

A Genome-wide RNAi Screen Draws a Genetic Framework for Transposon Control and Primary piRNA Biogenesis in *Drosophila*

Felix Muerdter,^{1,3,6} Paloma M. Guzzardo,^{1,6} Jesse Gillis,^{1,2} Yicheng Luo,¹ Yang Yu,¹ Caifu Chen,⁴ Richard Fekete,⁵ and Gregory J. Hannon^{1,*}

¹Watson School of Biological Sciences, Howard Hughes Medical Institute

²The Stanley Institute for Cognitive Genomics

Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA

³Zentrum für Molekularbiologie der Pflanzen, Entwicklungs-genetik, University of Tübingen, 72076 Tübingen, Germany

⁴Genetic Analysis R&D, Life Technologies Corporation, Foster City, CA 94404, USA

⁵Molecular Cell Biology R&D, Life Technologies Corporation, Austin, TX 78744, USA

⁶These authors contributed equally to this work

*Correspondence: hannon@cshl.edu

<http://dx.doi.org/10.1016/j.molcel.2013.04.006>

SUMMARY

A large fraction of our genome consists of mobile genetic elements. Governing transposons in germ cells is critically important, and failure to do so compromises genome integrity, leading to sterility. In animals, the piRNA pathway is the key to transposon constraint, yet the precise molecular details of how piRNAs are formed and how the pathway represses mobile elements remain poorly understood. In an effort to identify general requirements for transposon control and components of the piRNA pathway, we carried out a genome-wide RNAi screen in *Drosophila* ovarian somatic sheet cells. We identified and validated 87 genes necessary for transposon silencing. Among these were several piRNA biogenesis factors. We also found CG3893 (*asterix*) to be essential for transposon silencing, most likely by contributing to the effector step of transcriptional repression. Asterix loss leads to decreases in H3K9me3 marks on certain transposons but has no effect on piRNA levels.

INTRODUCTION

Transposable elements populate virtually every eukaryotic genome. Although their presence can impart some benefits, when not properly controlled, transposons can compromise the genomic integrity of their host and its offspring. Hence, in all higher animals there are control mechanisms in place that prevent wholesale transposon mobilization (Malone and Hannon, 2009).

In *Drosophila*, the principal pathway that protects the inheritable genome is comprised of a catalog of small RNAs that interact with an animal-specific clade of Argonaute family proteins: the Piwi proteins (Ishizu et al., 2012). This pool of Piwi-in-

teracting RNAs (piRNAs), which contains millions of distinct sequences bearing homology to transposable elements, functions as a molecular memory of transposon identity (Brennecke et al., 2007). Using their bound piRNAs as a guide, Piwi proteins (Piwi, Aubergine, and Argonaute 3) recognize and silence their targets. Failure of this pathway leads to defects in germline development and to sterility (Khurana and Theurkauf, 2010).

Genetic studies have uncovered a number of loci that are essential for the proper function of the piRNA pathway. Besides the core proteins of the Piwi clade, *flamenco* has long been implicated in transposon control. This discrete locus on the X chromosome of *Drosophila* was found to be a major determinant for silencing of the retroelement *gypsy* almost two decades ago, though its underlying molecular nature remained mysterious (Bucheton, 1995). Sequencing small RNAs bound to Piwi proteins and mapping these sequenced reads back to the genome revealed the true nature of the *flamenco* locus. Rather than being a protein coding gene, the *flamenco* transcript is the precursor to the majority of piRNAs expressed in the follicle cells of the ovary (Brennecke et al., 2007). This and other sites of abundant piRNA generation were termed piRNA clusters. What mechanisms mark *flamenco* and other cluster transcripts for processing into piRNAs remains largely unclear. Several studies have shed some light on this topic by identifying some of the protein factors that play a role in piRNA cluster transcription and transport. Rhino and Cutoff, as well as histone methylation marks deposited by Eggless (EGG), are necessary for cluster transcription (Klattenhoff et al., 2009; Pane et al., 2011; Rangan et al., 2011). In addition, UAP56, a helicase implicated in splicing and RNA export, was found to bind a germline piRNA cluster RNA and may escort the transcript from the nucleus to the nuage for processing (Zhang et al., 2012). Intriguingly, Rhino, Cutoff, and UAP56 all were reported to be specific to germline piRNA clusters, leaving factors involved in somatic cluster determination a mystery.

Mutagenesis screens for sterility phenotypes in *Drosophila* also highlighted factors that later were found to be elements of the piRNA pathway (Schübach and Wieschaus, 1989, 1991). Molecular analyses have begun to place these factors at specific

steps of the pathway, such as being required during the initiation and biogenesis phase or during the effector phase and silencing of transposons. However, very little is known about specific biochemical functions or enzymatic activities of the proteins that act at each step. One major insight came from the discovery of a trimming activity in insect cell lysates that shortens the 3' end of putative piRNA precursors to their mature length (Kawaoka et al., 2011). However, the protein responsible for this activity remains unknown. Biochemical and structural studies of Zucchini (ZUC), which was previously implicated in the piRNA pathway, suggest it as a promising candidate for the nuclease that creates the 5' ends of primary piRNAs (Ipsaro et al., 2012; Nishimasu et al., 2012). Whether Zucchini and a trimming enzyme together comprise the complete primary biogenesis machinery or other endo- or exonucleolytic cleavage events create intermediates that are further matured remains unknown.

Another enigmatic aspect of the pathway is precisely how Piwi-piRNA complexes silence their targets. In the case of somatic cells of the ovary, it has become clear that control of transposons occurs at the transcriptional level through Piwi-directed deposition of epigenetic marks. Recently, three conclusive studies showed that, upon Piwi depletion, transposons engage in active transcription and show a depletion of H3K9 trimethyl (H3K9me3) (Le Thomas et al., 2013; Rozhkov et al., 2013; Sienski et al., 2012). In one of these studies, the authors place Maelstrom (MAEL) at the effector step of transcriptional repression (Sienski et al., 2012). Interestingly, loss of *mael* derepresses transposons without preventing H3K9me3 deposition, indicating that this modification may not be the definitive silencing mark. What the final silencing mark may be, and which proteins are responsible for establishing repressive chromatin marks over transposons, has yet to be determined.

Much of what has been learned of the piRNA pathway that operates in follicle cells relied on the use of a cultured ovarian somatic sheet (OSS) cell line (Niki et al., 2006). This cell line expresses microRNAs (miRNAs), endogenous small interfering RNAs (endo-siRNAs), and piRNAs (Lau et al., 2009; Saito et al., 2009). With an active siRNA and primary piRNA pathway in place, genetic requirements for primary piRNA biogenesis and transposon silencing can be investigated.

Here, we describe a genome-wide screen that builds a foundation for addressing many open questions relevant to piRNA biogenesis and effector functions. By individually assaying more than 41,000 long, double-stranded RNAs (dsRNAs), targeting every annotated gene in the *Drosophila* genome, and examining their effect on transposon expression levels, we describe a comprehensive genetic framework for transposon control. We reveal piRNA biogenesis factors and place proteins, which to our knowledge have not been implicated in the piRNA pathway, at the effector step.

RESULTS

An RNAi Screen for Elements of the Somatic piRNA Pathway

In order to assay transposon derepression upon knockdown (KD) of any given target gene, we established a sensitive assay for *gypsy* messenger RNA (mRNA) levels. Based on quantitative

PCR (qPCR) with hydrolysis probes (TaqMan), this assay specifically detects the spliced subgenomic transcript of the retrotransposon (Figure 1A). The expression of this transcript is known to be highly sensitive to disruption of the piRNA pathway, even more so than its unspliced counterpart (Péligsson et al., 1994).

Knockdown of target genes was accomplished by transfecting dsRNAs from two independent genome-wide libraries with a total of 41,342 dsRNAs. The average count of dsRNAs per gene was 2.28, targeting 13,914 genes with valid IDs in FlyBase (McQuilton et al., 2012). Additionally, the two libraries contained 1,045 negative controls, 2,097 dsRNAs without an annotated target, and 2,301 dsRNAs targeting the Heidelberg collection of predicted genes.

We transfected OSS cells with individual dsRNAs in 48-well plates, lysed the cells 5 days later, and used the lysate for reverse transcription (Figure 1A). The qPCR results were normalized to their respective plate using Z scores (Ramadan et al., 2007). As a secondary metric, we calculated the fold change in relation to the median. The knockdown of *piwi* in this setting led to a *gypsy* mRNA signal that was detectable by qPCR much earlier than the average of the plate. In four biological replicates of *piwi* knockdown, the average normalized signal for *gypsy* was almost 5 SDs away from the median of its plate (Figure S1A). Hence, our assay for transposon derepression is both sensitive and robust, at least given that the RNA interference (RNAi) trigger is of good quality. When comparing several independent dsRNAs, we saw that there was considerable variance in this respect. dsRNAs against known components of the pathway, such as *armitage* (*armi*), led to consistent derepression of *gypsy*; however, levels varied from 25- to 70-fold (Figure S1B). Since we assayed several dsRNAs against each annotated gene, we felt confident that the majority of pathway components could be identified.

Out of 41,342 tested dsRNAs, 33,780 met our criteria for inclusion in our analysis; of these, 320 dsRNAs met the criteria for primary hit selection (Figure 1B and Table S1). Included in this list were 18 dsRNAs without annotated targets, which were disregarded. All genes that were previously implicated in *gypsy* control were outliers in the primary screen (Figure S1B). Knockdown of pathway components such as *capsuleen*, *hen1*, *egg*, or *squash* was not expected to cause strong *gypsy* derepression based on existing literature and indeed did not cause derepression of *gypsy* in our assay (Olivieri et al., 2010; Rangan et al., 2011). After choosing Z score and fold change cutoffs for hit selection based on 217 green fluorescent protein (GFP) negative controls, only 3 out of 645 (0.5%) additional negative controls scored as weak hits (Figure S1C).

