corresponding to ∼150 heads) was used for each chromatin immunoprecipitation (ChIP) assay. For each ChIP assay, 30 μl Dynabeads Protein G (Invitrogen, Carlsbad, CA) was washed with ice-cold PBS 0.1% Tween 20 (PBST) and incubated in 500 μl PBST with the appropriate antibody (6 μl anti-H3K9me3, Active Motif, catalogue number 39161; 6 μg anti-H3K27me3, Diagenode, catalogue number 069-050) for 2 hr on a rotating wheel (10 rpm) at 4°. After washing with PBST, the Protein G beads prebound with or without (negative control) the antibody were resuspended in

850 μl ice-cold RIPA buffer, added to chromatin, and left in rotation (10 rpm) at 4° for 16 hr. After immunopurification, beads were washed (each wash 10 min, 20 rpm at 4°) four times with 1 ml RIPA, once with 1 ml LiCl wash buffer (250 mM LiCl, 10 mM Tris-HCl, pH 8.0, 1 mM EDTA, 0.5% NP-40, 0.5% DOC), and once with 1 ml TE buffer (10 mM Tris-HCl, pH 8, 1 mM EDTA). Beads were resuspended in 100 μl elution buffer (50 mM Tris-HCl, pH 8.0, 10 mM EDTA, 1% SDS) and incubated on a Termomixer at 65° and 800 rpm. Beads were removed by a magnetic rack and the cross-link reversal of immunoprecipitated (IP) DNA was carried out overnight at 65° in a PCR machine. Then, 100 μl TE was added, RNA was degraded by 4 μl RNase A (10 mg/ml) for 1 hr at 37° in a Termomixer (600 rpm), and proteins were digested with 4 μl Proteinase K (10 mg/ml) at 55° for 2 hr in a Termomixer (750 rpm). IP DNA was purified using QIAquick PCR Purification columns (QIAGEN, Valencia, CA) and eluted with 50 μl EB Buffer. A volume of 1 μl was used for qPCR analysis. qPCR was performed with the Fast SYBR Green Master Mix qPCR Kit (Applied Biosystems, Foster City, CA) using 10 μl of total reaction, and analyzed on a Light Cycler 480 Real-Time PCR System (Roche Applied Science). Relative quantifications as a percentage of starting material (percentage of input) were determined using the following equation: % (ChIP/total input) = 2^[(*Ct*(*x*% input) − log(*x*%)/log$_2$) − *Ct*(ChIP)] × 100%. *Ct*(ChIP) and *Ct*(*x*%input) are threshold values obtained from the exponential phase of qPCR for the IP DNA sample and input sample, respectively; (log*x*%/log$_2$) accounts for the dilution 1:*x* of the input to balance the difference in amounts of ChIP and input DNA from qPCR. For qPCR primers see Table S1.

Quantification of piRNA expression

Small RNAs were isolated from ∼20 ovaries dissected from 3-day-old females or ∼100 heads from 0- to 24-hr-old females, using the miRVana small RNA Isolation Kit (Ambion) according to the manufacturer's instructions. RNA was eluted in 100 μl of RNAase-free water and quantified by spectrophotometric analysis. Samples were then diluted to 2 ng/μl. TaqMan Small RNA Assays (Applied Biosystem) was used for the analysis of piRNAs expression. Specific TaqMan primers and probes for piRNA A (5′ TATTGAGCTCACCGAGAAAGGGCTGGC 3′), piRNA B (5′ TGTTGTGAGTGTATCCAGATGGAG 3′), and for the reference 2S ribosomal RNA (Applied Biosystem) were used in quantitative RT-PCR (qRT-PCR). Reverse transcription for piRNA A and B and the reference 2S ribosomal RNA was performed using the TaqMan microRNA Reverse Transcription Kit and the target specific primers provided with the TaqMan small RNA Assay. qPCR was performed using primers and probes provided with the TaqMan small RNA Assay and using TaqMan Universal Master Mix II, no Uracil-N Glycosylase (Applied Biosystems). Real-time PCRs were performed in Fast 96-well Reaction Plates (Applied Biosystems) using an Applied Biosystems 7900 HT Fast Real-Time PCR System according to the manufacturer's suggested procedure. piRNA levels were normalized to the internal standard 2S ribosomal RNA. To measure the fold change of expression levels between control and experiment(s), the ΔΔCt method was used. Unless otherwise reported, three biological replicates were used and the average and SD were calculated. Differences between experiments and controls were tested with two-tailed one-sample Student's *t*-tests (*** $P < 0.005$).

Data availability

The authors state that all data necessary for confirming the conclusions presented in the article are represented fully within the article. *Drosophila* strains are available upon request.

Results

An ecdysis phenotype manifesting during adult eclosion is induced by a gypsy insertion in the cut locus

We explored the identification of a male with a cut-wing phenotype in a strain carrying a *P*-element insertion in the *AstC-R2* gene, which is located in the third chromosome. The cut-wing phenotype was induced by a novel spontaneous mutation affecting the *cut* locus in the *X* chromosome, since it did not complement the ct^6 allele (Figure 1A). The new *ct* allele was named ct^A. Homozygous mutations in the *suppressor of Hairy wing* [*su(Hw)*] gene revert cut-wing phenotypes only when induced by *gypsy* (Harrison *et al.* 1993). The wing phenotype induced by the ct^A mutation was suppressed by $su(Hw)^8$ (Figure 1A), suggesting that the mutation was effectively induced by *gypsy*. Indeed a *gypsy* element was identified in the *cut* regulatory region (Figure S1, A and B and File S1).This result was not surprising because it is known that the *cut* locus is a hotspot for *gypsy* insertions (Jack 1985).

A ct^A homozygous strain in the $AstC\text{-}R2^{f03116}$ background was obtained (Figure S2A). In addition to the *cut*-specific phenotypes (Jack 1985), we found that the *gypsy* mutation induced high levels of mortality during adult eclosion, which was not found in control flies (Figure 1, B and C). In fact, when ct^A/ct^A flies heterozygous for the $AstC\text{-}R2^{f03116}$ mutation were crossed with each other, we found that ∼40% of the $AstC\text{-}R2^{f03116}$ heterozygous flies (*w* ct^A $flam^A$; $AstC\text{-}R2^{f03116}/TM6b$) and >90% of $AstC\text{-}R2^{f03116}$ homozygous flies (*w* ct^A $flam^A$; $AstC\text{-}R2^{f03116}/AstC\text{-}R2^{f03116}$) of the progeny failed to emerge from the puparium (Figure 1C). A significant level of lethal ecdysis deficiency was also observed in flies with only one copy of the ct^A allele (Figure 1C, *w* ct^A $flam^A/w$ $flam^A$; $AstC\text{-}R2^{f03116}/AstC\text{-}R2^{f03116}$ flies), suggesting that one *gypsy* copy was sufficient to induce the phenotype and that this phenotype does not depend on a *cut* loss of function. Regarding the phenotype, we found that pharates exhibited characteristic eclosion movements that are common during adult eclosion, but they appeared trapped in the pupal case (Figure 1D). When the pupal case of affected flies was removed, we found that pharates were trapped in the pupal cuticle (Figure 1E).

To exclude that the $AstC\text{-}R2^{f03116}$ mutation was necessary to determine the ecdysis phenotype, the third chromosome

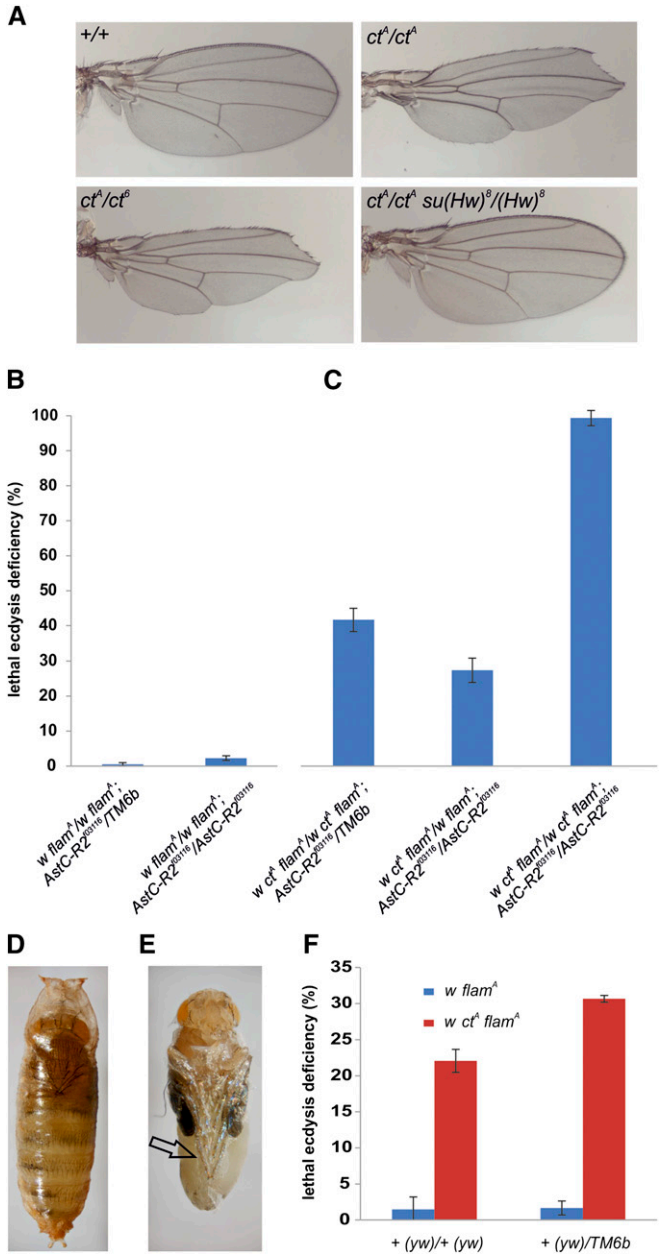

Figure 1 Lethal ecdysis deficiency is induced by a *gypsy* insertion in the *cut* locus. (A) Cuticle preparations of wings from $flam^A$/ $flam^A$ (+/+), ct^A $flam^A$/ct^A $flam^A$ (ct^A/ct^A), ct^A $flam^A$/ct^6 (ct^A/ct^6), and ct^A $flam^A$/ct^A $flam^A$; $su(Hw)^8$/$su(Hw)^8$ (ct^A/ct^A; $su(Hw)^8$/$su(Hw)^8$). (B, C) Mortality during adult eclosion of flies in the *AstC-R2* ^{f03116} background. (B) Mortality during adult eclosion of heterozygous and homozygous flies derived from the mating of *AstC-R2* ^{f03116} heterozygous parents. (C) Mortality during adult eclosion of ct^A heterozygotes and homozygotes, and *AstC-R2* heterozygous and homozygous flies derived from mating of *AstC-R2* heterozygous parents. (D, E) representative pictures of ct^A/ct^A; *AstC-R2* ^{f03116}/*AstC-R2* ^{f03116} pharate adults. (D) Fly trapped inside the puparium ~12 hr after the start of the adult ecdysis, showing the characteristic pigmentation of eclosed adults. (E) Fly extracted from puparium showing pupal membrane attached to legs, wings, and other parts of the body (arrow). (F) Mortality during adult eclosion of flies with the third chromosomes deriving from the *y* w^{67c23} strain (indicated in brackets) in absence or presence of the ct^A mutation and of the *TM6b* balancer chromosome.

carrying the *P*-element insertion was substituted with a wild-type chromosome from the *y* w^{67c23} strain (Figure S2B). The ct^A mutation in this wild-type context showed >20% mortality during adult eclosion (Figure 1F), confirming that the *gypsy* insertion in the *cut* locus *per se* induces lethal ecdysis deficiency. Mortality was, however, lower than in *AstC-R2*^{f03116} homozygous or heterozygous conditions, suggesting a high sensitivity of this phenotype to the genetic background. In fact, the presence of *TM6b* in combination with a wild-type third chromosome was sufficient to increase mortality level (Figure 1F), confirming that lethal ecdysis deficiency induced by the *gypsy* insertion in the *cut* locus is highly sensitive to the genetic background.

Developmental phenotypes are produced in flamenco permissive backgrounds in the presence of gypsy-induced mutations

Since *flamenco* permissive alleles are unable to suppress *gypsy* (Prud'homme *et al.* 1995), the identification of the *gypsy* insertion in the *cut* locus suggested that ct^A flies carry a *flamenco* permissive allele. *flamenco* permissivity can be evaluated by monitoring the activity of transgenes carrying fragments of *gypsy* fused to the *lacZ* reporter (Sarot *et al.* 2004). We observed strong activation of the *gypsy-lacZ* construct when the original *X* chromosome from the *AstC-R2*^{f03116} strain was combined with an *X* chromosome carrying the *Df(1)l11* deficiency (Figure 2, A and B), which encompasses the *flamenco* locus (Prud'homme *et al.* 1995; Pélisson *et al.* 1997). *flamenco* restrictive alleles are dominant and repress the activation of *gypsy* even when in combination with permissive alleles (Pélisson *et al.* 1997). The *X* chromosome carrying the $flam^A$ allele combined with an *X* chromosome carrying the $flam^{FM7c}$ restrictive allele repressed the *gypsy-lacZ* construct (Figure 2C), while combined with the permissive $flam^1$ allele activated the expression of *gypsy-lacZ* (Figure 2D). Based on these results, the *flamenco* permissive allele in the ct^A-carrying chromosome was named $flam^A$.

