Genome-wide mapping of Piwi association with specific loci in *Drosophila* ovaries

Na Liu,[1,2,†] Nils Neuenkirchen,[1,2,†] Mei Zhong,[1,2] and Haifan Lin[1,2,*]

[1]Yale Stem Cell Center, Yale University School of Medicine, New Haven, CT 06520-8073, USA
[2]Department of Cell Biology, Yale University School of Medicine, New Haven, CT 06520-8073, USA

*Corresponding author: Yale Stem Cell Center, Yale University School of Medicine, 10 Amistad Street, Room 237E, New Haven, CT 06509-8073, USA. haifan.lin@yale.edu
[†]These authors contributed equally to this work.

Abstract

Small noncoding RNA pathways have been implicated in diverse mechanisms of gene regulation. In *Drosophila* ovaries, Piwi binds to Piwi-interacting RNAs (piRNAs) of mostly 24–28 nucleotides (nt) and plays an important role in germline stem cell maintenance, transposon repression, and epigenetic regulation. To understand the mechanism underlying these functions, we report the application of the DamID-seq method to identify genome-wide binding sites of Piwi in *Drosophila* ovaries. Piwi localizes to at least 4535 euchromatic regions that are enriched with piRNA target sites. Surprisingly, the density of Piwi binding to euchromatin is much higher than in heterochromatin. Disrupting the piRNA binding of Piwi results in an overall change of the genomic binding profile, which indicates the role of piRNAs in directing Piwi to specific genomic sites. Most Piwi binding sites were either within or in the vicinity of protein-coding genes, particularly enriched near the transcriptional start and termination sites. The methylation signal near the transcriptional termination sites is significantly reduced when Piwi was mutated to become defective in piRNA binding. These observations indicate that Piwi might directly regulate the expression of many protein-coding genes, especially through regulating the 3′ ends of targeted transcripts.

Keywords: Piwi; piRNA; chromatin association; DamID-seq

Introduction

Epigenetic factors play a central role in establishing and maintaining heritable chromatin states that dictate gene expression. A prerequisite for the proper function of epigenetic factors is their faithful binding to correct target sites in the genome. However, many epigenetic factors do not possess DNA sequence recognition ability (Holoch and Moazed 2015; Soshnev *et al.* 2016). Thus, how these epigenetic factors are recruited to their target sites remains largely unknown. In recent years, small noncoding RNAs have been implicated in epigenetic programming as DNA-sequence recognition molecules, as best characterized in the fission yeast *S. pombe* (Grewal 2010). In this process, small noncoding RNAs work cooperatively with specific RNA-binding proteins such as the PAZ-PIWI-Domain (PPD) protein family, which is composed of AGO (Argonaute) and PIWI (P-element-Induced WImpy testes) subfamilies (Cox *et al.* 1998; Thomson and Lin 2009). AGO subfamily proteins bind to microRNAs (miRNAs) or small interfering RNAs (siRNAs) in both animals and plants and are ubiquitously expressed in most tissues (Bartel 2004). PIWI subfamily proteins bind to a distinct class of small RNAs ranging from 24 to 32 nt called PIWI-interacting RNAs (piRNAs) and are predominantly expressed in the animal gonads (Aravin *et al.*

2006; Girard *et al.* 2006; Grivna *et al.* 2006; Lau *et al.* 2006; Watanabe *et al.* 2006).

In *Drosophila*, the eponymous member of the PIWI subfamily, Piwi, is mainly present in the nucleus of both germline and gonadal somatic cells (Cox *et al.* 2000; Saito *et al.* 2009). Several functions have been attributed to Piwi, including the repression of transposable elements (Brennecke *et al.* 2007), germline stem cell maintenance (Cox *et al.* 1998), oogenesis, spermatogenesis (Lin and Spradling 1997; Gonzalez *et al.* 2015), canalization (Gangaraju *et al.* 2011), and epigenetic regulation (Brower-Toland *et al.* 2007; Yin and Lin 2007; Saito and Siomi 2010; Huang *et al.* 2013; Saito 2013).

The epigenetic function of Piwi has been associated with Piwi binding to specific genomic sequences as guided by its cognate piRNA. This was first shown for Piwi binding to a subtelomeric site in the genome (Yin and Lin 2007; Lin and Yin 2008). Two subsequent reports indicated that Piwi recruits Heterochromatin Protein 1a (HP1a) and the H3K9 methyltransferase Su(var)3-9 to the chromatin (Brower-Toland *et al.* 2007; Huang *et al.* 2013). Furthermore, the absence of Piwi leads to the selective loss of H3K9me3 marks (Pal-Bhadra *et al.* 2004; Sienski *et al.* 2012; Huang *et al.* 2013). Several other studies further illustrated that Piwi is guided to nascent transcripts by its associated piRNA

(Brower-Toland *et al.* 2007; Sienski *et al.* 2012; Huang *et al.* 2013; Le Thomas *et al.* 2013; Rozhkov *et al.* 2013; Post *et al.* 2014; Sytnikova *et al.* 2014). However, piRNA-independent binding of Piwi to a small number of specific genomic site has also been observed (Peng *et al.* 2016).