To ask whether genes that affect transposon control show preference toward specific annotation groups, we performed functional enrichment analysis on our primary data set. After multiple testing correction, 215 functions were associated with transposon silencing (corrected $p < 0.05$), many with strong potential relevance and very significant enrichment (the top 20 functions have a corrected $p < 1 \times 10^{-6}$; Table S2). Among the most significant, we find expected cellular components like the Yb body, but also more surprising functions like regulation of growth of symbiont in host. While genes implicated in the piRNA pathway drive several of these enrichments, all scoring highly in the primary screen, our candidate hits intersect with these in some of the

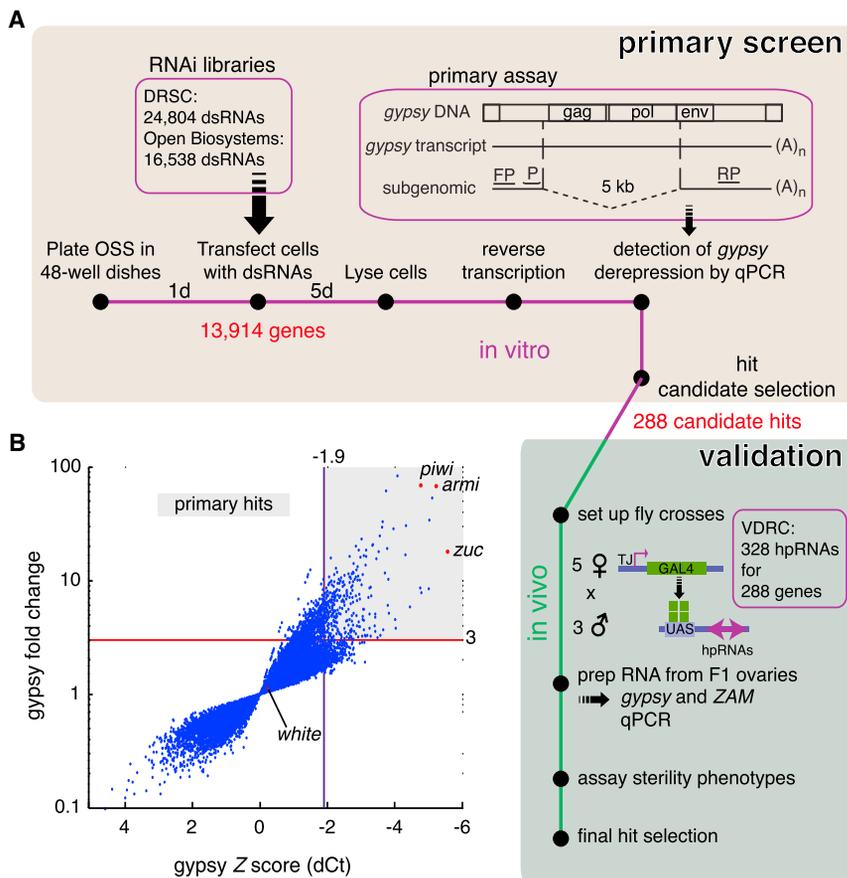


Figure 1. A Genome-wide RNAi Screen for piRNA Pathway Components Acting in the Somatic Compartment of *Drosophila* Ovaries

(A) A workflow of the primary RNAi screen in ovarian somatic sheet (OSS) cells and validation of primary hit candidates *in vivo* is shown. Each gene in the *Drosophila* genome was knocked down with one or more dsRNAs. At 5 days after transfection, cells were tested for increased levels of the *gypsy* retrotransposon. The primers and the hydrolysis probe used for the qPCR are shown (FP, forward primer; P, hydrolysis probe; RP, reverse primer). The dashed line indicates the ~5 kb segment not present in the subgenomic transcript. We further tested 288 genes *in vivo* using the Gal4/UAS system to drive hairpin RNAs (hpRNAs) within the *traffic jam* (TJ) expression domain.

(B) All transfected wells were assayed for levels of *gypsy* and one reference gene for normalization. Levels of *gypsy* expression are displayed as Z scores and fold change. The cutoffs for both Z score (<-1.9) and fold change (>3) are indicated as red lines. The shaded area shows the selection of primary hit candidates. Three positive controls (*piwi*, *armi*, *zuc*) and one negative control (*white*) are marked as red dots. Only wells that passed the filter for primary data point selection are shown. For all primary data points see Table S1. See also Figure S1.

enriched functions. For example, under dorsal appendage formation, *smt3* (SUMO) joins *armi* and *zuc* (Nie et al., 2009). Further work will be necessary to determine to what extent these unexpected intersections relate to biologically relevant connections.

Next, we compared protein interaction data to the full-ranked fold change list. For every gene in the genome, the degree to which that gene's interaction partners scored highly in the fold change list was measured (as receiver operating characteristics [ROC]). Of the top 20 genes with interaction partners significantly elevated in the fold change list, 18 are annotated as belonging to the proteasomal complex (which had 65 genes in total). This observation was remarkably significant ($p < 1 \times 10^{-40}$ after multiple testing correction), which may be due partially to the correlated interaction profiles of the proteasome complex. The two remaining genes not belonging to the proteasomal complex were *bx42*, a homolog of mammalian Skip (SKI-interacting protein), a protein implicated in splicing, and *calypso*, a histone 2A deubiquitinase (Makarov et al., 2002; Scheuermann et al., 2010). Both genes are highly expressed in ovaries, according to the modENCODE tissue expression data. Whether their interaction with particularly high scoring genes is indicative of any regulatory function remains to be tested.

In Vivo Validation of Primary Hits

We obtained 328 fly lines from the Vienna *Drosophila* RNAi Center (VDRC) containing inducible hairpin RNAs (hpRNAs) against

(*traffic jam*), these hpRNAs can effectively knock down any given target gene within the same expression domain (Tanentzapf et al., 2007). Using hpRNAs against *aub* as a negative control, we observed highly significant changes in *gypsy* expression by qPCR when knocking down *armi* (Figure S2A). We confirmed the effect on the piRNA pathway by measuring two additional transposons, *ZAM* and *gypsy3*. All known pathway components that were identified as primary hits were validated using this approach except for Piwi, which showed developmental defects upon knockdown (Figure S2B). Harnessing this *in vivo* system, we validated 87 out of the 288 primary hit candidates (Figure 2A and Table S3). In order to be considered as validated, knockdown of the target gene had to result in upregulation of *gypsy* or *ZAM* by at least 2-fold. For crosses with male flies originating from the GD library (first generation) from VDRC, we used *ZAM* as a metric; for the KK library (second generation) we measured *gypsy* levels. This decision was based on the finding that negative controls from the GD library already showed higher basal levels of *gypsy* subgenomic transcript when compared to the KK library negative controls (Figures S3A–S3C).

Out of the 288 candidate hits, knockdown of 52 genes, including Piwi, led to such severe developmental defects that dissection of ovaries and confirmation of the initial screen result was not possible. However, several arguments can be made that this category harbors a substantial number of true hits. First, the genes in this category had an average *gypsy* fold change of 5.8 in

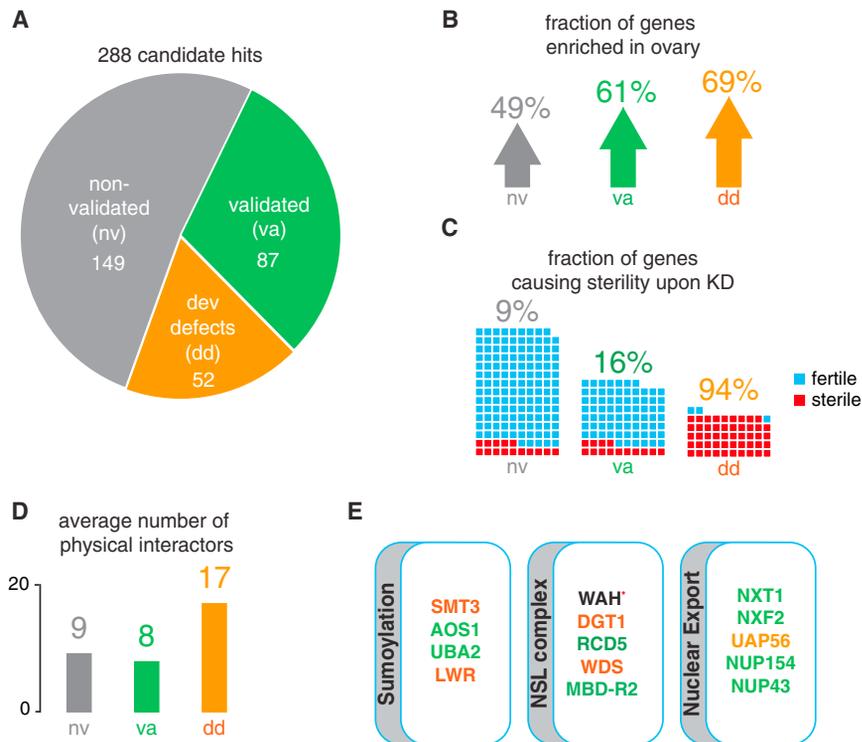


Figure 2. Primary Candidates Were Validated In Vivo

(A) The number of hit candidates that validated (va) or did not validate (nv) in vivo is shown. Genes that caused severe developmental defects upon knockdown and therefore could not be assayed are also indicated (dd). A full list of validated fly lines and corresponding transposon derepression information is available in Table S3.