Since lethal ecdysis deficiency has never been reported to be a phenotype induced by *gypsy* insertions in the *cut* locus (Jack 1985), we hypothesized that these phenotypes in *w* ct^A $flam^A$ flies should depend on a genetic interaction between the *gypsy* insertion and the *flamenco* permissive allele. To test this hypothesis, the ct^A chromosome was recombined with a chromosome carrying the permissive $flam^{BG}$ allele (Mevel-Ninio *et al.* 2007). Mortality during adult eclosion of *w* ct^A $flam^{BG}$ flies was higher than in control *w* $flam^{BG}$ flies (Figure 2F), and very similar to that found in *w* ct^A $flam^A$ flies (Figure 1F). We also tried to confirm these data using two other *ct* alleles and the permissive $flam^1$ allele. To achieve this goal, the ct^6 allele, which is also induced by a *gypsy* insertion (Tchurikov *et al.* 1989), was recombined with $flam^1$. These recombinant flies (ct^6 $flam^1$) also showed an increased lethal ecdysis deficiency with respect to $flam^1$ and ct^6 flies (Figure 2G). In addition, we found that the heteroallelic combination

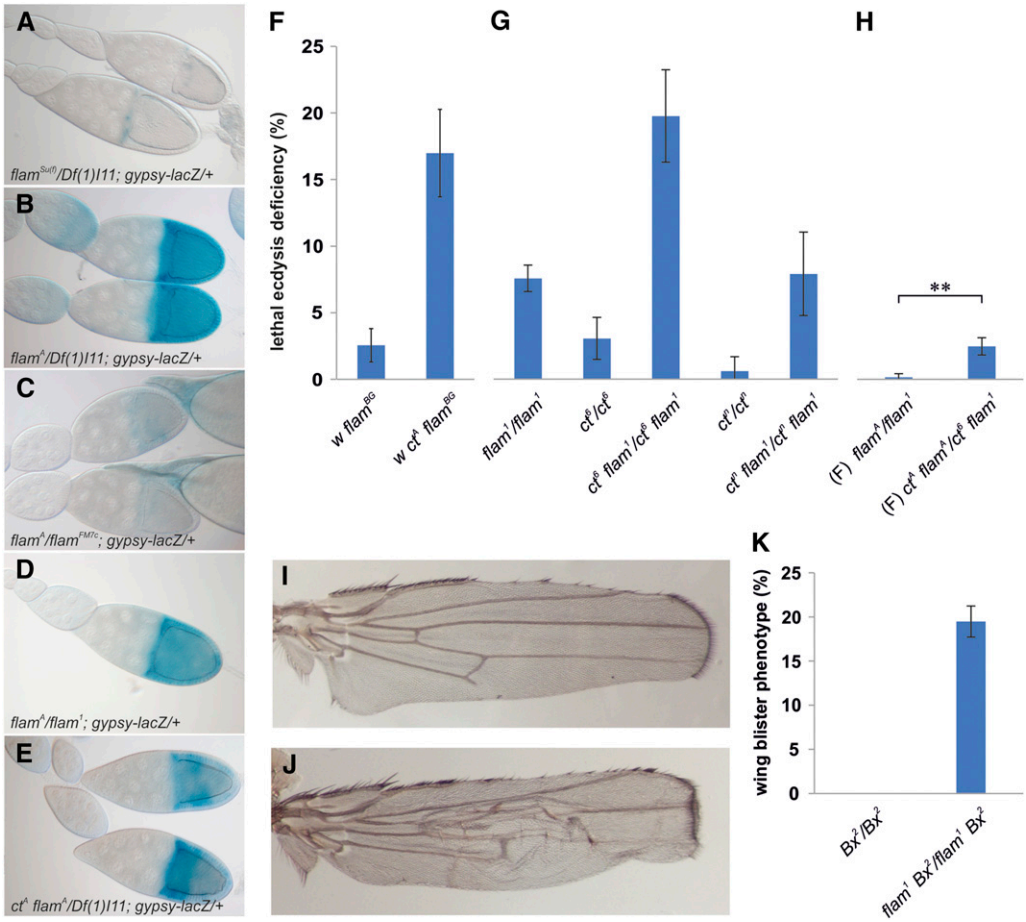

Figure 2 Genetic interactions between *flamenco* permissive alleles and *gypsy* insertions induce specific phenotypes. (A–E) Representative pictures of egg chambers subjected to β-galactosidase staining as readout for *gypsy-lacZ* reporter activity. (A) *gypsy-lacZ* reporter activity is repressed when the $flam^{su(f)}$ restrictive allele is combined with the *Df(1)l11* deficiency encompassing the *flamenco* locus. (B) When the same deficiency is combined with the *X* chromosome currying the $flam^{A}$ allele, *gypsy-lacZ* reporter activity is strongly induced. (C) *gypsy-lacZ* reporter activity is repressed when the $flam^{FM7c}$ restrictive allele is combined with the $flam^{A}$ allele. (D) *gypsy-lacZ* reporter activity is induced in $flam^{A}/flam^{1}$ ovaries. (E) Combination of the *Df(1)l11* deficiency with the ct^{A} allele in the $flam^{A}$-containing chromosome significantly reduces β-galactosidase staining. (F) Mortality during adult eclosion of ct^{A} $flam^{BG}$ flies compared to the $flam^{BG}$ strain. (G) Mortality during adult eclosion of ct^{6} and ct^{n} flies in their original genetic background and in combination with $flam^{1}$. (H) Mortality during adult eclosion of $flam^{A}/flam^{1}$ and ct^{A} $flam^{A}/ct^{6}$ $flam^{1}$ female (F) flies (** $P < 0.01$). (I, J) Wings derived from male flies hemizygous for Bx^{2} allele in (I) its original genetic background and (J) in $flam^{1}$ background. (K) Frequency of wing-blister phenotype of Bx^{2} wings in their original genetic background and in combination with $flam^{1}$.

ct^{A} $flam^{A}/ct^{6}$ $flam^{1}$ produced a low but significant increase of lethal ecdysis deficiency with respect to $flam^{A}/flam^{1}$, confirming the specificity of the genetic interaction (Figure 2H). Since *roo*-specific small RNAs are not produced by the *flamenco* locus (Yang *et al.* 2010), we explored whether the ct^{n} allele, which is induced by insertion of a *roo* element (Tchurikov *et al.* 1989), genetically interacts with $flam^{1}$. The ct^{n} $flam^{1}$ recombinant flies did not show an increased level of ecdysis failure with respect to $flam^{1}$ and ct^{n} flies (Figure 2G). This demonstrated that *flamenco* does not interact with a *cut* mutation induced by a TE that is not under its control, confirming that the ecdysis deficiency phenotype is independent from the *cut* loss of function. Finally, we tested whether a *gypsy*-induced mutation in a different gene was able to genetically interact with *flamenco*. To achieve this, a *gypsy* mutation in the *Beadex* (*Bx*) gene was recombined with $flam^{1}$. We used the gain-of-function mutation Bx^{2}, which is known to induce scalloped wing margins (Figure 2I) (Shoresh *et al.* 1998). We found that the Bx^{2} $flam^{1}$ recombinant showed wing blistering in addition to the wing-margin phenotype (Figure 2, J and K), confirming that a *gypsy* insertion in a gene different from *cut* in combination with a *flamenco* permissive allele can produce a specific phenotype.

All these observations support the hypothesis that the phenotypes we observed were specifically induced by a genetic interaction between a *gypsy* element inserted in a euchromatic site and a *flamenco* permissive allele.

flamenco activity is modulated by *gypsy* insertions into distant euchromatic loci

Due to the increase of *gypsy* copy number, a new *gypsy* insertion in a *flamenco* permissive background should result in a higher level of *gypsy* expression. Surprisingly, *gypsy* expression appeared significantly reduced in all stages and tissues analyzed from ct^{A} $flam^{A}$ homozygous flies (Figure 3, B–E). This suggested that the presence of the ct^{A} insertion produced a less permissive background for *gypsy*, which was confirmed by the reduction in *gypsy-lacZ* reporter activity in ovaries of ct^{A} $flam^{A}/Df(1)l11$ females (Figure 2E) compared to control (Figure 2B). Besides repression of the *gypsy* reporter, we found a significant reduction in the expression of endogenous *gypsy* elements (Figure S3A). To confirm that the downregulation of *gypsy* expression is specifically dependent on *flamenco* permissive alleles, the heteroallelic combination $flam^{A}/flam^{1}$ was also tested. We found a *gypsy* downregulation both when the *X* chromosomes carried two different *gypsy* insertions in the same locus, and when the *X*

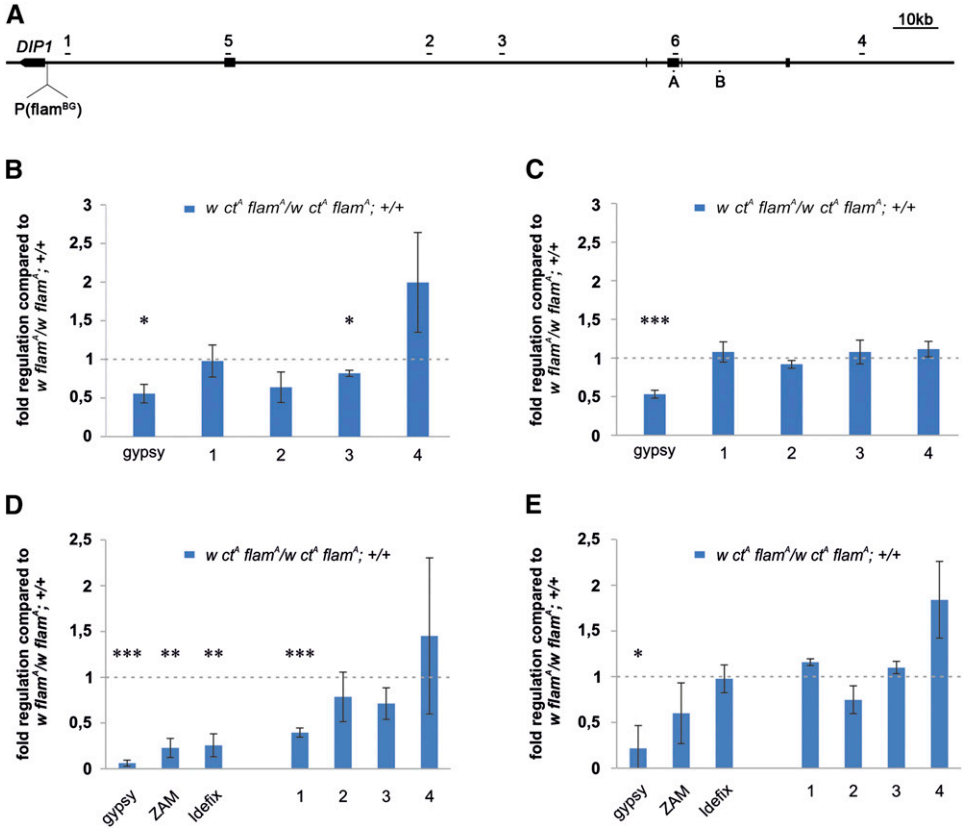

Figure 3 The *gypsy* insertion in the *cut* locus induces changes in the expression of *gypsy* and *flamenco*. (A) Map of the *flamenco* cluster showing the *DIP1* gene and the $flam^{BG02658}$ ($flam^{BG}$) insertion at the 5′ end of the locus, the *gypsy* transposon fragments (thick bars), the positions of primers used to analyze *flamenco* primary transcripts expression (numbers 1–6), and the position of tested piRNAs (letters A and B). (B–E) qRT-PCR analysis of *gypsy*, *ZAM*, *Idefix*, and specific *flamenco* regions in RNA isolated from (B) 48-hr-old pupae, (C) 96-hr-old pupae, (D) 0- to 24-hr-old female adult heads, and (E) 3-day-old adult ovaries. Shown are average levels ($n = 3$), and error bars indicate SD (* $P <$ 0.05, ** $P < 0.01$, *** $P < 0.005$).

chromosomes carried two different *gypsy* insertions in different loci (Figure S3B); confirming the specificity of the effect. Then, we explored whether the new *gypsy* insertion in the *cut* locus had consequences on the regulation of other retrotransposons regulated by *flamenco*. To this aim, the expression levels of *ZAM* and *Idefix* were analyzed in adult heads and ovaries. We found that the expression of these two retrotransposons was significantly decreased in heads while it appeared unaffected in ovaries (Figure 3, D and E). These data confirmed that a euchromatic insertion of *gypsy* could modify the expression of other retrotransposons regulated by *flamenco*, but that this regulation is tissue-dependent. It is possible that in ovaries, where the Piwi pathway is active, the regulatory mechanism induced by the new *gypsy* insertion takes place in a different way.

Since *gypsy* and other TEs regulated by *flamenco* showed a downregulation in the $flam^A$ permissive background, we hypothesized that the new *gypsy* insertion in the *cut* locus could also modify *flamenco* activity. Since *flamenco* is a source of small RNAs derived from the processing of a primary transcript, analysis of the expression levels at the *flamenco* locus could give an indication of the activity of the locus. Expression of the *flamenco* precursor transcript was analyzed by qRT-PCR using four pairs of primers designed to target different regions of the locus (Figure 3A). Two pairs were selected in sequences that map uniquely to the *flamenco* locus and are located at the 5′ end (primers 1) and in the central part (primers 3). The two other pairs of primers were specific for sequences that do not map uniquely to the *flamenco* locus (primers 2 and 4). The ct^A mutation had a minor or no effect on the transcription of *flamenco* during pupal stages (Figure 3, B and C) or in ovary tissues (Figure 3E). However, in heads, we observed a significant reduction at the 5′ region of the *flamenco* locus (Figure 3D). These data are in agreement with the hypothesis that *flamenco* activity can be modified by a new *gypsy* insertion in a euchromatic region.