To study how extensively Piwi is involved in epigenetic regulation and to what extent the regulation depends on piRNAs in the *Drosophila* ovary, systematic mapping of Piwi binding sites in the genome needs to be achieved. At present, only two chromatin immunoprecipitation followed by deep sequencing (ChIP-seq) studies have been reported for Piwi (Huang *et al.* 2013; Lin *et al.* 2015; Marinov *et al.* 2015; Peng *et al.* 2016). However, the quality of the ChIP-seq data was insufficient for concluding whether Piwi binds to piRNA target sites in the genome. The difficulty of ChIP-seq of Piwi is likely because it binds to genomic sites via piRNA hybridization with nascent RNAs (Brower-Toland *et al.* 2007; Sienski *et al.* 2012; Huang *et al.* 2013; Le Thomas *et al.* 2013; Rozhkov *et al.* 2013; Post *et al.* 2014; Sytnikova *et al.* 2014). This mode of binding may cause several challenges for performing Piwi ChIP: (1) the targeting is RNase-sensitive; (2) as nascent transcripts constantly travel through the transcription unit, the Piwi-ChIP signal spreads from one focused site to the entire downstream transcribed region, drastically reducing the overall signal intensity; and (3) nontethered transcripts in the nucleus can compete for Piwi-piRNA binding. Together, these mechanisms could effectively reduce the signal-to-noise ratio of ChIP-seq.

To bypass the challenges in Piwi ChIP-seq, here we combined the DNA adenine methylation identification (DamID) method (van Steensel and Henikoff 2000) with deep-sequencing to map the binding sites of Piwi in the whole genome in the *Drosophila* ovary. This method does not require the use of antibodies, can identify chromatin binding sites mediated by both stable as well as transient protein-DNA interactions, and thus provides an independent assessment of Piwi interaction with chromatin. The DamID method was recently applied to the *Drosophila* ovary (Ilyin *et al.* 2017). Here we report the mapping of Piwi binding in the genome by comparing wild-type Piwi and a mutant Piwi protein deficient in piRNA binding. Using this approach, we identified thousands of Piwi binding sites. Most of the sites are localized in euchromatic regions corresponding to piRNA-target sites. The targeting sites are preferentially located in protein-coding genes, especially at the transcription start and termination sites (TSS and TTS). Finally, disrupting the Piwi-piRNA interaction diminished the Piwi binding density in the 3' region of bound genes. These findings provide strong evidence supporting the piRNA-guided Piwi association with specific genomic sites and reveal the pattern of such association.

Materials and methods
DNA constructs and site-directed mutagenesis
The coding sequence of *myc*-tagged genomic *piwi* was transferred from pCa4B2G_*myc-piwi* (Darricarrere *et al.* 2013) into pENTR (GenBank accession number: GU931384.1) using restriction enzymes *Xba*I and *Spe*I resulting in pENTR_*myc-piwi*. Following the removal of an internal *Nsi*I site from the coding sequence of the *Escherichia coli* DNA methyltransferase (Dam) in pLgw_RFC1-V5-EcoDam (van Steensel and Henikoff 2000) using QuikChange Site-Directed Mutagenesis [Agilent Genomics; Dam(-NsiI)_up, Dam(-NsiI)_low; see Supplementary Table S3 for primer sequences], the Dam sequence was amplified (Dam-NsiI_F, Dam-Nsi_R) and inserted into the linearized pENTR_*myc-piwi* vector (myc-NsiI_F, pENTR-NsiI_R) via the restriction enzyme *Nsi*I resulting in

pENTR_Dam-*myc-piwi*. Subsequently, the *Dam-myc* sequence was amplified (Dam-XbaI_F, myc-SpeI_R) and subcloned into an empty pENTR vector using restriction enzymes *Xba*I and *Spe*I to generate the construct pENTR_*Dam-myc*. Mutations to disrupt piRNA binding in Piwi [Y551L, K555E; as described in (Le Thomas *et al.* 2013)] were introduced to the pENTR_*Dam-myc-piwi* plasmid by QuikChange (YK_UP, YK_LOW) and verified by Sanger sequencing. Plasmids for injection into fly embryos were generated by using the Gateway cloning system (Invitrogen). Donor vectors pENTR_*Dam-myc*, pENTR_Dam-*myc-piwi*[(WT)], and pENTR_*Dam-myc-piwi*[(YK)] and the destination vector pPW (containing a UASp upstream of the clonase insertion site) were recombined following the manufacturer's instructions.