(B) Validated hits are preferentially expressed in ovaries. The percentage of genes that are enriched in ovaries compared to whole fly is shown for the three categories. These data are based on mRNA signals on Affymetrix expression arrays available from FlyAtlas (Chintapalli et al., 2007).

(C) The fraction of genes causing sterility upon knockdown is shown. Each small box represents one gene, with blue and red indicating if flies were fertile or sterile, respectively, upon knockdown.

(D) The degree to which a gene may represent a node in a network in each of the classifications was measured by the number of physical interactors. Interaction data from BioGRID was used for this analysis (Breitkreutz et al. 2008).

(E) Components of the *Drosophila* sumoylation pathway, the nonspecific lethal complex, and proteins involved in nuclear export are primary hits that validate in vivo. WAH could not be validated in vivo because no RNAi fly was available at the time of submission (red asterisk). The text coloring of each gene indicates the result of the validation screen and is consistent with the categories in (A). See also Figure S2.

the primary data set, as compared to 3.3 for the primary hits that failed secondary validation. This average fold change was even higher than the validated subset (5.1-fold). Since primary fold changes as well as Z scores are a function of precision (the likelihood of a primary hit to be validated), this is indicative of the biological relevance of these hits (Figure S2C). Second, while the nonvalidated genes had a representation count in the dsRNA libraries of 2.84, the developmental defect set had a count of 2.42, which is significantly different for the two sets ($p \approx 0.0035$, Wilcoxon test). Thus, the genes of the developmental defect category were disadvantaged according to their annotation class with respect to their possibility to be a primary hit by chance in comparison to the candidates that failed validation, yet they had a much higher average fold change.

Both the validated and developmental defect sets were significantly enriched for genes with higher expression levels in ovaries as compared to whole fly (Figure 2B). While knockdown of genes within the nonvalidated category only led to sterility of 9% of the crosses, the fraction was 16% for the validated and 94% for the developmental defect set (Figure 2C). The extreme phenotypes we observe in the developmental set could imply more generic functions for these genes than simply action in the piRNA pathway. Indeed, around 90% have an average of 17 physical interactors, which is significantly higher than the other categories (Figure 2D).

When ranked by their fold changes in the validation screen, the majority of the somatic piRNA pathway components were among the strongest hits (Table 1). However, genes that, to our

knowledge, have not been implicated in the pathway, scored highly as well. *nxt1*, a nuclear export factor, ranks first with *gypsy* levels almost 2,500-fold higher than the negative control (Figure 2E) (Herold et al., 2001). In addition, depletion of NXT1 also led to sterility. Knockdown of the RNA helicase *uap56*, which acts in the same export pathway (Herold et al., 2003), showed derepression of *gypsy* of up to 8-fold in the primary screen. First implicated in splicing, *uap56* was recently shown to be involved in transport of the primary piRNA transcript of dual-stranded clusters to the nuclear pore (Zhang et al., 2012). The knockdown of *uap56* in follicle cells affected germline development to such an extent that in vivo verification of the primary screen results was not possible. We were able to validate another mRNA export factor, *nxf2*, which interacts with *nxt1* (Table S3) (Herold et al., 2001).

NXT1 was previously reported to affect interactions with the nuclear pore complex as well (Lévesque et al., 2001). Hence, the presence of several nuclear pore components within the top 20 hits is not surprising: both *nup154* and *nup43* knockdowns caused similar derepression of *gypsy* (Table 1). Additionally, NUP154-deficient flies were sterile in our assay.

Another two genes ranking among the top 10 are uncharacterized as of yet: CG3893 and CG2183. CG3893 shows homology to mammalian GTSF1. Even though no direct link to the piRNA pathway has been shown, this germline-specific factor also seems to be indispensable for transposon control in mice (Yoshimura et al., 2009). CG2183 is predicted to be a homolog of GASZ. This protein was previously implicated in the piRNA

Table 1. Top 20 Validated Hits

FlyBase ID	Symbol	Primary Screen Average Fold Change	Validation Screen Fold Change			Fertility	Comments
			<i>Gypsy</i>	<i>Zam</i>	<i>Gypsy3</i>		
FBgn0028411	<i>nxt1</i>	2	2452	3566	41	–	Involved in mRNA export from nucleus
FBgn0000928	<i>fs(1)Yb</i>	11	96	700	335	+	piRNA pathway component
FBgn0041164	<i>armi</i>	48	197	846	112	+	piRNA pathway component
FBgn0261266	<i>zuc</i>	19	809	549	9	+	piRNA pathway component
FBgn0263143	<i>vret</i>	4	74	315	22	+	piRNA pathway component
FBgn0036826	<i>CG3893 (asterix)</i>	42	80	207	10	+	Contains two CHHC zinc fingers
FBgn0016034	<i>mael</i>	3	159	452	16	+	piRNA pathway component
FBgn0033273	<i>CG2183</i>	4	173	158	11	+	Fly homolog of GASZ
FBgn0029800	<i>lin-52</i>	5	153	153	8	+	dREAM complex subunit
FBgn0038016	<i>MBD-R2</i>	16	85	48	4	+	NSL complex subunit
FBgn0029113	<i>uba2</i>	3	84	12	8	+	Sumoylation E1 ligase
FBgn0034617	<i>CG9754</i>	2	26	61	14	+	No conserved domains
FBgn0027499	<i>wde</i>	3	40	120	12	+	Cofactor of Eggless
FBgn0021761	<i>nup154</i>	6	30	186	3	–	Structural constituent of nuclear pore
FBgn0003612	<i>Su(var)2-10</i>	2	9	20	4	–	dPIAS, putative SUMO E3 ligase
FBgn0001624	<i>dlg1</i>	2	16	2	1	+	Guanylate kinase
FBgn0003401	<i>shu</i>	7	14	416	1	+	piRNA pathway component
FBgn0038739	<i>CG4686</i>	4	13	1	1	+	Part of ribokinase/pfkB and DUF423 superfamilies
FBgn0038609	<i>nup43</i>	3	12	3	1	+	WD40-repeat-containing domain
FBgn0038925	<i>cchl</i>	0	12	1	2	–	Cytochrome c heme lyase

pathway in mice (Ma et al., 2009) and has now been validated as a piRNA pathway component in flies (Czech et al., 2013).

smt3 (SUMO), which was one of the highest scoring genes in the primary screen, could not be validated in vivo because of developmental defects that occurred upon knockdown (Talamillo et al., 2008). However, the depletion of its E1 activating enzymes AOS1 and UBA2 caused consistent transposon derepression in the validation screen (Table 1 and Figure 2E). Knockdown of the E2 conjugating enzyme *lesswright* also caused developmental defects and could not be validated.

Another notable validated hit is *windei* (*wde*), which was previously reported as a cofactor of EGG in H3K9 trimethylation (Koch et al., 2009). While EGG depletion had no effect on *gypsy* expression in our assay, knockdown of *wde* caused derepression of *gypsy*, although to a lower extent than *ZAM* (Table S3).

Additionally, two genes involved in transcriptional regulation were identified. MBD-R2 is part of the nonspecific lethal complex (Figure 2E) (Raja et al., 2010). All members of this complex except for *rcd1* scored in the primary screen. *lin-52*, which is part of the dREAM transcriptional regulator complex, also scored highly with *gypsy* and *ZAM* (Lewis et al., 2004).

Characterization of Newly Described piRNA Pathway Components

In order to place some of the validated hits at particular steps within the piRNA pathway, we constructed RNA sequencing (RNA-seq) libraries from ovaries of biological replicates of *tj-Gal4*-driven hpRNA crosses. Mapping RNA-seq reads to transposon consensus sequences revealed the same levels of

gypsy and *ZAM* derepression observed by qPCR (Figure 3A). When analyzed on a global scale, mainly transposons dominant within the somatic compartment of the ovary reacted significantly to the respective knockdown of the target gene in a *tj-Gal4*-dependent manner (false discovery rate [FDR] < 0.05) (Malone et al., 2009). Transposons like *HeT-A*, *roo*, or *Rt1b*, which were previously shown to be germline dominant, did not change expression levels (Figure 3A) (Malone et al., 2009). The patterns of derepression that we observed in knockdowns of known components of the pathway (*armi* and *mael*) remarkably resembled those observed in *CG3893* and *wde* knockdown.

We observed a similar behavior of *CG3893* and known piRNA pathway components in their impacts on the expression of protein coding genes. While proteins like *NXT1* or *UBA2*, with potentially more general functions beyond the piRNA pathway, impacted the expression of hundreds to thousands of genes, this was not the case for *ARMI*, *MAEL*, or *CG3893*, which show effects that are much more restricted to transposon transcripts (Figure 3B). We interpreted these results as an indication that *CG3893* might act specifically within the piRNA pathway.