To confirm the causal effect of *gypsy* euchromatic insertions on the *flamenco* activity, we looked for novel spontaneous mutations in the original $AstC\text{-}R2^{f03116}$ strain. We isolated a new mutation producing a forked-bristle phenotype (Figure S4A). This mutation, caused by a new *gypsy* insertion in the *forked* (*f*) gene, was named f^A. When placed in in a $y\ w^{67c23}$ background, the f^A allele only produced a small increase in mortality during adult eclosion (Figure S4B), and induced significant changes in the expression levels of *gypsy* and of a sequence that does not map uniquely to the *flamenco* locus (Figure S4C). Differences in expression levels in regions mapping uniquely to the *flamenco* locus were also found (Figure 5B), confirming the causal effect of *gypsy* euchromatic insertions on the *flamenco* activity.

The ecdysis phenotype induced by the gypsy insertion in the cut locus is based on an RNA-silencing mechanism

The involvement of an RNA-silencing mechanism is a possible hypothesis to explain in which way the *gypsy* insertion in the *cut* locus induces lethality during adult eclosion. Therefore

we decided to analyze the effects of mutations affecting the production of endo-siRNAs and piRNAs on this lethal phenotype. However, since loss-of-function mutations of endo-siRNA and piRNA genes increase heterochromatin formation at *flamenco* and at other small RNA clusters (Moshkovich and Lei 2010), we wondered if mutations affecting heterochromatin formation *per se* could modify the adult ecdysis phenotype induced by ct^A. Interestingly, we found that heterozygous mutations in *Su(var)3-9*, *Su(var)205*, and *Su(var)3-7*, which also suppress heterochromatic silencing at the *flamenco* locus (Moshkovich and Lei 2010), induced a strong reduction in the ecdysis phenotype (Figure 4A). Furthermore, we found that the increased lethal ecdysis deficiency found in ct^A $flam^A$ carrying the *TM6b* balancer chromosome (Figure 1F) was associated with an enhancement of variegation both at the *flamenco* locus (Figure S5, A and B and File S1) and at pericentromeric loci (Figure S5C). All these data confirm that modifications in heterochromatin silencing at the *flamenco* locus affect the ecdysis phenotype induced by ct^A. In addition, these observations suggested that analysis of homozygous mutants of the endo-siRNA and piRNA pathways should be avoided, since their effects on the heterochromatin of *flamenco* could mask any effects on post-transcriptional silencing. Notably, an increase of heterochromatin formation at the *flamenco* locus in *AGO2* and *piwi* heterozygous mutants seems to be excluded by ChIP assays performed using α-HP1 antibodies (Moshkovich and Lei 2010). Furthermore, there are indications that some TEs are also activated in heterozygous RNA interference-mutant conditions (Xie *et al.* 2013), suggesting that a heterozygous condition can be sufficient to produce a partial TE activation. Based on these data, we decided to analyze the effects of mutations affecting endo-siRNAs and piRNAs in heterozygous conditions. We found that heterozygous $AGO2^{414}$ and $Dcr\text{-}2^{L811fsX}$ mutations induced a significant reduction in lethal ecdysis phenotype (Figure 4B), supporting the hypothesis that the endo-siRNA pathway is involved in the post-transcriptional silencing of *flamenco*. Heterozygous mutations in *aub* and *AGO3* genes, which are expressed in some somatic tissues (Specchia *et al.* 2008; Perrat *et al.* 2013), also produced a partial rescue of the lethal ecdysis deficiency induced by ct^A (Figure 4B); indicating that piRNAs are also involved in the post-transcriptional silencing of *flamenco*. We also analyzed the effect of heterozygous mutations for both endo-siRNA and piRNA genes. The double heterozygous mutant produced an almost-complete rescue of the phenotype (Figure 4B), confirming the hypothesis that both these pathways are involved in the induction of the lethal ecdysis deficiency phenotype.

A possible explanation of the reduced *flamenco* expression in heads of $flam^A$ ct^A flies was that the *gypsy* insertion in the *cut* locus could induce heterochromatic silencing at the *flamenco* locus. However, ChIP experiments did not reveal increases in levels of the heterochromatic marks H3K9me3 and H3K27me3 in $flam^A$ ct^A heads (Figure 4, C–E), both in the analyzed regions of the *flamenco* locus and in the selected TE sequences. Although repressive marks were found from the *gypsy* sequence, we observed background levels of H3K9me3 and no increase in H3K27me3 into the genomic regions flanking the *gypsy* insertion (Figure 4, C–E, cut-up and cut-down primers). Based on these results, the observed changes in expression levels of *gypsy* and at the *flamenco* locus induced by ct^A are not accompanied by heterochromatin changes, suggesting post-transcriptional silencing.

All these data suggest that the lethal ecdysis deficiency phenotype requires TE silencing pathways. The RNA-silencing mechanism inducing this phenotype seems to act at the post-transcriptional level but is sensitive to heterochromatin modifications.

Euchromatic insertions of gypsy modify the flamenco transcript processing

We found significant reductions in *gypsy* expression induced by euchromatic insertions of *gypsy* in all developmental stages and tissues analyzed (Figure 3 and Figure S3), and confirmed by the *gypsy-lacZ* reporter assay (Figure 2E). This downregulation occurs without an increase in heterochromatic marks (Figure 4, C–E), suggesting post-transcriptional silencing. Furthermore, silencing of *gypsy* is accompanied by alterations in expression of specific *flamenco* regions and of other TEs regulated by *flamenco* (Figure 3D). An attractive hypothesis was that the sequences complementary to *gypsy* in the *flamenco* locus, required to produce small RNAs against *gypsy*, were downregulated along with *gypsy*. To confirm this hypothesis, we explored whether *gypsy* insertions in euchromatic regions caused a decrease in *gypsy* fragment sequences inside the *flamenco* transcript. To this aim, we analyzed the expression levels of two separate *gypsy* fragments (see Figure 3A, primers 5 and 6) in *w* ct^A $flam^A$ and *w* f^A $flam^A$ 48-hr-old pupae, using specific primers that only amplify *gypsy* fragments inside *flamenco*. We observed a significant reduction in the expression of the *gypsy* fragments in *w* ct^A $flam^A$ and in *w* f^A $flam^A$ pupae (Figure 5, A and B). Similar results were observed when analyzing the expression pattern of the *gypsy* fragments in *w* ct^A $flam^A$ adult heads (Figure 5C). Since the expression of the distal *gypsy* fragment was considerably low in the *w* $flam^A$ ovary (Figure 5D), we were unable to reliably test a reduction due to the ct^A mutation by qPCR. Expression of *gypsy* fragments was also analyzed in *80EF* and *42AB* dual-strand piRNA clusters, both in adult heads and in ovaries, without finding significant changes (Figure S6, A and B). This suggests that *flamenco* has specific properties that make it a preferential target of the silencing promoted by a new euchromatic *gypsy* insertion.

It has been reported that a reduction in the expression of *flamenco* primary transcript is accompanied by a reduction in the levels of mature piRNAs originating from this sequence (Brennecke *et al.* 2007). Expression analysis by qRT-PCR of the level of piRNA A, which maps to the distal *gypsy* fragment at ∼850 bp upstream to the region amplified by primers 6 (Figure 3A, region A), revealed a significant reduction in ovaries (Figure 5E). A similar reduction was found in the somatic tissues of the head (Figure S7A). To confirm the

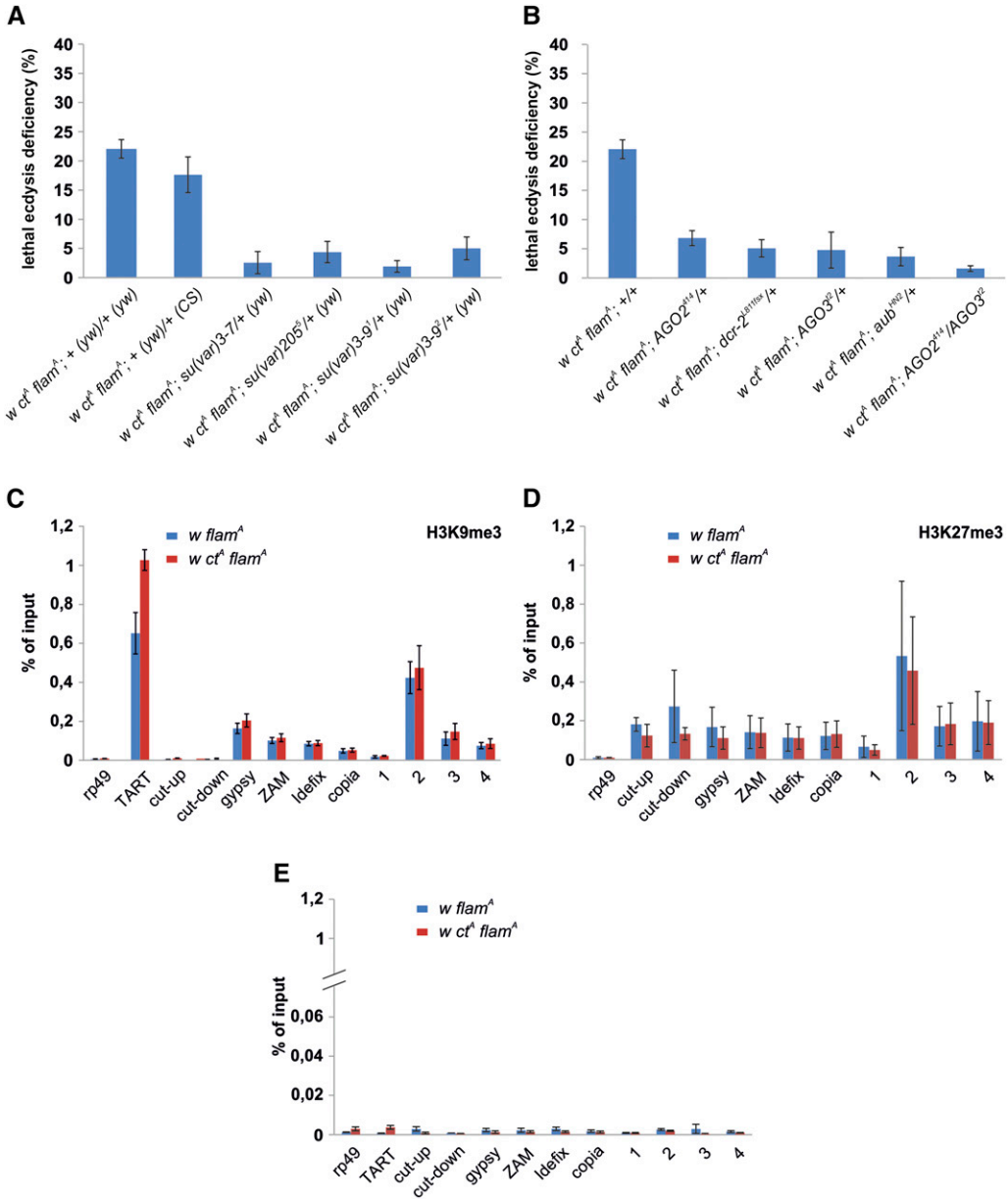

Figure 4 Lethal ecdysis deficiency and *flamenco* modulation induced by *gypsy* seems to depend on post-transcriptional silencing. (A, B) Mortality during adult eclosion of ct^A $flam^A$ flies carrying mutations affecting heterochromatin or small RNA pathways. (A) Analysis of mutations affecting heterochromatin formation. As an additional control, the heterozygous combination of a third chromosome from the yw^{67c23} strain and a third chromosome from the Canton-S strain was also analyzed. The genotype of the flies from which the third wild-type chromosome originates is enclosed in brackets. (B) Analysis of mutations affecting endo-siRNA and piRNA production. (C–E) ChIP analysis. Cross-linked chromatin of adult heads from *w* $flam^A$ and *w* ct^A $flam^A$ females was IP with antibodies specific to (C) H3K9me3, (D) H3K27me3, and (E) no antibody as negative control. The IP DNA was analyzed by qPCR. Protein binding is expressed as the percentage of input and is shown for each primer set. cut-up and cut-down primers amplify upstream and downstream of the *gypsy* insertion at the *cut* locus, respectively. *TART* is used as a positive region for H3K9me3, and *rp49* as a negative region showing background levels. Shown are average levels from two independent experiments and error bars indicate average deviation.

reduction of piRNA A levels, the Piwi protein was IP from ovary extracts (Saito *et al.* 2006) and piRNAs from Piwi IP were subjected to northern blot analysis using a specific probe (Figure S7B and File S1). Although at a lower extent compared to the RT-qPCR analysis, this alternative experimental approach confirmed the downregulation of piRNAs mapping to region A in ct^A $flam^A$ ovaries. Since *ZAM* expression is not significantly decreased in ct^A $flam^A$ ovaries (Figure 3E), we analyzed the expression level of piRNA B originating from a *ZAM* fragment of *flamenco* (Figure 3A, region B) by qRT-PCR, without observing a significant reduction (Figure 5E). These data confirm that new *gypsy* euchromatic insertions promote a reduction in the expression of complementary regions in the *flamenco* transcript and of a piRNA originating from these sequences. The simplest explanation of the downregulation in the expression of *gypsy* fragment sequences inside the *flamenco* transcript and the reduction in specific piRNAs is a degradation of *gypsy* complementary sequences inside the *flamenco* primary transcript.