Fly stocks and generation of transgenic flies
Cultures were reared at 25°C on standard molasses agar medium. The fly stocks BL#2077 (w^+; P{*w*[+*mC*]=GAL4-Hsp70.PB}2), carrying the heat-shock-inducible *gal4* transgene, BL#3703 [w^{1118}/Dp(1; Y)y^+; CyO/nub[1] b^1 Sco lt^1 stw[3]; MKRS/TM6B, Tb[1]], a double balancer line, and BL#5138 (y^1 w[*]; P{*w*[+*mC*]=tubP-GAL4}LL7/TM3, Sb^1 Ser[1]), expressing ubiquitous Gal4 under the *alphaTub84B* promoter, were obtained from the Bloomington *Drosophila* Stock Center at Indiana University.

Transgenic flies were generated by BestGene, Inc. (Chino Hills, CA) by injecting pPW-based constructs containing the coding sequences for either Myc-Piwi[(WT)], Myc-Piwi[(YK)], Dam, Dam-Piwi[(WT)], or Dam-Piwi[(YK)] into y^1 w^{67c23}; P{CaryP}attP2 (BL#8622) fly embryos. DNA insertion occurred via the PhiC31 integrase at the Cyto Site 68A4 (FlyBase *Drosophila melanogaster*, R5.47). Transgenic flies and the *hsp70-gal4* driver line (BL#2077) and *tub-gal4* driver line (BL#5138) were balanced by crossing with above-mentioned BL#3703 flies. Finally, transgenic fly lines containing an *hsp70*-promoter or *tubulin*-driven *gal4* driver and one of the respective transgenes containing an upstream *UASp* [*hs-gal4/CyO*; *UASp*: *myc-piwi*[(WT)],w^+/TM6B, and *hs-gal4/CyO*; *UASp*: *myc-piwi*[(YK)],w^+/TM6B; *hs-gal4/CyO*; *UASp*: *dam,* w^+/TM6B, *hs-gal4/CyO*; *UASp*: *dam-piwi*[(WT)],w^+/TM6B, and *hs-gal4/CyO*; *UASp*: *dam-piwi*[(YK)],w^+/TM6B] were obtained.

Genotyping of transgenic flies
Genomic DNA of transgenic flies was isolated according to Gloor *et al.* (1993). One microliter of this reaction was used in a 15 μl PCR reaction comprising 2× GO-Taq (Promega), and 5 pmol of forward and reverse primer (see Supplementary Table S3). PCR program: 95°C, 30 s; (95°C, 30 s; 55°C, 30 s; 72°C, 30 s) ×34; 72°C, 30 s; 12°C, ∞. PCR reactions were separated on a 1% agarose gel.

Immunofluorescence of Drosophila ovaries
Transgenic flies expressing EGFP-Piwi[(WT)] and EGFP-Piwi[(YK)] were obtained from the Aravin lab (Le Thomas *et al.* 2013). Transgenic flies for this work containing *UASp: myc-piwi*[(WT)],w^+ and *UASp: myc-piwi*[(YK)],w^+ were crossed with flies harboring an *hsp70-gal4* driver. The resulting flies were heat-shocked twice at 37°C for 90 min with 6 h in between the heat-shocks. Ovaries of 3–4-day-old flies were dissected in ice-cold 1× phosphate-buffered saline (PBS) and fixed for 10 min in 1× PBS, 0.3% (v/v) Triton X100, 4% (w/v) paraformaldehyde at room temperature (RT). After washing (3×) in 1× PBS, the tissue was incubated in 1× PBST [1× PBS and 0.3% (v/v) Triton X100] for 10 min at RT. Samples were blocked in 1× PBST and 5% (v/v) normal goat serum (NGS) for 1 h at RT and then incubated in 1× PBST, 5% NGS, and anti-Myc antibody (9B11, Cell Signalling; dilution: 1: 1000) at 4°C overnight. After washing (3×) in 1× PBS, the tissue was incubated in 1× PBST, 5%

NGS, and secondary antibody antimouse (goat), Alexa Fluor® 488 (Thermo Fisher; dilution: 1:400) for 1 h at RT. Finally, ovaries were stained with DAPI (4',6-diamidino-2-phenylindole; 1:1000) and mounted in VECTASHIELD (Vector Laboratories). For the EGFP-Piwi samples, the antibody-staining steps were omitted as the transgenes already carried a fluorescent marker. Samples were imaged using a Leica TCS SP5 Spectral Confocal Microscope on sequential scanning mode. Image collection was carried out using the Leica Application Suite imaging software.