We further annotated validated hits based upon their impacts on piRNA levels. Interestingly, we saw a substantive drop in the abundance of piRNAs uniquely mapping to the soma-dominant *flamenco* piRNA cluster in the *nxt1*, *uba2*, and *wde* knockdowns (Figure 4A). To avoid skewing this result based on normalization to a piRNA-producing locus, which theoretically should not change in soma-specific knockdowns (i.e., *cluster 1/42AB*), we examined the rankings within each library of piRNA clusters based on the overall abundance of corresponding piRNAs. Given

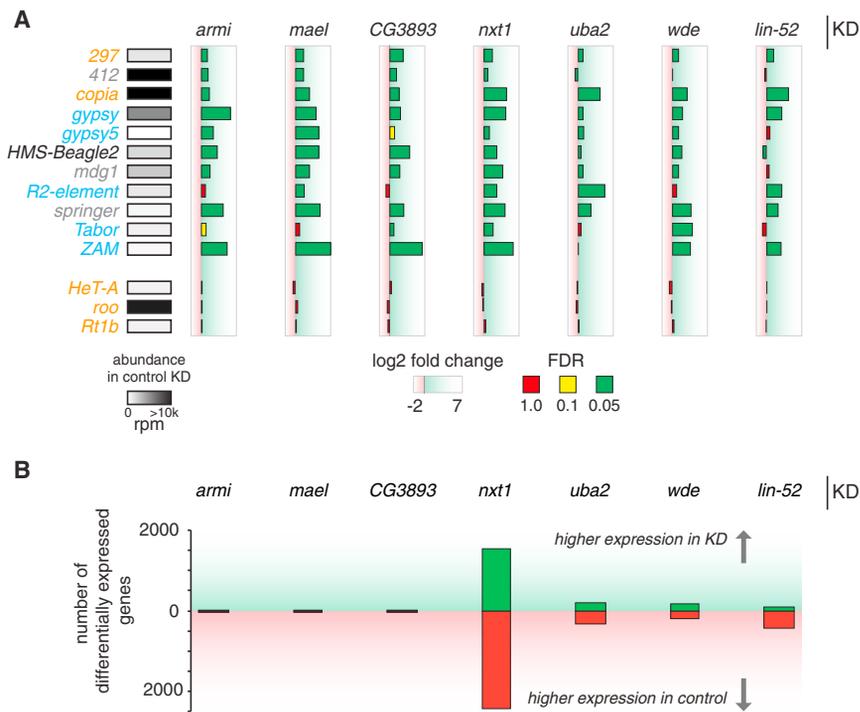


Figure 3. RNA-Seq Shows Changes in Gene and Transposon Expression upon Knockdown of Top Candidates In Vivo

(A) A subset of somatically expressed transposons is derepressed in the indicated KD. The classification of transposons according to Malone et al. (2009) is indicated in orange (germline dominant), gray (intermediate), blue (soma dominant), and black (unclassified). The absolute abundance of reads in control knockdown mapping to each transposon is shown in shades of gray. The log₂ fold change of each target gene versus a negative control (*aub*) is shown. Color of the bars represents the significance of these fold changes and is indicated as an adjusted p value (FDR). Green indicates highly significant differences ($p \leq 0.05$), yellow indicates moderately significant changes ($0.05 < p \leq 0.1$), and red indicates nonsignificant changes ($0.1 < p \leq 1$) based on two biological replicates. Each knockdown is normalized to *aub* knockdown controls from their corresponding library (GD or KK). For differences in transposon abundance levels between both *aub* controls see Figure S3.

(B) The number of genes differentially expressed (adjusted $p < 0.05$) in each knockdown with respect to the control is shown. Green bars indicate the number of genes that have higher expression levels in the knockdown fly line, while red bars designate the number of genes with higher levels in the *aub* negative control.

that we only knock down each gene in the somatic compartment, only clusters within that expression domain (i.e., *flamenco*) should change their ranking. Validating this approach, we observed that in the *armi* knockdown, *flamenco* was the tenth most abundant cluster while it was the third most abundant in total RNA libraries from negative-control ovaries. Conversely, *42AB* and *X-upstream* (cluster 20A) remained at the top of the list in all tested knockdowns (Figure 4B). The two genes that mirrored *armi* are *nxt1* and *uba2*. *wde* and *lin-52* showed changes in *flamenco* piRNAs that altered its ranking, but not to the same extent. In none of the knockdowns did the length profiles of the remaining piRNAs from *flamenco* change (Figure 4C). When compared to their negative controls, piRNA levels did not change in the *mael* and *CG3893* knockdown. The same conclusions could be drawn when mapping to transposons consensus sequences: antisense populations of piRNAs with homology to soma-dominant transposons showed severe reductions in the *nxt1*, *uba2*, and *wde* knockdowns, which resembled patterns seen for the biogenesis factor *armi* (Figure 4D). Depletion of *mael*, *lin-52*, or *CG3893* did not show the same effect. Intriguingly, in the case of *lin-52* this did not coincide with the effects seen for *flamenco* mappers. None of the assayed knockdowns showed any changes in mature miRNA levels, indicating that the observed effects were specific to the piRNA pathway and not a general trend of all small RNA populations (Figure S4).

CG3893 Is Indispensable for Transposon Silencing in the Germline

piwi and *mael* have recently been shown to silence transposons in the somatic compartment of the ovary through effects on

transposon transcription (transcriptional gene silencing or TGS) (Le Thomas et al., 2013; Rozhkov et al., 2013; Sienski et al., 2012). *CG3893*, displaying patterns in global transposon derepression similar to *mael* and unaffected mature piRNA populations, appeared to be a promising candidate for a pathway component acting at the TGS effector step.

CG3893, a ~20 kDa protein, is a member of a family of proteins with unknown function (UPF0224), characterized by the presence of highly conserved zinc finger domains (Figure 5A). All five proteins of this family are weakly expressed in OSS cells and show germline-specific moderate to high expression in the ovary (Figure 5B, modENCODE tissue expression data [Graveley et al., 2011]). Out of all five members of the family, only *CG3893* showed strong effects on transposon mRNA abundance when knocked down in OSS cells (Figure 5C).

In order to obtain an additional model for a loss of function of *CG3893*, we searched for available transposon insertion lines. We investigated two available lines (204406, Kyoto *Drosophila* Genetic Resource Center [DGRC]; 22464, Bloomington *Drosophila* Stock Center at Indiana University). Line 204406 has a P element insertion into the first exon of *CG3893*, disrupting its N-terminal CHHC zinc finger domain (Figure S5A). Consistent with the insertion site, we identified truncated *CG3893* mRNAs by RNA-seq in libraries from homozygous animals. The levels of *CG3893* mRNA expression were also reduced in animals homozygous for this mutation when compared to heterozygous siblings. The results obtained by RNA-seq were confirmed by qPCR (Figure S5B). The second line (22462) has a P element insertion upstream of the first exon in either the promoter or the 5' untranslated region

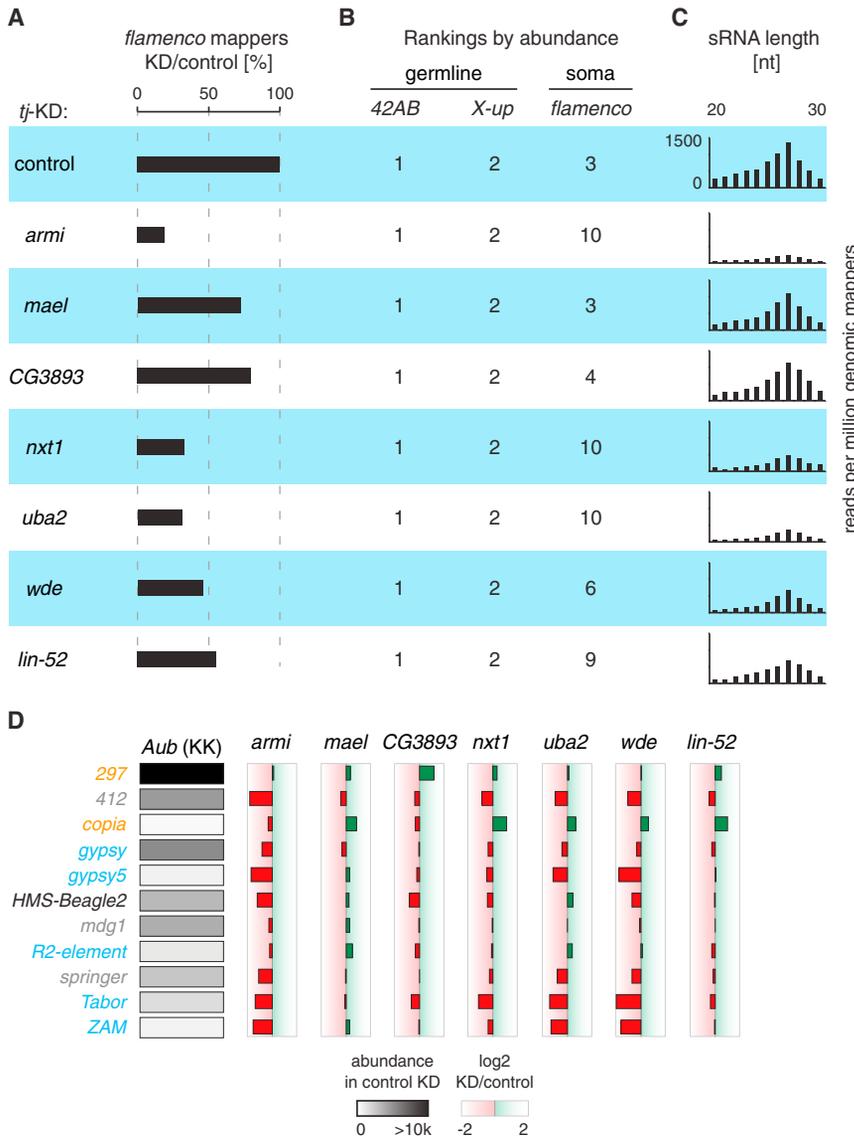


Figure 4. Biogenesis of Small RNAs from Somatic Clusters and Transposons Is Affected in Knockdowns of a Subset of Top Candidate Genes

(A) Percentages of total unique mappers (sense species, >23 nt) to *flamenco* in each knockdown (as indicated) in relation to the control knockdown are shown.

(B) The internal rankings for three representative piRNA clusters based on their representation in piRNA populations are displayed. Expression bias toward either domain (soma or germline) is indicated. Cluster definitions are in concordance with Brennecke et al. (2007).

(C) The size profiles of piRNAs mapping in sense orientation to *flamenco* in each knockdown (as in [A]) are plotted as total read count per million genomic mappers. As a control, we show that levels of microRNAs do not change in knockdowns versus negative control (Figure S4).