The gypsy insertion in the cut locus promotes production of fused transcripts containing gypsy and flanking genomic sequences

Decreased levels of *gypsy* and *gypsy* sequences in the *flamenco* transcript may be explained by a cosuppression triggered by the new *gypsy* insertion in the *cut* locus. This suggested that the *gypsy* insertion site becomes a new small RNA-generating locus able to silence *gypsy* sense and antisense sequences. This hypothesis would be supported by the fact that in the *Drosophila* germline, new euchromatic TE insertions induce a bidirectional transcription with formation of RNA containing TE and flanking genomic sequences, triggering

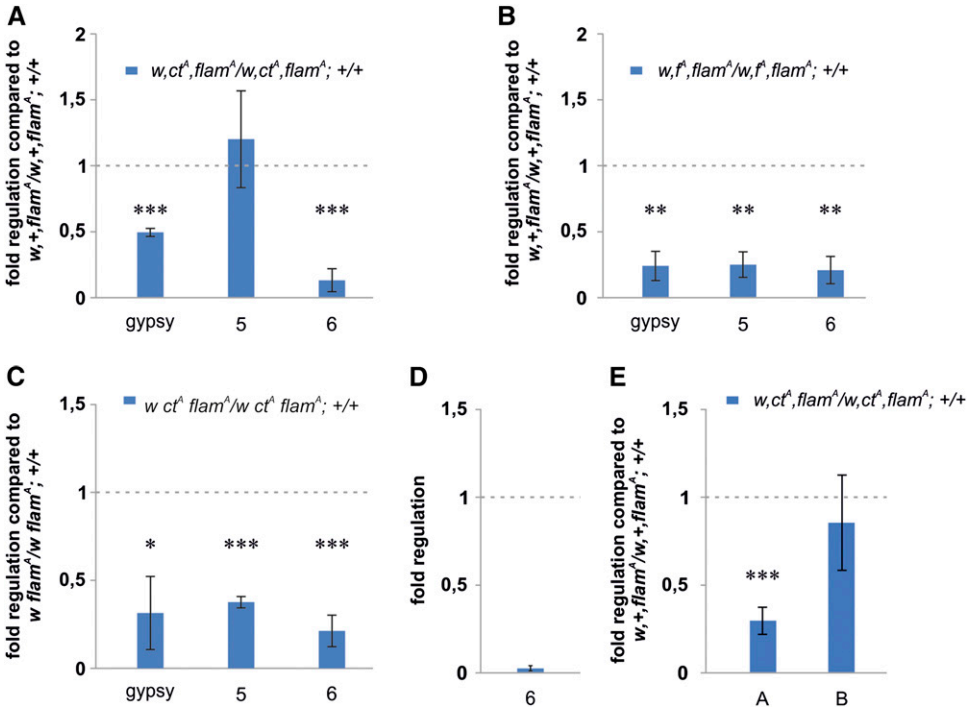

Figure 5 Cosuppression of *gypsy* and *flamenco* sequences from which small RNAs against *gypsy* originate. qRT-PCR analysis. Shown are the transcript levels of *gypsy* and two *gypsy* fragments in the *flamenco* locus (primers 5–6, described in Figure 3A) in RNA isolated from 48-hr-old pupae carrying the homozygous (A) ct^A or (B) f^A mutation induced by *gypsy*. (C) Expression of *gypsy* and *gypsy* fragments inside *flamenco* in 0- to 24-hr-old adult heads from flies carrying the ct^A homozygous mutation. (D) Expression of the distal *gypsy* fragment (primers 6) in *w* $flam^A$ adult ovaries compared to that in 48-hr-old pupae. (E) piRNA levels in 3-day-old adult ovary detected by TaqMan specific probes (see *Materials and Methods* for details). Position of analyzed piRNAs loci are indicated in Figure 3A. Shown are average levels ($n = 3$), and error bars indicate SD (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.005$).

the production of novel endo-siRNAs and pi-RNAs (Shpiz *et al.* 2014). We tested whether the *gypsy* insertion in the *cut* locus promotes the production of RNAs containing *gypsy* and flanking sequences by qPCR experiments performed using cDNAs obtained with strand-specific primers (Figure 6A). Expression in ct^A $flam^A/ct^A$ $flam^A$ and ct^A $flam^A/f^A$ $flam^A$ heads was compared with that in control heads ($flam^A/flam^A$). We used ct^A $flam^A/f^A$ $flam^A$ flies because they show a very low level of mortality during adult eclosion (Figure S4B), while they also carry two *gypsy* insertions as ct^A $flam^A$ homozygous flies. This suggests that the effect on the *flamenco* locus is different in flies with a different pattern of *gypsy* insertions. We analyzed expression of the genomic region where the *gypsy* insertion of the ct^A allele maps, finding a significant expression in the telomere to centromere (forward) direction and a very weak expression in the centromere to telomere (reverse) direction (Figure 6, B and C). Surprisingly, the presence of even a single ct^A allele was sufficient to produce a strong activation in the reverse expression in the cut-up site and a strong suppression of the forward expression in the cut-down site (Figure 6, B and C). These data demonstrate that the *gypsy* insertions can modify expression of the regions surrounding the insertion site. We also found transcription from the flanking sequences into the *gypsy* sequence in the forward and reverse direction (Figure 6, D and E). Expression in the forward direction appeared quite similar, while the reverse expression of the ct^A homozygous heads was about two times higher than that of ct^A f^A double heterozygous heads (Figure 6E). Reverse expression produces a fused transcript carrying sense *gypsy* sequences that can potentially be processed to produce small RNAs against *gypsy* antisense sequences. If these *gypsy* sense sequences served as a precursor for small RNAs, a lower level of *gypsy* sequences in the *flamenco* transcript from ct^A homozygous heads compared to ct^A f^A double heterozygous heads would be expected. Indeed, expression of regions 5 and 6 was lower in ct^A homozygous heads as well as expression of the 5′ of the *flamenco* transcript (Figure 6F). This suggests a higher level of degradation of the *flamenco* transcript in ct^A homozygous heads. *gypsy* expression was not significantly different (Figure 6F), suggesting that the equal level of *gypsy* antisense sequences produces the same effect in ct^A homozygotes and in ct^A f^A double heterozygotes. These data support the hypothesis that the *gypsy* insertion in the *cut* locus triggers the production of novel small RNAs directed against the sense sequence of the *gypsy* retrotransposon element and the *gypsy* antisense sequences that are part of the *flamenco* transcript.

Discussion

RNA-silencing mechanisms based on the production of small noncoding RNAs deriving from small RNA clusters have evolved to protect cells against TEs. Our understanding of the function of these clusters in the repression of TE activity is also derived from studies on the *D. melanogaster flamenco* locus. In the present work we found that new euchromatic *gypsy* insertions modify the activity of the *flamenco* locus and induce developmental alterations. This suggests that *flamenco* directly or indirectly regulates other processes in addition to TE silencing, and in fact it has been shown previously that insertional mutations in the *flamenco* locus induce alterations in oogenesis (Mevel-Ninio *et al.* 2007).

Phenotypes induced by euchromatic *gypsy* insertions seem to be very sensitive to the genetic background. In fact, we

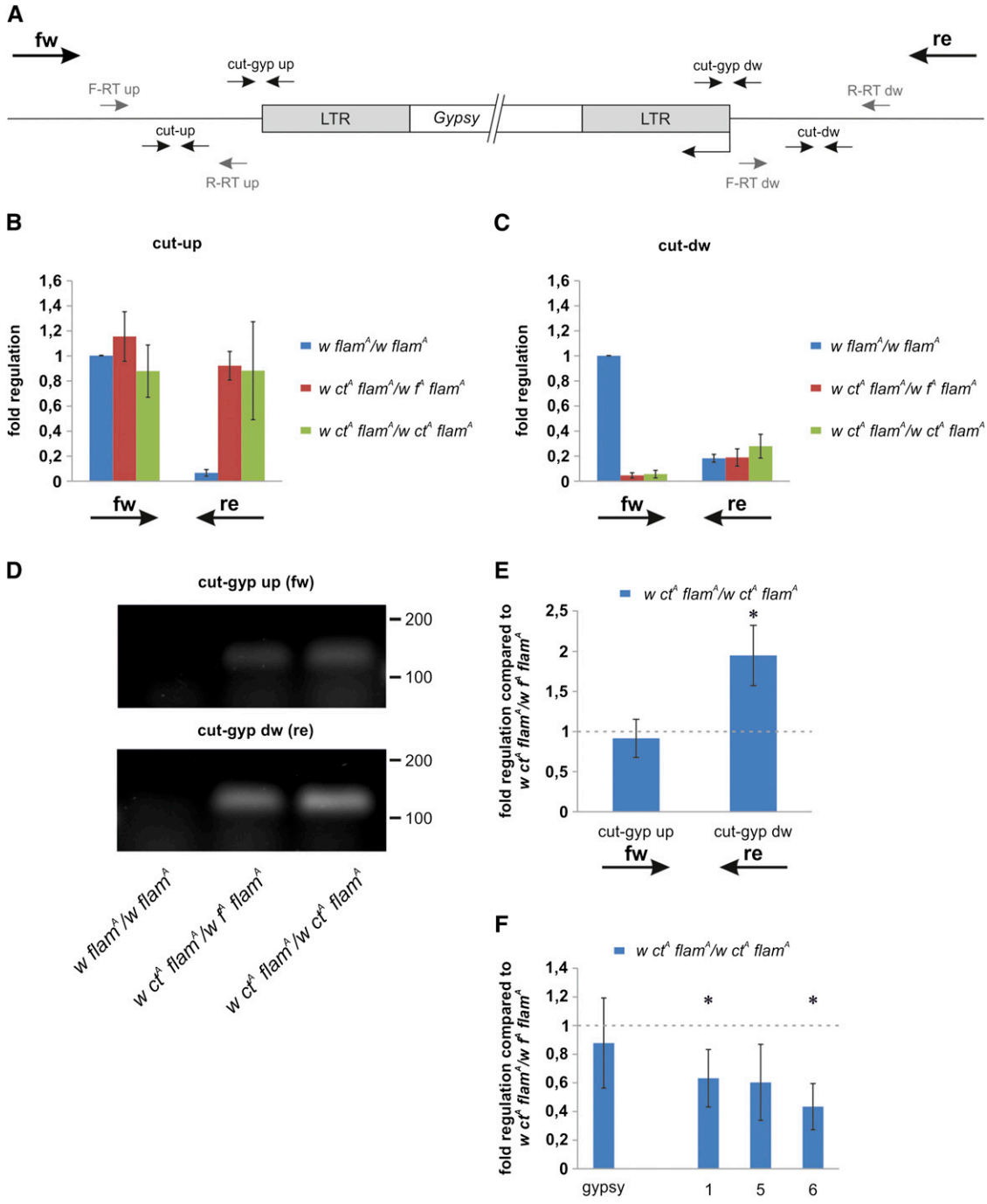

Figure 6 The *gypsy* insertion in the *cut* locus triggers the expression of fused transcripts from flanking genomic regions into the *gypsy* sequences. (A) Map of the *gypsy* insertion site in the *cut* locus and of the primers used in RT-PCR experiments (small arrows). Stand specific RT primers are represented in grey while qPCR primers are represented in black. fw arrow indicates direction of telomere-to-centromere expression while re arrow indicates direction of centromere-to-telomere expression. (B–F) RT-PCR experiments performed using RNA isolated from 0- to 24-hr-old female heads. (B, C) Strand-specific qRT-PCR reveals transcription in the genomic regions surrounding the *gypsy* insertion site in the *cut* locus. Transcription levels are variable and influenced by the pattern of the *gypsy* mutations present in each strain. Shown are average levels (n = 3), and error bars indicate SD. Retrotranscription of the cut-up site was primed by (B) primers F-RT up and R-RT up, while retrotranscription of the cut-down site was primed by (C) primers F-RT dw and R-RT dw. Primers cut-up were used to amplify the cut-up region while primers cut-dw were used to amplify the cut-down region. (D) Strand-specific RT-PCR products loaded in a 2% agarose gel confirming the presence of fused transcripts only in flies carrying the *gypsy* insertion in the *cut* locus. Primers cut-gyp up were used to amplify the forward fused transcript while primers cut-gyp dw were used to amplify the reverse fused transcript. DNA ladder scale is in base pairs. (E) Strand-specific qRT-PCR analysis of the transcription levels of the regions cut-gyp up and cut-gyp dw. Shown are average levels (n = 3), and error bars indicate SD (* $P < 0.01$). (F) qRT-PCR analysis. Shown are the transcript levels of *gypsy* and of three regions inside the *flamenco* locus (primers 1, 5, and 6; described in Figure 3A). Shown are average levels (n = 3), and error bars indicate SD (* $P < 0.05$).

found significant changes in the frequency of the ecdysis deficiency phenotype induced by *gypsy* in presence of mutations or a balancer chromosome inducing epigenetic modifications. The low level of lethal ecdysis deficiency found in the heteroallelic combination ct^A $flam^A/ct^6$ $flam^1$ (Figure 2H) could be linked to the hybrid vigor, which also seems to depend on epigenetic changes that are found in hybrids (Chen 2013). Since mortality during adult eclosion induced by the *gypsy* insertion in the *cut* locus is modified by epigenetic modifications, it is likely that the weaker hybrid phenotypes have epigenetic causes. The number and localization of preexisting and new *gypsy* insertions in the different genetic backgrounds could also contribute to explain the variability (type and frequency) of the particular phenotypes arising from the genetic interaction between *gypsy* and *flamenco*. Regarding the localization of the *gypsy* insertions, we found different levels of lethal ecdysis deficiency in flies with the same genetic background, the same number of new *gypsy* mutations, and different *gypsy* localizations (Figure 1F and Figure S4B). Different effects were also found when the expression of *gypsy* and *gypsy* fragments inside the *flamenco* transcript were compared between two of these strains (Figure 6F).