(D) piRNAs mapping to a subset of somatically expressed transposons are reduced when gene expression of a subset of top hits is disrupted. The classification of transposons according to Malone et al. (2009) is indicated in orange (germline dominant), gray (intermediate), blue (soma dominant), and black (unclassified). The absolute abundance of antisense piRNAs in an *aub* control mapping to each transposon is shown in shades of gray. The log2 fold change of each target gene versus a negative control is shown. See also Figure S4.

ported to be mainly cytoplasmic in adult testes (Yoshimura et al., 2007). However, when overexpressed in OSS cells, GFP fusion proteins of CG3893 colocalized with Piwi in the nucleus (Figure 5E).

Our results to this point demonstrated the involvement of CG3893 in the somatic compartment of the ovary. In order to investigate its role in all tissues of the female germline, we generated RNA-seq

(UTR) of CG3893. Our RNA-seq data in control flies indicated a slightly extended CG3893 transcript when compared to the gene model presented in FlyBase. Even though this insertion is not in any coding sequence, the observed phenotypes were severe: homozygous females were completely sterile and characterized by a complete absence of ovarian structures. This correlated with undetectable levels of CG3893 transcript when assayed by qPCR, indicating a complete loss of function (Figure S5B). The phenotypes observed in females homozygous for the first insertion (204406) were milder, with ovaries developing to a rudimentary stage (Figure 5D). Nevertheless, this potentially hypomorphic mutation caused females to be sterile, demonstrating the negative impact of the insertion on CG3893 function.

According to our current model of transposon silencing as a nuclear phenomenon, effectors at this step are expected to be nuclear as well. The mouse homolog of CG3893, Gtsf1, is re-

and small RNA libraries from females heterozygous and homozygous for the exonic P element insertion. RNA-seq revealed a remarkable change in global transposon expression. Almost all classes of annotated transposons populating the *Drosophila* genome, except for the P element itself, showed upregulation in homozygous females when compared to their heterozygous sisters (Figure 5F). This derepression effect was equally strong for germline- and soma-dominant transposon classes. Yet, when mapping antisense piRNA reads to transposon consensus sequences, we see no change in the homozygous animals (Figure 5G).

Three recent publications demonstrate not only that piRNA-mediated transposon silencing is a nuclear phenomenon occurring through TGS, but also that it acts through deposition of silencing H3K9 trimethyl marks on active copies of transposons (Le Thomas et al., 2013; Rozhkov et al., 2013; Sienski et al., 2012). Given its potential involvement in this

step, we investigated the effects of CG3893 disruption on this histone mark by performing chromatin immunoprecipitation sequencing (ChIP-seq) analysis for H3K9me3 in ovaries from heterozygous and homozygous females. Strikingly, homozygous females showed a marked reduction of H3K9me3 levels over a subset of peaks identified in heterozygous animals. These differential peaks overlapped with full-length insertions of both germline- and soma-dominant transposons while neighboring peaks over transposon fragments did not change (Figure 5H; fragments correspond to *accord2* and *diver*, center panel). We also observed the same changes when we mapped to consensus sequences of the corresponding transposon families, thereby aggregating signal from all genomic insertions (Figure 5H). We found that 75% of all peaks showed lower levels of H3K9me3 deposition in homozygous flies, with 28 peaks exhibiting a reduction to less than 50% of the read density compared to heterozygous siblings (Figure S5C). These peaks corresponded mostly to retrotransposons (LTR, LINE), while none of the peaks over DNA elements were affected. However, the loss of H3K9me3 in homozygous mutant animals was not exclusive to transposable elements, as a limited number of protein coding genes, such as CG8964, showed similar patterns (Figure S5D). Our RNA-seq data were in accordance with this finding: the CG8964 transcript was increased ~26-fold in homozygote versus heterozygote flies. Whether this effect is due to effects on a transposon insertion within its genomic locus or within another gene that controls CG8964 expression, or whether this represents a piRNA-independent function of CG3893, remains to be seen. Because of its small size, yet powerful role in transposon silencing, we name CG3893 *asterix* (*arx*).

DISCUSSION

We are just beginning to understand the precise molecular steps necessary for piRNA biogenesis and successful silencing of transposons. In an effort to shed light on all aspects of the pathway, from piRNA biogenesis to the effector mechanisms of transposon control, we performed an unbiased, genome-wide RNAi screen in cultured ovarian somatic cells. The primary in vitro screen proved to be a robust and specific assay for transposon derepression, with all expected piRNA components scoring strongly. To assess the validity of the primary data set, we tested our top candidate hits for effects in vivo. In order to make this resource more accessible to the scientific community, we created a web resource with all primary and validation data points (<http://somatic-pirnascreen.cancan.cshl.edu/>).

Within our list of 87 validated genes, we have promising candidates for filling almost every gap in our current understanding of the piRNA pathway (Figure 6). For example, the identified RNA export factors and nucleoporins could act in the export of primary cluster transcripts to the cytoplasm. We also identify genes that are likely to affect transcription of these piRNA precursors. WDE was shown to be a cofactor of *eggless*, a gene required for transcription of clusters (Koch et al., 2009; Rangan et al., 2011). While depletion of EGG did not result in mobilization of *gypsy* (as previously shown), *wde* knockdown led to high levels

of *gypsy* expression, hinting toward a role for WDE independent of EGG. *lin-52* was previously described as a transcriptional activator of Piwi (Georlette et al., 2007). However, Li et al. (2013) recently showed that A-MYB, which provides activity orthologous to the dREAM complex in mice, controls the expression of key pathway components as well as piRNA precursor transcripts.

Following nuclear export, piRNA precursors have to be further processed by an endonuclease to create the 5' end of the mature piRNA. ZUC, which was recently shown to be a cytoplasmic, single-stranded RNA-specific endonuclease, is the most likely candidate for this function (Ipsaro et al., 2012; Nishimasu et al., 2012). The fact that we do not identify any other annotated endonuclease with comparable derepression phenotypes in our screen supports the role for ZUC in this step. Ribonuclease (RNase) P and RNase Z, both endonucleases implicated in transfer RNA (tRNA) processing, did score in the primary screen but could not be validated in vivo because of their severe developmental defects (Dubrovsky et al., 2004; Frank and Pace, 1998). After 5' end formation and loading into Piwi, each piRNA is proposed to be further trimmed to its mature length. The only genes exhibiting exonuclease activity and scoring highly in our screen were *csf4* and *rrp6*, both components of the exosome (Andrulis et al., 2002). However, neither of the two genes could be validated in vivo due to arrested gonadal development in knockdowns.

Transcriptional silencing of transposons through Piwi is a nuclear process, and previous data have demonstrated that unloaded Piwi remains in the cytoplasm (Saito et al., 2010). One protein possibly involved in reimportation of loaded Piwi is Karybeta3, a homolog of Importin 5 (Mosammaparast and Pemberton, 2004), which emerged as a hit from our screen.

Upon reentry into the nucleus, Piwi is able to recognize transcription of active transposons through its bound piRNA and consequently silence them (Le Thomas et al., 2013; Rozhkov et al., 2013; Sienski et al., 2012). So far, only Piwi itself and MAEL have been implicated in this step. With Asterix, we present a component of the nuclear piRNA silencing machinery that is indispensable for transcriptional repression. However, even though we see lower levels of H3K9me3 in mutant animals compared to heterozygous siblings, Asterix most likely is not directly responsible for depositing these marks. The only conserved domains within the protein are predicted to be RNA binding (Andreeva and Tidow, 2008). This still leaves a need for identifying chromatin remodelers and histone methyltransferases (HMTs) that act as piRNA effectors in TGS. No annotated HMTs were hit in the screen; specifically, disruption of Su(var)3-9, which was recently implicated in epigenetic programming directed by piRNAs, did not lead to any significant transposon derepression, suggesting a possible redundancy in proteins that act in histone methylation (Huang et al., 2013).

Elongation factors also have been shown to play a role in chromatin modification. The elongation factor SPT6, which interacts with the nuclear exosome, emerged as another validated hit of the screen (Andrulis et al., 2002). Indeed, data from fission yeast implicated Spt6 in the silencing of heterochromatic repeats (Kiely et al., 2011). In the *spt6* mutants, decreased recruitment

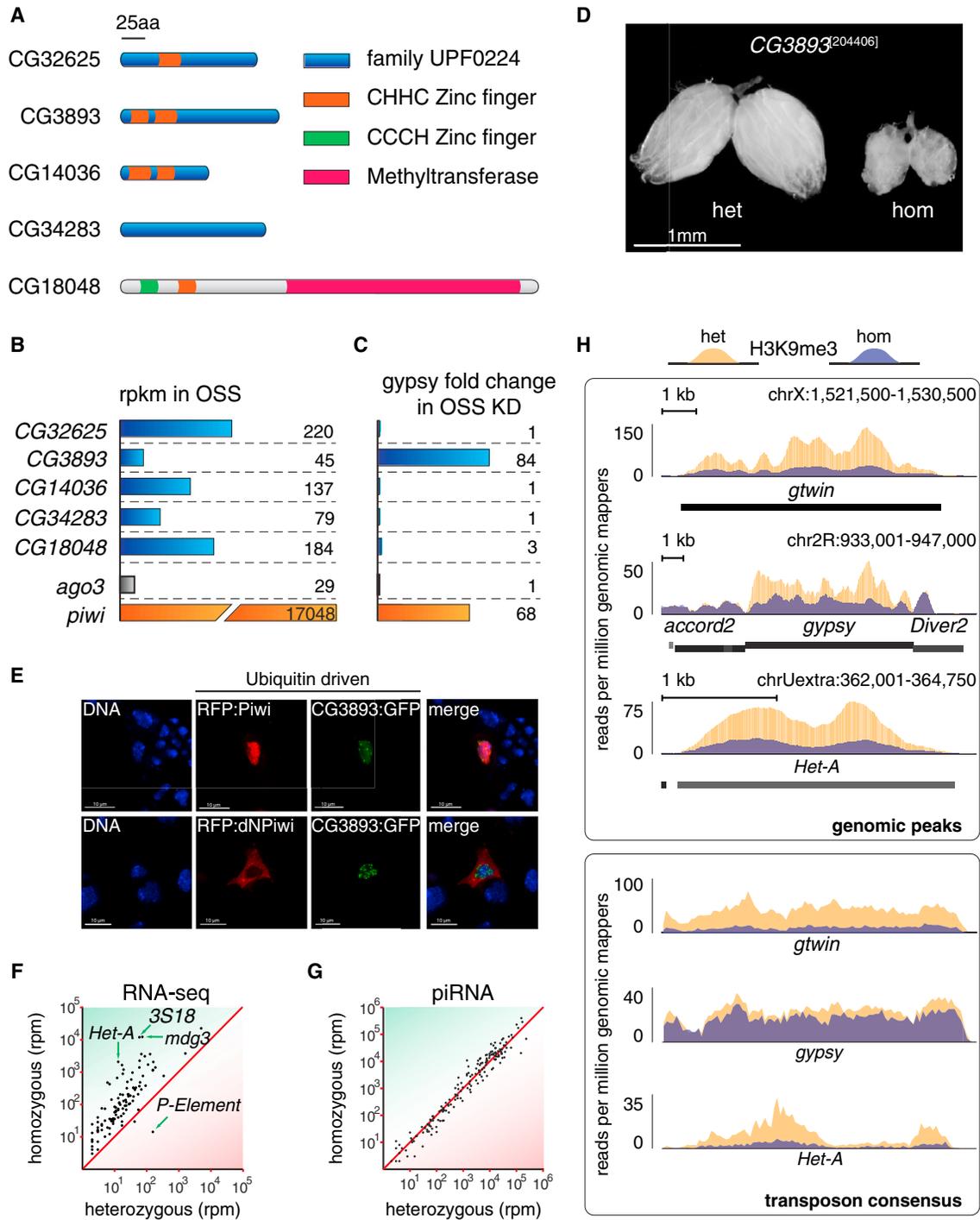


Figure 5. Disruption of CG3893 Function Has a Severe Impact on Transposon Silencing

(A) The five members of the *Drosophila* uncharacterized protein family UPF0224 and their domain structures are diagrammed. The conserved domains are highlighted as colored boxes.

(B) All five family members are weakly expressed in OSS cells. *piwi* and *ago3* expression levels are shown for comparison. Expression levels are based on the modENCODE cell line expression data and are displayed as reads per kilobase per million mapped reads (rpkm).

(C) CG3893, but no other members of its protein family, has a strong impact on transposon silencing upon knockdown in OSS cells. Effects of knockdown of *ago3* and *piwi* are shown for comparison. Numbers represent fold changes of *gypsy* levels with respect to the median fold change of the corresponding plate in the primary screen.

(D) The ovarian morphology of flies heterozygous or homozygous for a P element insertion in CG3893 is shown (204406, Kyoto DGRC). For a more detailed view of the insertion and expression levels see Figure S5.

(legend continued on next page)

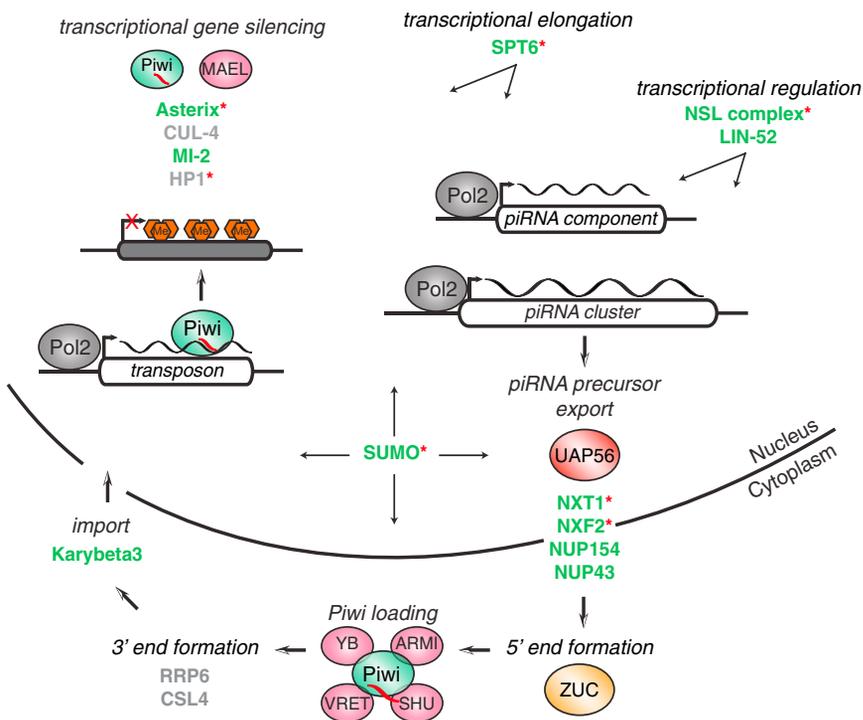


Figure 6. Potential Roles for Newly Identified piRNA Pathway Components

Known piRNA components are shown as bubbles. The newly identified genes are shown in bold text colored according to their validation status (color code as in Figure 2; Pol2, RNA polymerase 2; red hexagons represent H3K9me3). Red asterisks denote genes that validated in the germline screen done in parallel (Czech et al., 2013).

dent to have identified a comprehensive set of pathway components, which to our knowledge have not been implicated in the piRNA pathway. Our data support the current view that the piRNA pathway is the major pathway exerting transposon control, given that both the primary and validation screens were dominated by known components of the piRNA machinery. Our meta-analysis on the primary data set as a whole, as well as the list of validated genes, will provide a resource for the field in efforts toward a greater depth of understanding of piRNA

of the CLRC complex led to loss of H3K9me3 and, as a consequence, lower occupancy of the HP1 homolog Swi6. Cul-4, a homolog of a CLRC member in yeast that is involved in histone methylation in *Drosophila*, was a primary hit in the screen but led to developmental defects during validation (Higa et al., 2006; Hong et al., 2005).

Another validated hit involved in chromatin remodeling was *mi-2*, yet its knockdown only led to modest effects on transposon derepression (Brehm et al., 2000). Heterochromatin protein 1 (HP1), which interacts with Piwi, could not be validated due to developmental defects (Brower-Toland et al., 2007). Intriguingly, MI-2 and HP1 are both sumoylation substrates, implying a function for SUMO beyond the one in piRNA biogenesis demonstrated here (Nie et al., 2009).

In summary, our unbiased, genome-wide approach was successful in identifying likely candidates to fill in many of the gaps in our understanding of the molecular mechanisms of transposon control in *Drosophila*. Together with the findings of a transcriptome-wide screen for germline piRNA pathway components done in parallel (Czech et al., 2013), which showed substantial overlap with the top hits of our screen, we are confi-

production and the mechanisms by which piRNAs silence transposons.

EXPERIMENTAL PROCEDURES

Cell Culture

OSS cells were cultured as previously described and transfected using Xfect Transfection Reagent according to manufacturer's guidelines (Clontech 631317) (Niki et al., 2006).

DNA Plasmids

Expression vectors of CG3893:GFP, RFP:Piwi, and RFP: Δ NTPiwi were made using the *Drosophila* Gateway Collection.

Imaging of Fluorescent Fusion Proteins in OSS

OSS cells were cotransfected with plasmids expressing the indicated fusion proteins using Cell Line Nucleofector Kit V (Amaxa Biosystems; program T-029). Fixed cells were stained with Hoechst 33342 (Invitrogen, R37601).

RNAi Libraries

Two *Drosophila* dsRNA libraries were used in this study: the Open Biosystems *Drosophila* RNAi Collection and the *Drosophila* RNAi Screening Center Genome-wide RNAi library (DRSC 2.0).

(E) Tagged CG3893 colocalizes with Piwi in the nucleus of OSS cells when overexpressed in transient transfections. Nuclear Hoechst staining is blue, GFP-tagged CG3893 is green, and red fluorescent protein (RFP)-tagged Piwi or Δ NT-Piwi is shown in red (Saito et al., 2009).

(F) Transposons are highly upregulated upon disruption of CG3893 in the P element insertion line. A scatter plot of reads per million (rpm) is shown for RNA-seq of heterozygous versus homozygous flies. Each dot represents one transposon consensus sequence. Only sequences mapping in the sense orientation are taken into account.

(G) piRNA levels are not affected by CG3893 disruption. The number of piRNA reads mapped to the same transposon consensus sequences as in (F) is expressed in reads per million.

(H) Levels of H3K9me3 decrease dramatically on a subset of transposons upon depletion of CG3893. Density plots for normalized H3K9me3 ChIP-seq reads over three transposons, *gtwin*, *gypsy*, and *Het-A*, are shown. Yellow distributions correspond to levels in heterozygous flies, and blue distributions correspond to the homozygous state. The upper box shows three distinct genomic peaks over transposon insertions; the lower box shows the corresponding consensus sequences. For all identified H3K9me3 peaks and their read densities see Figure S5C.

RNAi Screening

A detailed description of the primary screen can be found in the [Supplemental Experimental Procedures](#). A basic workflow is shown in [Figure 1A](#). All primers used are listed in [Table S4](#).

Drosophila Stocks and Husbandry

Fly stocks are listed in the [Supplemental Experimental Procedures](#). A description of husbandry and validation screen procedures can be found in the [Supplemental Experimental Procedures](#).