The involvement of the piRNA pathway in the induction of the lethal ecdysis phenotype produced by ct^A would confirm an important role of piRNA in somatic tissues. A contribution of maternal inheritance of piRNAs cannot be excluded, since it has been reported that epigenetic information can be transmitted to the offspring in the form of small RNAs (Akkouche *et al.* 2013). A deeper genetic analysis will be required to better understand how piRNAs operate using this mechanism.

A cosupression mechanism seems to be at the base of the altered flamenco activity induced by gypsy

We examined the effect of a new mutation induced by a *gypsy* insertion in the *cut* locus, which resulted in significant changes in the expression of specific sequences of the *flamenco* locus. We found that the expression of *gypsy* is downregulated as well as the expression of sequences that map uniquely to the *flamenco* locus inside regions that give rise to small RNAs which specifically target *gypsy*. This mechanism is reminiscent of the phenomenon described as cosuppression in plants, where a transgene reduces the expression of homologous sequences through a RNA-mediated silencing mechanism (Smyth 1997; Montgomery and Fire 1998). Overall, our data suggest that *gypsy* elements may cause post-transcriptional silencing of *flamenco*. This does not seem strange because small RNAs antisense to the *flamenco* transcript are found in databases (Yang *et al.* 2010), suggesting that small RNA clusters are natural targets of RNA silencing processes. Small RNAs antisense to *flamenco* also map to *gypsy* fragment sequences. It is known that endo-siRNAs and piRNAs also derive from TE sequences (Saito and Siomi 2010), and that in the *Drosophila* germline new euchromatic TE insertions trigger the production of endo-siRNAs and piRNAs from TE and surrounding sequences (Shpiz *et al.* 2014). This suggests that the *gypsy* insertion in the *cut* locus may become a new small RNA-generating locus. Production of sense- and antisense-fused transcripts containing *gypsy* and the flanking genomic sequences at the *gypsy* insertion site in the *cut* locus (Figure 6, D and E) is in agreement with this hypothesis. Generation of new small RNAs from these *gypsy* sequences can explain the concomitant silencing of the *gypsy* retrotransposon and antisense *gypsy* fragment sequences. This can also explain the reduction in *gypsy* piRNA A level that we observed in ovaries and in adult heads. Reduction of piRNAs also exclude that the decreased levels of *gypsy* transcript and of *flamenco* precursor transcript derives from an increased production of *flamenco* piRNAs.

We did not find significant alterations in the expression of *gypsy* sequences located in other small RNA clusters (Figure S6). This may depend on the fact that *80EF* and *42AB* are dual-stranded clusters while *flamenco* is a single-stranded cluster, and they may have a different regulation. Expression of other *gypsy* fragments in different uni-strand and dual-strand clusters should be tested to confirm this hypothesis. Finally, we found that the region amplified by primers 2, which is a repeated sequence, may undergo significant changes (Figure S4C); suggesting that other genomic regions are involved in this regulatory mechanism. To have a broader view, it will be interesting to test if other TEs can modulate the activity of the small RNA clusters responsible for their suppression.

Reciprocal regulation of gypsy and flamenco probably influences the regulation of distal genes and sequences

Different experimental evidence indicates that *flamenco*, like other small RNAs clusters, is a heterochromatic region (Moshkovich and Lei 2010). Heterochromatin modifications can modulate *flamenco* activity, changing the expression pattern of the locus. However, we did not observe major changes in H3K9me3 and H3K27me3 enrichment at the *flamenco* locus in ct^A homozygous flies (Figure 4, C–E). Furthermore, we did not find a spreading of heterochromatin in genomic sequences surrounding the *gypsy* insertion at the *cut* locus, as occurs when *gypsy* or other TEs are transcriptionally silenced by heterochromatic formation (Sienski *et al.* 2012). These findings suggest that the silencing of *gypsy* and *flamenco* that we found in the $flam^A$ background depends on a post-transcriptional mechanism. Since *flamenco* seems to be transcribed as a continuous single-stranded precursor spanning the locus (Brennecke *et al.* 2007; Goriaux *et al.* 2014a), differences in expression levels of specific *flamenco* regions are likely the result of differences in the processing of the precursor. In fact, differential changes in expression of specific regions of the *flamenco* locus have also been found in *zucchini* mutants, which have a role in the primary processing of piRNAs from the *flamenco* locus (Haase *et al.* 2010) and in comparing *flamenco* restrictive and permissive strains (Lavrenov *et al.* 2014). We propose that the differences in expression levels found at the specific *flamenco* regions analyzed, which are induced by new euchromatic *gypsy* insertions, depend on a degradation of specific sequences of the *flamenco* precursor transcript.

It is plausible that the high complexity of the small RNA clusters resulted in the acquisition of novel regulatory functions besides the original TE silencing function. A quantitative change in the amount of small RNAs, derived from *flamenco* and other small RNA clusters, can have an effect on the regulation of multiple target sequences. Potential targets regulated by *flamenco* are mRNAs that share sequence complementarity with small RNAs originating from *flamenco* and genes near the insertion site of *gypsy*, or of other TEs regulated by the small RNA cluster. It has been demonstrated that some TE-derived piRNAs and endo-siRNAs function as regulators of host genes (Rouget *et al.* 2010; McCue *et al.* 2012), and it has been suggested that TE-derived small noncoding RNAs may have a large influence on the transcriptome (McCue and Slotkin 2012). It is also known that when TEs are inserted into euchromatic regions, the expression of neighboring host genes can be modulated by ectopic recruitment of silencing factors (Slotkin and Martienssen 2007; Cowley and Oakey 2013). Although we did not find modifications in heterochromatic marks near to the *gypsy* insertion site into the *cut* locus, we cannot exclude that heterochromatin modifications can occur in other places near to other *gypsy* or TE sequences. This is supported by the fact that *flamenco* permissive alleles induce a genome-wide redistribution of the HP1 protein, as well as suppression of heterochromatic silencing in *trans* (Moshkovich and Lei 2010). We observed about a onefold increase in H3K9me3 level at the telomere-associated retrotransposon *TART* (Figure 4C), suggesting possible changes in heterochromatin distribution.

Identification of the target genes responsible for the different phenotypes induced by the *gypsy* insertions will be important to understand the mechanism by which the reciprocal regulation of *gypsy* and *flamenco* modifies fundamental developmental processes. This will help to clarify whether the developmental phenotypes induced by euchromatic *gypsy* insertions depend on the regulatory effects on the *flamenco* activity, or whether they directly depend on the production of new small RNAs from the *gypsy* insertion site. In the latter case, the lack of repression of *gypsy* due to the *flamenco* permissive background would favor the activation of the new small RNA cluster.

Acknowledgments

A special thanks to the late Franco Graziani, without whom this work would never have been possible. We thank Valeria Cavaliere, Giuseppe Gargiulo, Adriana La Volpe, Paolo Bazzicalupo, Umberto di Porzio, Aurora Storlazzi, and Massimo Di Giulio for helpful discussion, advice, and critical reading of the manuscript. We thank Silvia Andone and Yukiko Murota for technical support, Tamas Schauer for advice about chromatin preparation and immunoprecipitation, Marilena Ignesti for her useful advice, and Giovanna Lattanzi for her help during the revision of the manuscript. We thank the Institute of Genetics and Biophysics for use of the Integrated Microscopy Facility. We also thank Alain Pélisson for fly strains. This work was supported by the Italian Ministry of Economy and Finance "Medical Research Italy" (MERIT-RBNE08LN4P-002), the Italian Ministry of Economy and Finance grant project FaReBio di Qualità, the 2012 "5 per mille" project to the Rizzoli Laboratory of Musculoskeletal Cell Biology, and Deutsche Forschungsgemeinschaft (SFB 1064, SPP1356).

Literature Cited

Akkouche, A., T. Grentzinger, M. Fablet, C. Armenise, N. Burlet *et al.*, 2013 Maternally deposited germline piRNAs silence the tirant retrotransposon in somatic cells. EMBO Rep. 14: 458–464.

Bergman, C. M., H. Quesneville, D. Anxolabehere, and M. Ashburner, 2006 Recurrent insertion and duplication generate networks of transposable element sequences in the *Drosophila melanogaster* genome. Genome Biol. 7: R112.

Brennecke, J., A. A. Aravin, A. Stark, M. Dus, M. Kellis *et al.*, 2007 Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. Cell 128: 1089–1103.

Bucheton, A., 1995 The relationship between the flamenco gene and gypsy in *Drosophila*: how to tame a retrovirus. Trends Genet. 11: 349–353.

Buchon, N., and C. Vaury, 2006 RNAi: a defensive RNA-silencing against viruses and transposable elements. Heredity 96: 195–202.

Chalvet, F., L. Teysset, C. Terzian, N. Prud'homme, P. Santamaria *et al.*, 1999 Proviral amplification of the Gypsy endogenous retrovirus of *Drosophila melanogaster* involves env-independent invasion of the female germline. EMBO J. 18: 2659–2669.

Chen, Z. J., 2013 Genomic and epigenetic insights into the molecular bases of heterosis. Nat. Rev. Genet. 14: 471–482.

Cowley, M., and R. J. Oakey, 2013 Transposable elements re-wire and fine-tune the transcriptome. PLoS Genet. 9: e1003234.

Czech, B., C. D. Malone, R. Zhou, A. Stark, C. Schlingeheyde *et al.*, 2008 An endogenous small interfering RNA pathway in *Drosophila*. Nature 453: 798–802.

de Koning, A. P., W. Gu, T. A. Castoe, M. A. Batzer, and D. D. Pollock, 2011 Repetitive elements may comprise over two-thirds of the human genome. PLoS Genet. 7: e1002384.

Desset, S., C. Conte, P. Dimitri, V. Calco, B. Dastugue *et al.*, 1999 Mobilization of two retroelements, ZAM and Idefix, in a novel unstable line of *Drosophila melanogaster*. Mol. Biol. Evol. 16: 54–66.

Desset, S., C. Meignin, B. Dastugue, and C. Vaury, 2003 COM, a heterochromatic locus governing the control of independent endogenous retroviruses from *Drosophila melanogaster*. Genetics 164: 501–509.

Elbarbary, R. A., B. A. Lucas, and L. E. Maquat, 2016 Retrotransposons as regulators of gene expression. Science 351: aac7247.

Fedoroff, N. V., 2012 Presidential address. Transposable elements, epigenetics, and genome evolution. Science 338: 758–767.

Ghildiyal, M., H. Seitz, M. D. Horwich, C. J. Li, T. T. Du *et al.*, 2008 Endogenous siRNAs derived from transposons and mRNAs in *Drosophila* somatic cells. Science 320: 1077–1081.

Goriaux, C., S. Desset, Y. Renaud, C. Vaury, and E. Brasset, 2014a Transcriptional properties and splicing of the flamenco piRNA cluster. EMBO Rep. 15: 411–418.

Goriaux, C., E. Theron, E. Brasset, and C. Vaury, 2014b History of the discovery of a master locus producing piRNAs: the flamenco/COM locus in Drosophila melanogaster. Front. Genet. 5: 257.

Gunawardane, L. S., K. Saito, K. M. Nishida, K. Miyoshi, Y. Kawamura *et al.*, 2007 A slicer-mediated mechanism for repeat-associated siRNA 5′ end formation in *Drosophila*. Science 315: 1587–1590.

Haase, A. D., S. Fenoglio, F. Muerdter, P. M. Guzzardo, B. Czech *et al.*, 2010 Probing the initiation and effector phases of the somatic piRNA pathway in Drosophila. Genes Dev. 24: 2499–2504.

Harrison, D. A., D. A. Gdula, R. S. Coyne, and V. G. Corces, 1993 A Leucine-Zipper domain of the suppressor of hairy-wing protein mediates its repressive effect on enhancer function. Genes Dev. 7: 1966–1978.

Jack, J. W., 1985 Molecular organization of the cut locus of *Drosophila melanogaster*. Cell 42: 869–876.

Kalmykova, A. I., M. S. Klenov, and V. A. Gvozdev, 2005 Argonaute protein PIWI controls mobilization of retrotransposons in the *Drosophila* male germline. Nucleic Acids Res. 33: 2052–2059.

Lavrenov, A. R., L. N. Nefedova, N. I. Romanova, and A. I. Kim, 2014 Expression of hp1 family genes and their plausible role in formation of flamenco phenotype in *D. melanogaster*. Biochemistry (Mosc.) 79: 1267–1272.

Li, C., V. V. Vagin, S. Lee, J. Xu, S. Ma *et al.*, 2009 Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies. Cell 137: 509–521.

Malone, C. D., J. Brennecke, M. Dus, A. Stark, W. R. Mccombie *et al.*, 2009 Specialized piRNA pathways act in germline and somatic tissues of the *Drosophila* ovary. Cell 137: 522–535.

McCue, A. D., and R. K. Slotkin, 2012 Transposable element small RNAs as regulators of gene expression. Trends Genet. 28: 616–623.

McCue, A. D., S. Nuthikattu, S. H. Reeder, and R. K. Slotkin, 2012 Gene expression and stress response mediated by the epigenetic regulation of a transposable element small RNA. PLoS Genet. 8: e1002474.

Mevel-Ninio, M., A. Pélisson, J. Kinder, A. R. Campos, and A. Bucheton, 2007 The flamenco locus controls the gypsy and ZAM retroviruses and is required for *Drosophila* oogenesis. Genetics 175: 1615–1624.

Mita, P., and J. D. Boeke, 2016 How retrotransposons shape genome regulation. Curr. Opin. Genet. Dev. 37: 90–100.

Montgomery, M. K., and A. Fire, 1998 Double-stranded RNA as a mediator in sequence-specific genetic silencing and co-suppression. Trends Genet. 14: 255–258.

Moshkovich, N., and E. P. Lei, 2010 HP1 recruitment in the absence of argonaute proteins in *Drosophila*. PLoS Genet. 6: e1000880.

Pélisson, A., S. U. Song, N. Prudhomme, P. A. Smith, A. Bucheton *et al.*, 1994 Gypsy transposition correlates with the production of a retroviral envelope-like protein under the tissue-specific control of the *Drosophila* flamenco gene. EMBO J. 13: 4401–4411.

Pélisson, A., L. Teysset, F. Chalvet, A. Kim, N. Prud'homme *et al.*, 1997 About the origin of retroviruses and the co-evolution of the gypsy retrovirus with the Drosophila flamenco host gene. Genetica 100: 29–37.

Perrat, P. N., S. Dasgupta, J. Wang, W. Theurkauf, Z. P. Weng *et al.*, 2013 Transposition-driven genomic heterogeneity in the *Drosophila* brain. Science 340: 91–95.

Prud'homme, N., M. Gans, M. Masson, C. Terzian, and A. Bucheton, 1995 Flamenco, a gene controlling the gypsy retrovirus of Drosophila melanogaster. Genetics 139: 697–711.

Rouget, C., C. Papin, A. Boureux, A. C. Meunier, B. Franco *et al.*, 2010 Maternal mRNA deadenylation and decay by the piRNA pathway in the early *Drosophila* embryo. Nature 467: 1128–1132.

Saito, K., and M. C. Siomi, 2010 Small RNA-mediated quiescence of transposable elements in animals. Dev. Cell 19: 687–697.

Saito, K., K. M. Nishida, T. Mori, Y. Kawamura, K. Miyoshi *et al.*, 2006 Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the *Drosophila* genome. Genes Dev. 20: 2214–2222.

Sarot, E., G. Payen-Groschene, A. Bucheton, and A. Pélisson, 2004 Evidence for a piwi-dependent RNA silencing of the gypsy endogenous retrovirus by the *Drosophila melanogaster* flamenco gene. Genetics 166: 1313–1321.

Schauer, T., P. C. Schwalie, A. Handley, C. E. Margulies, P. Flicek *et al.*, 2013 CAST-ChIP maps cell-type-specific chromatin states in the *Drosophila* central nervous system. Cell Rep. 5: 271–282.

Shoresh, M., S. Orgad, O. Shmueli, R. Werczberger, D. Gelbaum *et al.*, 1998 Overexpression Beadex mutations and loss-of-function heldup-a mutations in *Drosophila* affect the 3′ regulatory and coding components, respectively, of the Dlmo gene. Genetics 150: 283–299.

Shpiz, S., S. Ryazansky, I. Olovnikov, Y. Abramov, and A. Kalmykova, 2014 Euchromatic transposon insertions trigger production of novel Pi- and endo-siRNAs at the target sites in the drosophila germline. PLoS Genet. 10: e1004138.

Sienski, G., D. Donertas, and J. Brennecke, 2012 Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression. Cell 151: 964–980.

Slotkin, R. K., and R. Martienssen, 2007 Transposable elements and the epigenetic regulation of the genome. Nat. Rev. Genet. 8: 272–285.

Smyth, D. R., 1997 Gene silencing: cosuppression at a distance. Curr. Biol. 7: R793–R795.

Song, S. U., M. Kurkulos, J. D. Boeke, and V. G. Corces, 1997 Infection of the germ line by retroviral particles produced in the follicle cells: a possible mechanism for the mobilization of the gypsy retroelement of *Drosophila*. Development 124: 2789–2798.

Specchia, V., C. Benna, G. M. Mazzotta, A. Piccin, M. A. Zordan *et al.*, 2008 Aubergine gene overexpression in somatic tissues of aubergine(sting) mutants interferes with the RNAi pathway of a yellow hairpin dsRNA in *Drosophila melanogaster*. Genetics 178: 1271–1282.

Tchurikov, N. A., T. I. Gerasimova, T. K. Johnson, N. I. Barbakar, A. L. Kenzior *et al.*, 1989 Mobile elements and transposition events in the cut locus of *Drosophila melanogaster*. Mol. Gen. Genet. 219: 241–248.

Warren, I. A., M. Naville, D. Chalopin, P. Levin, C. S. Berger *et al.*, 2015 Evolutionary impact of transposable elements on genomic diversity and lineage-specific innovation in vertebrates. Chromosome Res. 23: 505–531.

Xie, W., R. C. Donohue, and J. A. Birchler, 2013 Quantitatively increased somatic transposition of transposable elements in *Drosophila* strains compromised for RNAi. PLoS One 8: e72163.

Yang, J. H., P. Shao, H. Zhou, Y. Q. Chen, and L. H. Qu, 2010 DeepBase: a database for deeply annotating and mining deep sequencing data. Nucleic Acids Res. 38: D123–D130.

Zanni, V., A. Eymery, M. Coiffet, M. Zytnicki, I. Luyten *et al.*, 2013 Distribution, evolution, and diversity of retrotransposons at the flamenco locus reflect the regulatory properties of piRNA clusters. Proc. Natl. Acad. Sci. USA 110: 19842–19847.

Communicating editor: J. A. Birchler

Figure S1

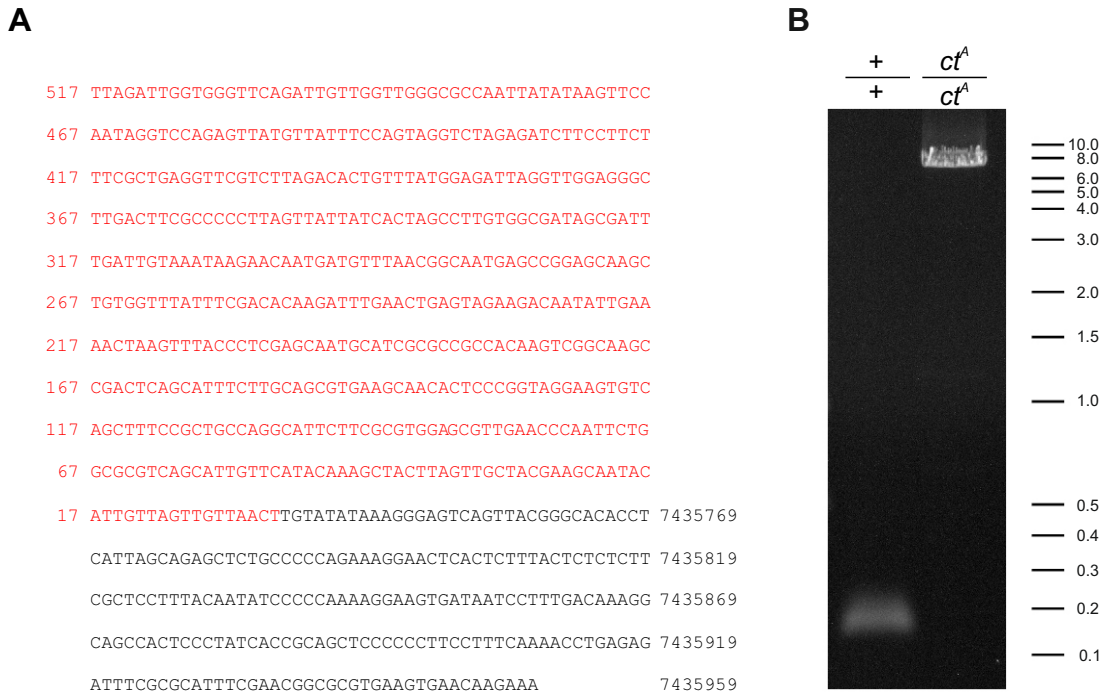

Figure S1. Identification of the *gypsy* insertion site at the *cut* locus.

(A) The g*ypsy* insertion site in the regulatory regions of the *cut* gene was identified by PCR using a primer inside the *gypsy* element and a set of primers designed upstream of the *cut* gene (see Supplemental Experimental Procedures). The *gypsy* insertion site maps in a region between the wing margin-specific enhancer and the promoter of the *cut* gene, a site where other *cut* wing mutations also map (Kim et al., 1996). In red is represented the 5'*gypsy* sequence and in black the flanking genomic region (FB2015_01 release) (St Pierre et al., 2014). (B) 0,8% agarose gel showing the presence of an amplicon of about 0,14 kb in *w flam*^{A}*/w flam*^{A} flies (+/+) and of about 7,5 kb in *w ct*^{A} *flam*^{A}*/w ct*^{A} *flam*^{A} flies (*ct*^{A}*/ ct*^{A}). DNA ladder scale is in kb.

Figure S2

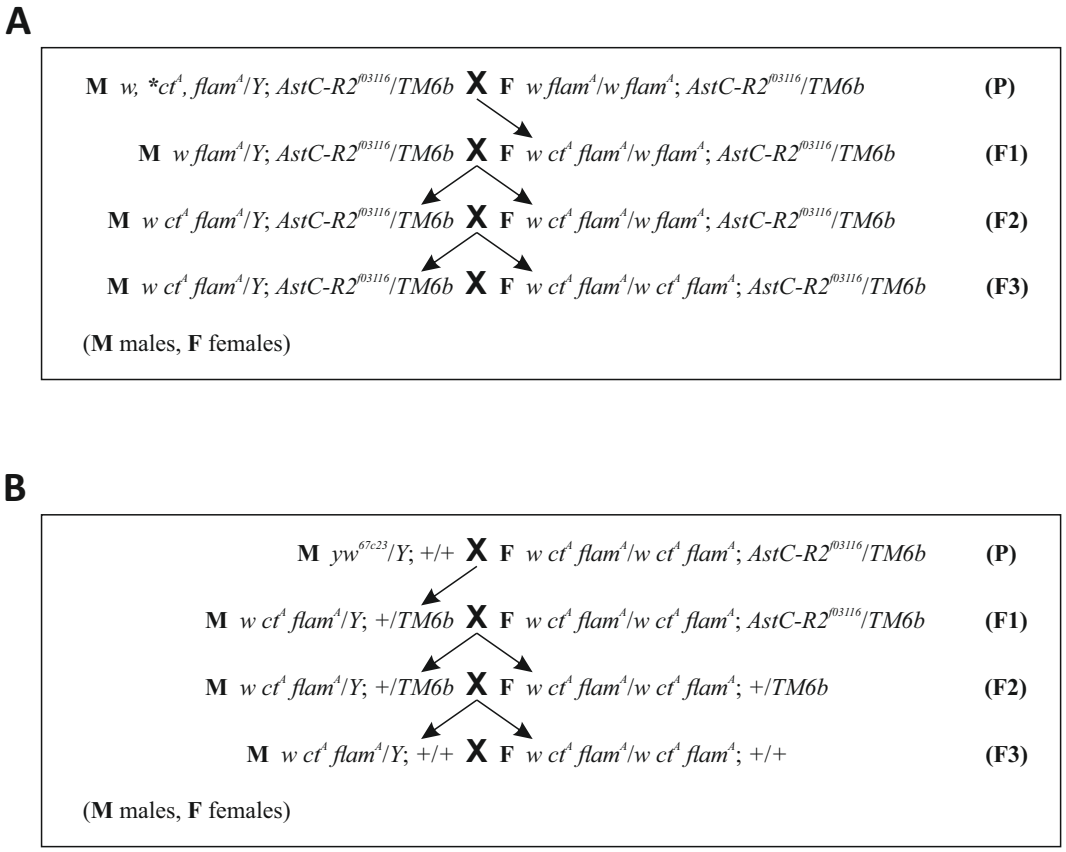

Figure S2. Genetic crosses followed to obtain the ct^A strains.