RNA Isolation and qPCR Assays

Total RNA from ten ovaries was extracted with Trizol and purified by organic extraction followed by isopropanol precipitation. After DNase treatment, complementary DNA (cDNA) was synthesized from 800 ng RNA using oligo dT primers and SuperScript III Reverse Transcriptase (Life Technologies). qPCR was performed to assay levels of *gypsy*, *ZAM*, *gypsy3*, and *rp49*. Fold changes for transposons were calculated using the delta Ct method ([Livak and Schmittgen, 2001](#)). All primers used are listed in [Table S4](#).

RNA-Seq and Analysis

For RNA-seq libraries, 2.5–5 μ g of total RNA was depleted of ribosomal RNA using the Epicentre Ribo-Zero rRNA Removal Kits (Human/Mouse/Rat) following the manufacturer's directions. Libraries were prepared using the Illumina ScriptSeq v2 RNA-Seq Library Preparation Kit and were sequenced on an Illumina HiSeq platform. Details on analysis can be found in the [Supplemental Experimental Procedures](#).

Small RNA Cloning and Analysis

For small RNA libraries, total RNA was depleted of 2S rRNA, and libraries were constructed using the Illumina TruSeq small RNA Sample Preparation Kit following the manufacturer's protocol. Details on analysis can be found in the [Supplemental Experimental Procedures](#).

ChIP-Seq

ChIP from 50 ovaries was done as described in [Ram et al. \(2011\)](#) and [Garber et al. \(2012\)](#), with some modifications. Details on the methodology and analysis can be found in the [Supplemental Experimental Procedures](#).

Statistical Procedures

Details on enrichment analysis and statistical procedures can be found in the [Supplemental Experimental Procedures](#).

ACCESSION NUMBERS

RNA-seq, ChIP-seq, and small RNA data have been deposited in the Gene Expression Omnibus database under accession number GSE46009.

SUPPLEMENTAL INFORMATION

Supplemental Information includes five figures, four tables, and Supplemental Experimental Procedures and can be found with this article online at <http://dx.doi.org/10.1016/j.molcel.2013.04.006>.

ACKNOWLEDGMENTS

We are greatly indebted to Norbert Perrimon (DGRC) for giving us reagents. We are grateful to Sabrina Boettcher who managed all logistics and supplies. We would like to thank Alon Goren for invaluable help with ChIP-seq cloning. Molly Hammell aided in analyzing RNA-seq and ChIP-seq data. Leah Sabin provided helpful insight and advice. We wish to thank Stephanie Muller, Assaf Gordon, and Astrid Haase for help with robotics and library normalization. We would like to thank Ben Czech, Jon Preall, Sho Goh, and Julius Brennecke for sharing data prior to publication. Assistance with sequencing was provided by Emily Lee. We would also like to thank Jo Leonardo and all members of the Hannon lab for vital support. P.M.G. is a NIH trainee on the CSHL WSBS NIH Kirschstein-NRSA predoctoral T32

GM065094 grant, a William Randolph Hearst Scholar, and a Leslie Quick Junior Fellow. A grant from T. and V. Stanley supported J.G.'s work. Work in the Hannon laboratory is supported by a grant from the National Institutes of Health (5R01GM062534) and a kind gift from Kathryn W. Davis. G.J.H. is an investigator of the HHMI.

Received: February 6, 2013

Revised: April 4, 2013

Accepted: April 5, 2013

Published: May 9, 2013

REFERENCES

- Andreeva, A., and Tidow, H. (2008). A novel CHHC Zn-finger domain found in spliceosomal proteins and tRNA modifying enzymes. *Bioinformatics* *24*, 2277–2280.
- Andrulis, E.D., Werner, J., Nazarian, A., Erdjument-Bromage, H., Tempst, P., and Lis, J.T. (2002). The RNA processing exosome is linked to elongating RNA polymerase II in *Drosophila*. *Nature* *420*, 837–841.
- Brehm, A., Längst, G., Kehle, J., Clapier, C.R., Imhof, A., Eberharter, A., Müller, J., and Becker, P.B. (2000). dMi-2 and ISWI chromatin remodelling factors have distinct nucleosome binding and mobilization properties. *EMBO J.* *19*, 4332–4341.
- Breitkreutz, B.J., Stark, C., Reguly, T., Boucher, L., Breitkreutz, A., Livstone, M., Oughtred, R., Lackner, D.H., Bähler, J., Wood, V., et al. (2008). The BioGRID Interaction Database: 2008 update. *Nucleic Acids Res.* *36* (Database issue), D637–D640.
- Brennecke, J., Aravin, A.A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., and Hannon, G.J. (2007). Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell* *128*, 1089–1103.
- Brower-Toland, B., Findley, S.D., Jiang, L., Liu, L., Yin, H., Dus, M., Zhou, P., Elgin, S.C., and Lin, H. (2007). *Drosophila* PIWI associates with chromatin and interacts directly with HP1a. *Genes Dev.* *21*, 2300–2311.
- Bucheton, A. (1995). The relationship between the flamenco gene and gypsy in *Drosophila*: how to tame a retrovirus. *Trends Genet.* *11*, 349–353.
- Chintapalli, V.R., Wang, J., and Dow, J.A.T. (2007). Using FlyAtlas to identify better *Drosophila melanogaster* models of human disease. *Nat. Genet.* *39*, 715–720.
- Czech, B., Preall, J.B., McGinn, J.T., and Hannon, G.J. (2013). A transcriptome-wide RNAi screen in the *Drosophila* ovary reveals novel factors of the germline piRNA pathway. *Mol. Cell* *50*. Published online May 9, 2013. <http://dx.doi.org/10.1016/j.molcel.2013.04.007>.
- Dietzl, G., Chen, D., Schnorrer, F., Su, K.-C., Barinova, Y., Fellner, M., Gasser, B., Kinsey, K., Ooppel, S., Scheiblauer, S., et al. (2007). A genome-wide transgenic RNAi library for conditional gene inactivation in *Drosophila*. *Nature* *448*, 151–156.
- Dubrovsky, E.B., Dubrovskaya, V.A., Levinger, L., Schiffer, S., and Marchfelder, A. (2004). *Drosophila* RNase Z processes mitochondrial and nuclear pre-tRNA 3' ends in vivo. *Nucleic Acids Res.* *32*, 255–262.
- Frank, D.N., and Pace, N.R. (1998). Ribonuclease P: unity and diversity in a tRNA processing ribozyme. *Annu. Rev. Biochem.* *67*, 153–180.
- Garber, M., Yosef, N., Goren, A., Raychowdhury, R., Thielke, A., Guttman, M., Robinson, J., Minie, B., Chevrier, N., Itzhaki, Z., et al. (2012). A high-throughput chromatin immunoprecipitation approach reveals principles of dynamic gene regulation in mammals. *Mol. Cell* *47*, 810–822.
- Georgette, D., Ahn, S., MacAlpine, D.M., Cheung, E., Lewis, P.W., Beall, E.L., Bell, S.P., Speed, T., Manak, J.R., and Botchan, M.R. (2007). Genomic profiling and expression studies reveal both positive and negative activities for the *Drosophila* Myb MuvB/dREAM complex in proliferating cells. *Genes Dev.* *21*, 2880–2896.
- Graveley, B.R., Brooks, A.N., Carlson, J.W., Duff, M.O., Landolin, J.M., Yang, L., Artieri, C.G., van Baren, M.J., Boley, N., Booth, B.W., et al. (2011). The developmental transcriptome of *Drosophila melanogaster*. *Nature* *471*, 473–479.