(A) To obtain a ct^A homozygous strain without changing the genetic background, the identified mutant male with the cut phenotype was crossed with sister females of the original strain (ct^+), and flies showing the cut phenotype were selected in the subsequent generations. Several females were used in each mating group and balancer X chromosomes were not used to avoid the fixation of possible second-site mutations. F1 and F2 females carried the ct^A mutation in heterozygosis and the X chromosomes could recombine. An identical cross scheme was followed to obtain the *w* f^A $flam^A$/*w* f^A $flam^A$; *AstC-R2*^{f03116}/*TM6b* strain. (B) To obtain a strain with the *w* ct^A $flam^A$ chromosome in a wild-type background, the third chromosome carrying the *AstC-R2*^{f03116} mutation was substituted with wild-type chromosomes from *y* w^{67c2} flies. Since the *AstC-R2*^{f03116} mutation is produced by the insertion of a *P*[w^+] element in the *AstC-R2* gene, males with w^- eyes were selected

in F1 and females and males with w^- eyes were selected in F2. Then, the *TM6b* balancer chromosome was eliminated from the strain selecting Tb^+ flies, since *TM6b* carries the Tb^1 mutant allele. An identical cross scheme was followed to obtain the control $w\ flam^A/w\ flam^A;$ +/+ strain and the $w\ f^A\ flam^A/w\ f^A\ flam^A;$ +/+ strain.

Figure S3

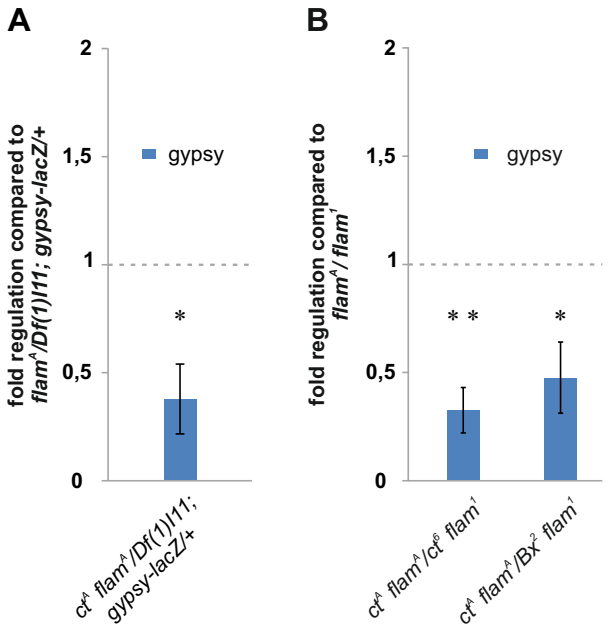

Fig S3. *Gypsy* expression in ovaries of flies carrying *gypsy* insertions in euchromatic regions is reduced in hemizygous and transheterozygous *flamenco* permissive backgrounds.

Quantitative RT-PCR analysis of *gypsy* expression in RNA isolated from adult ovaries ct^A $flam^A/Df(1)l11$; *gypsy-lacZ/+* (A) or ct^A $flam^A/ct^6$ $flam^1$ and ct^A $flam^A/Bx^2$ $flam^1$ (B). *Gypsy* expression is significantly reduced by the different *gypsy*-induced mutations. Shown are average levels (n=3) and error bars indicate standard deviation (*p<0.05, **p<0.01).

Figure S4

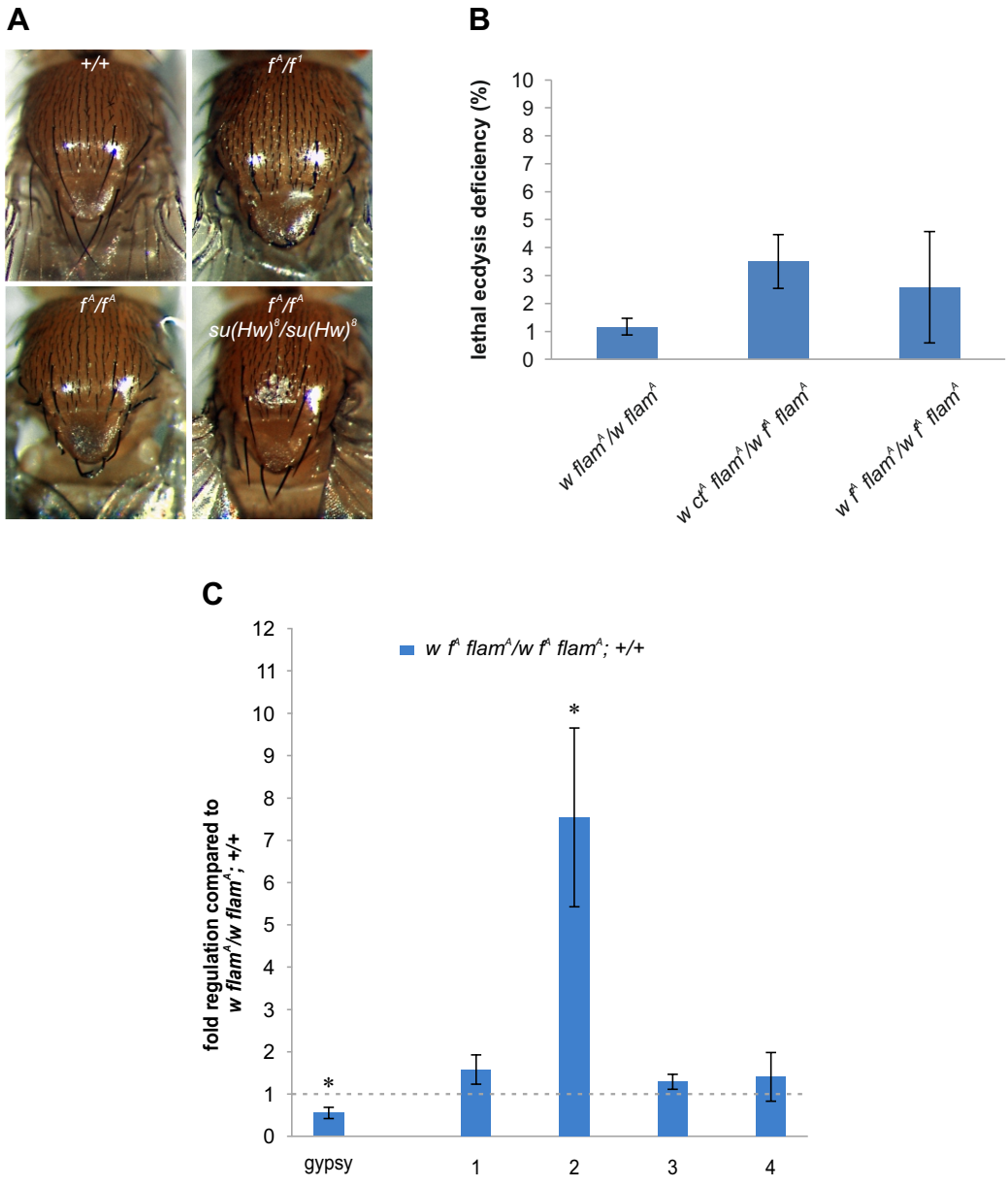

Fig S4. The *gypsy* insertion in the *forked* locus induces alterations in *flamenco* activity.

(A) Dorsal view of nota from $flam^A/flam^A$ (+/+); $f^A\ flam^A/f^A\ flam^A$ (f^A/f^A); $f^A\ flam^A/f^1$ (f^A/f^1) and $f^A\ flam^A/f^A\ flam^A$; $su(Hw)^8/su(Hw)^8$ (f^A/f^A; $su(Hw)^8/su(Hw)^8$). The new f^A mutation mapped on the X chromosome and did not complement the f^1 allele. *Gypsy* is known to induce mutations in the *forked* gene, producing phenotypes that are reversed by *Su(Hw)* mutations (Modolell et al., 1983). The phenotype induced by f^A in homozygous condition was suppressed by the homozygous $su(Hw)^8$ allele, confirming that the mutation was induced by *gypsy*. (B) Mortality during adult eclosion of $ct^A\ flam^A/f^A\ flam^A$; +/+ and $f^A\ flam^A/f^A\ flam^A$; +/+ females. (C) Quantitative RT-PCR analysis of *gypsy*

and specific *flamenco* regions (see Figure 3A) in RNA isolated from f^A $flam^A/f^A$ $flam^A$; +/+ 48 h old pupae. The *gypsy* insertion in the *forked* locus alters the expression of *gypsy*, and of the region amplified by primers 2, which does not map uniquely to the *flamenco* locus. Shown are average levels (n=3), and error bars indicate standard deviation (* $p<0.05$).

Figure S5

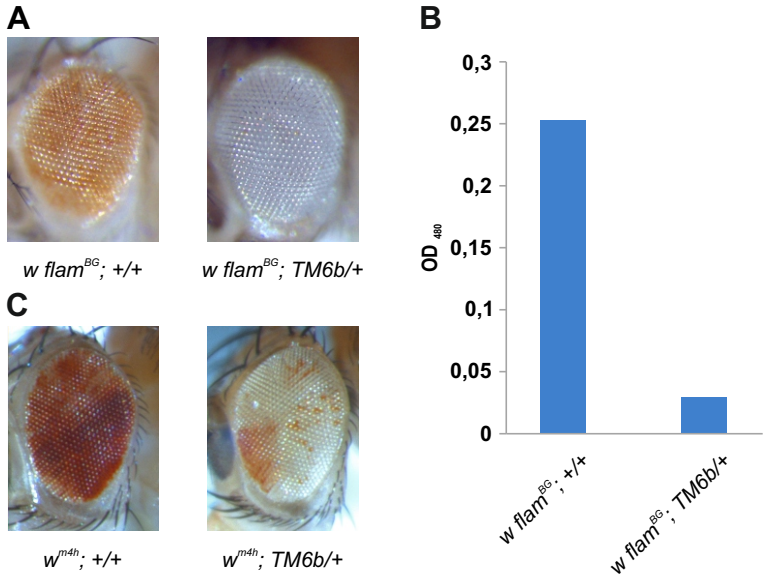

Fig S5. The *TM6b* chromosome enhances the silencing of a *mini-white* gene inserted inside the *flamenco* locus or pericentromeric chromatin.

(A) Adult eyes of wild type and *TM6b* flies carrying a *mini-white* transgene inserted in the *flamenco* locus ($flam^{BG}$). Presence of the *TM6b* balancer chromosome induces a silencing of *mini-white* expression resulting in decreased pigmentation. (B) Eye pigment levels extracted from heads of 6 days old males and measured at 480 nm. (C) Eyes of wild type and *TM6b* adult flies carrying the $In(1)w^{m4h}$ inversion in the X chromosome, which relocates the endogenous *white* gene in the pericentromeric chromatin region. The lower pigmentation in $In(1)w^{m4h}$; *TM6b/+* eye is an indication of an increase in heterochromatin formation at pericentromeric regions.

Figure S6

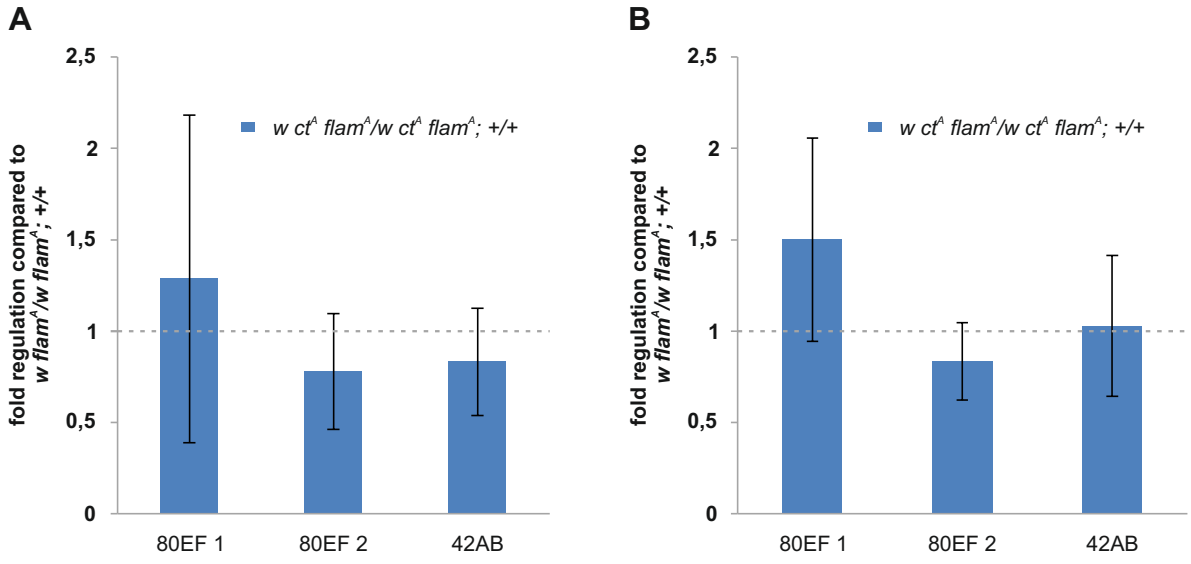

Fig S6. *80EF* and *42AB* dual-strand clusters do not seem to be regulated by *gypsy*.

Quantitative RT-PCR analysis of *gypsy* fragments located inside *80EF* and *42AB* clusters in RNA isolated from 0-24h female adult heads (A) and 3 days old adult ovaries (B). 80EF 1 and 80EF 2 primers were selected inside two different *gypsy* fragments that map uniquely to the *80EF* cluster. 42AB primers were selected inside a *gypsy* fragment that maps uniquely to the *42AB* clusters. Shown are average levels (n=3), and error bars indicate standard deviation.

Figure S7

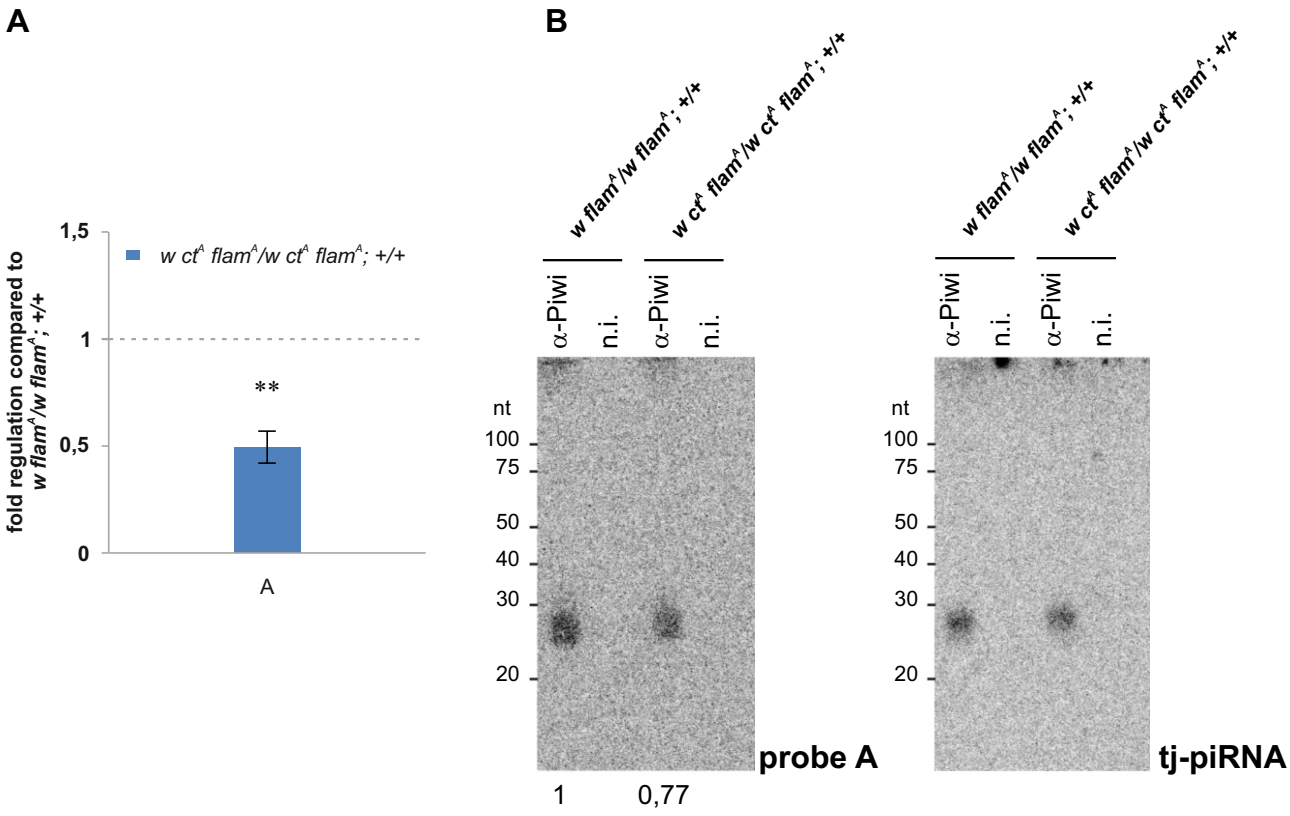

Fig S7. qRT-PCR and northern blot analysis of piRNAs from region A of the *flamenco* locus.

(A) piRNAs originating from region A of the *flamenco* locus (see Fig 3A) were detected by TaqMan specific probes in RNA isolated from 0-24 h old female heads. The experiment was performed on a single sample of about 100 heads using three technical replicates. Shown are average levels (n=3), and error bars indicate standard deviation (** $p<0.01$). (B) piRNAs in the Piwi complex immunopurified from ovaries of the indicated genotype were subjected to northern blot analysis (see also Supplemental Experimental Procedures). Probe A detects piRNAs originating from region A of the *flamenco* locus (see Fig 3A). *tj*-piRNA probe detects natural *tj*-piRNAs originating from the *tj*-3' UTR. Quantification of the bands intensity was performed using ImageJ software. Bands intensity is shown under the left figure as relative value after normalization to *tj*-piRNA. Control is set to 1. n.i: non-immune IgG (negative control).

Table S1. Primer sequences used in this study

	Forward Primer	Reverse Primer
***Gypsy* insertion**		
gyp1[1]	5' CTTAACTACCCTGTTTGTCG 3'	
cut1[1]		5' TAACATGTTCACCTCGTGGC 3'
cut2[1]		5' AGATCGAGGATCCCAGTATG 3'
cut4[1]		5' AATGGGATATGCCCGAGGAG 3'
cut6[1]		5' TGGTCGGAGATTTGTCCATG 3'
cut7[1]		5' CATGGACAAATCTCCGACCA 3'
cut8[1]		5' AACAAACAAATGGCCAGGGG 3'
cut9[1]		5' AACCAATTCGAAAGGCAGGG 3'
cut10[1]		5' CATATCGGTCCTACTTTCGG 3'
cutA	5' CCCAGCTTTCTTCGTTTTAAG 3'	5' GGGGCAGAGTTCTGCTAATG 3'
Primers for qPCR		
rp49[2, 3]	5' TCTGCATGAGCAGGACCTC 3'	5' ATCGGTTACGGATCGAACAA3'
gypsy[2, 3]	5' AGACGCTGCGACCATTCAC 3'	5' CGTGCTGCCTCCAGAATGAT 3'
flam1[2, 3]	5' TCAAAGCGATTCATTCCTCAG 3'	5' CCATTTGGCTATGAGGATCAG 3'
flam2[2, 3]	5' AGCGTGGGAATGCTGACTT 3'	5' CAGCAACACTTAGTTGCATTG 3'
flam3[2, 3]	5' CAGTGTACCCTCGATATATGCAC 3'	5' GGGCAAGTCGTTTAGACCAG 3'
flam4[2, 3]	5' TCGAATACGTCTGCAACAGG 3'	5'ATCAACACACCGCAAAATGA 3'
flam5[2]	5' CAGGCCCCCTATTGATTAGAT 3'	5' TGCTCGGGCTTTCTTAAAGT 3'
flam6[2]	5' GTATATCGGATGGCCGATTG 3'	5' GCACCGCAAATCATAACGTA 3'
80EF[1]	5' CGCCTCGATTAGTCATTTCA 3'	5' GCATTTTTCAGACGATGTGC 3'
80EF[2]	GGCTTAGGAATCGGGAAAT 3'	5' CGGTGCTGATCCTAATCCAT 3'
42AB[2]	ACCCCTGTGGTTTGGGTAAG 3'	5' ACACGAAGCTCTCCCTTTCA 3'
ZAM[2, 3]	5' TTCGCGTTAGGAGCCGTACT 3'	5' TTGACTTCGGTGTCGGAGAG 3'
Idefix[2, 3]	5' GAATGATTCCGCTCTAGTGG 3'	5' ATGCGGTCTCTTTCTTCTGC 3'
copia[3]	5' TCTGGTGCTAGTGACCATCT 3'	5' GCTTGGCCACTGCAATCTTA 3'
TART[3*]	5' GTACAACCAAAGTTGACGGG 3'	5' TGTATCGATATTTCGCGCTTTT 3'
cut-up[2, 3]	5' GGGCTGGGGAATAGAAAACT 3'	5' TTCATCCCAACTCTTAAAACGAA 3'
cut-down[2, 3]	5' ATCCCCCAAAAGGAAGTGAT 3'	5'AAATGCGCGAAATCTCTCAG 3'
cut-gyp up[2]	5' CCCAGCTTTCTTCGTTTTAAG 3'	5' CATAACTCTGGACCTATTGGAACTT 3'
cut-gyp dw[2]	5' GAGCGTTGAACCCAATTCTG 3'	5' AGGTGTGCCCGTAACTGACT 3'
rp49-RT[4]	5' GACAATCTCCTTGCGCTTCT 3'	
F-RT up[4]	5' CGGAGAGTTCGGCATTTG 3'	
R-RT up[4]		5' CAGCAAAAATAAAATACTTTTACTGTATTAAG 3'
F-RT dw[4]	5' AGTCAGTTACGGGCACACCT 3'	
R-RT dw[4]		5' CAAGGGGTTGCCTCTCATT 3'

Table S1. Primer sequences used in this study

Primers design was performed using Primer 3 (Untergasser et al., 2012).

[1] gyp1primer was used in combination with cut1-10 primers

[2] Primers used for qRT-PCR

[3] Primers used for ChIP

[4] RT-strand-specific primers

* (Moshkovich and Lei, 2010)

File S1. Supplemental Experimental Procedures

Identification of the *gypsy* insertion point in the *cut* locus by PCR

For identification of the *gypsy* insertion in the *cut* locus, DNA was extracted by crushing about 30 *w* ct^{A} $flam^{A}$*; TslC/TM6b* flies following the protocol described by E. Jay Rehm from the Berkeley *Drosophila* Genome Project. PCR reactions were performed using gyp1 and cut1-cut10 primers (Table). 25µl PCR reactions were processed using following conditions: 2 min at 95°C followed by 30 cycles of 30 sec at 95°C, 30 sec at 56°C, 2,5 min at 72°C; and 5 min at 72°C. Amplified products were visualized by ethidium bromide staining following electrophoresis on a 1% agarose gel. A DNA product of about 2500bp was obtained with the gyp1/cut10 couple of primers (Table S1). The DNA product was subsequently sequenced, and the *gypsy* insertion site was identified (Figure S2A). The g*ypsy* insertion region was amplified using the Expanded Long Templated PCR System (Roche) according to the manufacture's protocol, using cutA primers (Table S1). Amplified products were visualized by ethidium bromide staining following electrophoresis on 0,8% agarose gel.

Pigment assay

To measure eye pigmentation, 90 heads from 6-days old males of each genotype were dissected and homogenized in 100µl of methanol acidified with 0,1% HCl. The homogenates was placed at 4°C overnight, centrifuged and absorbance of supernatants were then measured at 480 nm.

Northern blotting

Immunoprecipitation from ovaries was performed as previously described (Nishida et al., 2009). Anti-Piwi antibody (P3G11) (Saito et al., 2006) was immobilized on Dynabeads protein G (Invitrogen). Ovaries lysates were prepared by crushing 50 ovaries in 200 µl of Binding buffer [30 mM HEPES-KOH (pH 7.3), 150 mM KOAc, 5 mM MgOAc, 5 mM DTT, 0.1% NP-40, 2 mg/ml

Pepstatin, 2 mg/ml Leupeptin, and 0.5% Aprotinin]. After centrifugation at 14000 g for 1min at 4°C, the supernatant was transferred to a new microcentrifuge tube and kept on ice. Pellet was then crushed again in 200 µl of Binding buffer. After centrifugation as above, the supernatant was combined with the first supernatant and kept on ice. These steps were repeated three times. The protein concentration of the lysates was measured and the same amount of proteins from both the control and mutant samples was adjusted to 2 mg/ml with Binding buffer. After incubation with 1 ml of ovary lysate, the beads were washed five times with Binding buffer. Total RNAs were isolated from the immunoprecipitates with phenol-chloroform and precipitated with ethanol. Northern blotting of piRNA was performed as described (Saito et al., 2006). The following probes were used for detection: 5'-GCCAGCCCTTTCTCGGTGAGCTCAATA-3' for piRNA from region A, and 5'-GGTAATGGGAATGCACTTCTCTTGAA-3' for *tj*-piRNA. Results were quantified with imageJ software (http://rsb.info.nih.gov/ij).

Supplemental References

KIM, J., SHEN, B., ROSEN, C. & DORSETT, D. 1996. The DNA-binding and enhancer-blocking domains of the Drosophila suppressor of Hairy-wing protein. *Mol Cell Biol,* 16**,** 3381-92.

MODOLELL, J., BENDER, W. & MESELSON, M. 1983. Drosophila melanogaster mutations suppressible by the suppressor of Hairy-wing are insertions of a 7.3-kilobase mobile element. *Proc Natl Acad Sci U S A,* 80**,** 1678-82.

MOSHKOVICH, N. & LEI, E. P. 2010. HP1 Recruitment in the Absence of Argonaute Proteins in Drosophila. *Plos Genetics,* 6.

NISHIDA, K. M., OKADA, T. N., KAWAMURA, T., MITUYAMA, T., KAWAMURA, Y., INAGAKI, S., HUANG, H. D., CHEN, D. H., KODAMA, T., SIOMI, H. & SIOMI, M. C. 2009. Functional involvement of Tudor and dPRMT5 in the piRNA processing pathway in Drosophila germlines. *Embo Journal,* 28**,** 3820-3831.

SAITO, K., NISHIDA, K. M., MORI, T., KAWAMURA, Y., MIYOSHI, K., NAGAMI, T., SIOMI, H. & SIOMI, M. C. 2006. Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the Drosophila genome. *Genes Dev,* 20**,** 2214-2222.

ST PIERRE, S. E., PONTING, L., STEFANCSIK, R. & MCQUILTON, P. 2014. FlyBase 102--advanced approaches to interrogating FlyBase. *Nucleic Acids Res,* 42**,** D780-8.

UNTERGASSER, A., CUTCUTACHE, I., KORESSAAR, T., YE, J., FAIRCLOTH, B. C., REMM, M. & ROZEN, S. G. 2012. Primer3--new capabilities and interfaces. *Nucleic Acids Res,* 40**,** e115.