- Herold, A., Klymenko, T., and Izaurralde, E. (2001). NXF1/p15 heterodimers are essential for mRNA nuclear export in *Drosophila*. *RNA* 7, 1768–1780.
- Herold, A., Teixeira, L., and Izaurralde, E. (2003). Genome-wide analysis of nuclear mRNA export pathways in *Drosophila*. *EMBO J.* 22, 2472–2483.
- Higa, L.A., Wu, M., Ye, T., Kobayashi, R., Sun, H., and Zhang, H. (2006). CUL4-DDB1 ubiquitin ligase interacts with multiple WD40-repeat proteins and regulates histone methylation. *Nat. Cell Biol.* 8, 1277–1283.
- Hong, E.J., Villén, J., Gerace, E.L., Gygi, S.P., and Moazed, D. (2005). A cullin E3 ubiquitin ligase complex associates with Rik1 and the Clr4 histone H3-K9 methyltransferase and is required for RNAi-mediated heterochromatin formation. *RNA Biol.* 2, 106–111.
- Huang, X.A., Yin, H., Sweeney, S., Raha, D., Snyder, M., and Lin, H. (2013). A Major Epigenetic Programming Mechanism Guided by piRNAs. *Dev. Cell* 24, 502–516.
- Ipsaro, J.J., Haase, A.D., Knott, S.R., Joshua-Tor, L., and Hannon, G.J. (2012). The structural biochemistry of Zucchini implicates it as a nuclease in piRNA biogenesis. *Nature* 491, 279–283.
- Ishizu, H., Siomi, H., and Siomi, M.C. (2012). Biology of PIWI-interacting RNAs: new insights into biogenesis and function inside and outside of germlines. *Genes Dev.* 26, 2361–2373.
- Kawaoka, S., Izumi, N., Katsuma, S., and Tomari, Y. (2011). 3' end formation of PIWI-interacting RNAs in vitro. *Mol. Cell* 43, 1015–1022.
- Khurana, J.S., and Theurkauf, W. (2010). piRNAs, transposon silencing, and *Drosophila* germline development. *J. Cell Biol.* 191, 905–913.
- Kiely, C.M., Marguerat, S., Garcia, J.F., Madhani, H.D., Bähler, J., and Winston, F. (2011). Spt6 is required for heterochromatic silencing in the fission yeast *Schizosaccharomyces pombe*. *Mol. Cell Biol.* 31, 4193–4204.
- Klattenhoff, C., Xi, H., Li, C., Lee, S., Xu, J., Khurana, J.S., Zhang, F., Schultz, N., Koppetsch, B.S., Nowosielska, A., et al. (2009). The *Drosophila* HP1 homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters. *Cell* 138, 1137–1149.
- Koch, C.M., Honemann-Capito, M., Egger-Adam, D., and Wodarz, A. (2009). Wendei, the *Drosophila* homolog of mAM/MCAF1, is an essential cofactor of the H3K9 methyl transferase dSETDB1/Eggless in germ line development. *PLoS Genet.* 5, e1000644.
- Lau, N.C., Robine, N., Martin, R., Chung, W.-J., Niki, Y., Berezikov, E., and Lai, E.C. (2009). Abundant primary piRNAs, endo-siRNAs, and microRNAs in a *Drosophila* ovary cell line. *Genome Res.* 19, 1776–1785.
- Le Thomas, A., Rogers, A.K., Webster, A., Marinov, G.K., Liao, S.E., Perkins, E.M., Hur, J.K., Aravin, A.A., and Tóth, K.F. (2013). Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. *Genes Dev.* 27, 390–399.
- Lévesque, L., Guzik, B., Guan, T., Coyle, J., Black, B.E., Rekosh, D., Hammarskjöld, M.L., and Paschal, B.M. (2001). RNA export mediated by tap involves NXT1-dependent interactions with the nuclear pore complex. *J. Biol. Chem.* 276, 44953–44962.
- Lewis, P.W., Beall, E.L., Fleischer, T.C., Georgette, D., Link, A.J., and Botchan, M.R. (2004). Identification of a *Drosophila* Myb-E2F2/RBF transcriptional repressor complex. *Genes Dev.* 18, 2929–2940.
- Li, X.Z., Roy, C.K., Dong, X., Bolcun-Filas, E., Wang, J., Han, B.W., Xu, J., Moore, M.J., Schimenti, J.C., Weng, Z., and Zamore, P.D. (2013). An Ancient Transcription Factor Initiates the Burst of piRNA Production during Early Meiosis in Mouse Testes. *Mol. Cell*.
- Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2⁻(Delta Delta C(T)) Method. *Methods* 25, 402–408.
- Ma, L., Buchold, G.M., Greenbaum, M.P., Roy, A., Burns, K.H., Zhu, H., Han, D.Y., Harris, R.A., Coarfa, C., Gunaratne, P.H., et al. (2009). GASZ is essential for male meiosis and suppression of retrotransposon expression in the male germline. *PLoS Genet.* 5, e1000635.
- Makarov, E.M., Makarova, O.V., Urlaub, H., Gentzel, M., Will, C.L., Wilm, M., and Lührmann, R. (2002). Small nuclear ribonucleoprotein remodeling during catalytic activation of the spliceosome. *Science* 298, 2205–2208.
- Malone, C.D., and Hannon, G.J. (2009). Small RNAs as guardians of the genome. *Cell* 136, 656–668.
- Malone, C.D., Brennecke, J., Dus, M., Stark, A., McCombie, W.R., Sachidanandam, R., and Hannon, G.J. (2009). Specialized piRNA pathways act in germline and somatic tissues of the *Drosophila* ovary. *Cell* 137, 522–535.
- McQuilton, P., St Pierre, S.E., and Thurmond, J.; FlyBase Consortium. (2012). FlyBase 101—the basics of navigating FlyBase. *Nucleic Acids Res.* 40 (Database issue), D706–D714.
- Mosammaparast, N., and Pemberton, L.F. (2004). Karyopherins: from nuclear-transport mediators to nuclear-function regulators. *Trends Cell Biol.* 14, 547–556.
- Nie, M., Xie, Y., Loo, J.A., and Courey, A.J. (2009). Genetic and proteomic evidence for roles of *Drosophila* SUMO in cell cycle control, Ras signaling, and early pattern formation. *PLoS ONE* 4, e5905.
- Niki, Y., Yamaguchi, T., and Mahowald, A.P. (2006). Establishment of stable cell lines of *Drosophila* germ-line stem cells. *Proc. Natl. Acad. Sci. USA* 103, 16325–16330.
- Nishimasu, H., Ishizu, H., Saito, K., Fukuhara, S., Kamatani, M.K., Bonnefond, L., Matsumoto, N., Nishizawa, T., Nakanaga, K., Aoki, J., et al. (2012). Structure and function of Zucchini endoribonuclease in piRNA biogenesis. *Nature* 491, 284–287.
- Olivieri, D., Sykora, M.M., Sachidanandam, R., Mechtler, K., and Brennecke, J. (2010). An in vivo RNAi assay identifies major genetic and cellular requirements for primary piRNA biogenesis in *Drosophila*. *EMBO J.* 29, 3301–3317.
- Pane, A., Jiang, P., Zhao, D.Y., Singh, M., and Schüpbach, T. (2011). The Cutoff protein regulates piRNA cluster expression and piRNA production in the *Drosophila* germline. *EMBO J.* 30, 4601–4615.
- Pélissou, A., Song, S.U., Prud'homme, N., Smith, P.A., Bucheton, A., and Corces, V.G. (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.
- Raja, S.J., Charapitsa, I., Conrad, T., Vaquerizas, J.M., Gebhardt, P., Holz, H., Kadlec, J., Fraterman, S., Luscombe, N.M., and Akhtar, A. (2010). The nonspecific lethal complex is a transcriptional regulator in *Drosophila*. *Mol. Cell* 38, 827–841.
- Ram, O., Goren, A., Amit, I., Shosh, N., Yosef, N., Ernst, J., Kellis, M., Gymrek, M., Issner, R., Coyne, M., et al. (2011). Combinatorial patterning of chromatin regulators uncovered by genome-wide location analysis in human cells. *Cell* 147, 1628–1639.
- Ramadan, N., Flockhart, I., Booker, M., Perrimon, N., and Mathey-Prevot, B. (2010). Design and implementation of high-throughput RNAi screens in cultured *Drosophila* cells. *Nat. Protoc.* 2, 2245–2264.
- Rangan, P., Malone, C.D., Navarro, C., Newbold, S.P., Hayes, P.S., Sachidanandam, R., Hannon, G.J., and Lehmann, R. (2011). piRNA production requires heterochromatin formation in *Drosophila*. *Curr. Biol.* 21, 1373–1379.
- Rozhkov, N.V., Hammell, M., and Hannon, G.J. (2013). Multiple roles for Piwi in silencing *Drosophila* transposons. *Genes Dev.* 27, 400–412.
- Saito, K., Inagaki, S., Mituyama, T., Kawamura, Y., Ono, Y., Sakota, E., Kotani, H., Asai, K., Siomi, H., and Siomi, M.C. (2009). A regulatory circuit for piwi by the large Maf gene traffic jam in *Drosophila*. *Nature* 461, 1296–1299.
- Saito, K., Ishizu, H., Komai, M., Kotani, H., Kawamura, Y., Nishida, K.M., Siomi, H., and Siomi, M.C. (2010). Roles for the Yb body components Armitage and Yb in primary piRNA biogenesis in *Drosophila*. *Genes Dev.* 24, 2493–2498.
- Scheuermann, J.C., de Ayala Alonso, A.G., Oktaba, K., Ly-Hartig, N., McGinty, R.K., Fraterman, S., Wilm, M., Muir, T.W., and Müller, J. (2010). Histone H2A deubiquitinase activity of the Polycomb repressive complex PR-DUB. *Nature* 465, 243–247.
- Schüpbach, T., and Wieschaus, E. (1989). Female sterile mutations on the second chromosome of *Drosophila melanogaster*. I. Maternal effect mutations. *Genetics* 121, 101–117.
- Schüpbach, T., and Wieschaus, E. (1991). Female sterile mutations on the second chromosome of *Drosophila melanogaster*. II. Mutations blocking oogenesis or altering egg morphology. *Genetics* 129, 1119–1136.

- Sienski, G., Dönertas, D., and Brennecke, J. (2012). Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression. *Cell* 151, 964–980.
- Talamillo, A., Sánchez, J., and Barrio, R. (2008). Functional analysis of the SUMOylation pathway in *Drosophila*. *Biochem. Soc. Trans.* 36, 868–873.
- Tanentzapf, G., Devenport, D., Godt, D., and Brown, N.H. (2007). Integrin-dependent anchoring of a stem-cell niche. *Nat. Cell Biol.* 9, 1413–1418.
- Yoshimura, T., Miyazaki, T., Toyoda, S., Miyazaki, S., Tashiro, F., Yamato, E., and Miyazaki, J.-i. (2007). Gene expression pattern of Cue110: a member of the uncharacterized UPF0224 gene family preferentially expressed in germ cells. *Gene Expr. Patterns* 8, 27–35.
- Yoshimura, T., Toyoda, S., Kuramochi-Miyagawa, S., Miyazaki, T., Miyazaki, S., Tashiro, F., Yamato, E., Nakano, T., and Miyazaki, J.-i. (2009). Gtsf1/Cue110, a gene encoding a protein with two copies of a CHHC Zn-finger motif, is involved in spermatogenesis and retrotransposon suppression in murine testes. *Dev. Biol.* 335, 216–227.
- Zhang, F., Wang, J., Xu, J., Zhang, Z., Koppetsch, B.S., Schultz, N., Vreven, T., Meignin, C., Davis, I., Zamore, P.D., et al. (2012). UAP56 couples piRNA clusters to the perinuclear transposon silencing machinery. *Cell* 151, 871–884.