Piwi DamID-seq

Biological triplicates of Dam-only, Dam-Piwi(WT), and Dam-Piwi(YK) ovaries were analyzed. 20 ovaries of 3–4-day-old transgenic flies raised at 25°C without heat-shock were dissected in ice-cold 1× PBS. Dissected ovaries were supplemented with 100 μl of buffer A [buffer A: 10 mM Tris/HCl (pH7.5), 60 mM NaCl, 10 mM EDTA, 0.2 mg/ml proteinase K] and ground with a sterile pestle. 400 μl of buffer A and 500 μl of buffer B [buffer B: 200 mM Tris/HCl (pH7.5), 30 mM EDTA, 2% (w/v) SDS, 0.2 mg/ml proteinase K] were added and the samples were incubated at 37°C for 1 h. Genomic DNA was isolated by phenol/chloroform extraction. Ethanol-precipitated DNA was resuspended on ice in 500 μl 1× Tris/EDTA (TE) buffer. The sample was treated with 100 μg/ml of RNase A (QIAGEN) to digest co-purified RNA molecules at 37°C for 30 min. Next, proteinase K was added (100 μg/ml) to degrade the RNase A at 37°C for 30 min. Following another phenol/chloroform extraction, the pellet was resuspended in 30 μl 1× TE buffer. The isolated DNA was analyzed by agarose gel electrophoresis for purity and stored at 4°C to prevent possible fragmentation by freeze-thawing cycles. To cleave methylated GATC sites in the isolated genomic DNA (gDNA), 2.5 μg gDNA was digested with *DpnI* enzyme (10 units, NEB) at 37°C for 2 h and the enzyme was heat-inactivated for 20 min at 80°C. For the annealing of DamID sequencing adapters, ssDNA oligos (AdRt, AdRb; see Supplementary Table S3) were incubated at 95°C for 5 min and slowly cooled down to 25°C at –1°C/min. 40 pmol of adapters were added to the digested genomic DNA and ligated at 16°C for 2 h (T4 DNA ligase, NEB) before the ligase was heat-inactivated at 65°C for 10 min. Next, unmethylated GATC sites were digested with *DpnII* (10 units, NEB) at 37°C for 2 h and the enzyme was inactivated at 65°C for 20 min. One percent of the thus prepared DamID sample (1 of 50 μl) was amplified by Q5 DNA polymerase (NEB; DamID_PCR): 98°C, 30 s; (98°C, 10 s; 55°C, 20 s; 72°C, 2 min) ×19; 72°C, 2 min; 12°C, infinite. A small fraction of the amplified DNA was analyzed by agarose gel electrophoresis and ethidium bromide staining. The rest was column purified and eluted in ddH$_2$O before removing the DamID adapters by *DpnII* digestion (10 units, NEB). Following a final round of column purification, 10 ng of the eluted DNA fragments was used in Illumina library preparation (Illumina TruSeq ChIP Sample Preparation Kit). The ligation products were purified and size-selected by agarose gel electrophoresis. Size-selected DNA was then purified and PCR-amplified to enrich for fragments with adapters on both ends. The final purified product was quantitated before cluster generation. Sequencing was carried out using a HiSeq2000 with one sample per lane at 100 bp single-end sequencing.

Bioinformatics analysis of DamID-seq data

Library preparation was conducted following Illumina's instruction. Deep sequencing was carried out with three biological replicates for each condition. The raw reads of 100 bp were mapped to the *Drosophila* genome release 5 (dm3 in UCSC) using bowtie2 allowing for no mismatches (Langmead and Salzberg 2012). In addition, all methylated GATC sites were also mapped to dm6 (Supplementary Table S4), a more updated version. Uniquely mapped sequencing reads were applied to the iDEAR algorithm (Gutierrez-Triana *et al.* 2016) in order to call enriched methylation sites. Identified sites with a P-value of ≤0.01 were selected for downstream analysis. The Ensembl gene and transposon annotation from UCSC RepeatMasker was used to annotate the identified methylation enrichment sites. A 1 kb flanking region was applied for the annotation. For euchromatic *vs.* heterochromatic annotation, chrXHet, chr2LHet, chr2RHet, chr3LHet, chr3RHet, and chrU were classified as heterochromatin, and chrX, chr2L, chr2R, chr3L, chr3R, and chr4 as euchromatin. The resulting reads were normalized by the individual chromosome length for the two classes. We used a piRNA dataset that was previously generated in our lab (Yin and Lin 2007) and is available on GEO under the accession number GSE9138. Only piRNAs that mapped to the *Drosophila* genome (dm3) with zero mismatches were considered. Multiply mapped piRNAs were weighted at each mapping site. Genomic regions of ±400bp of Piwi-led methylation enrichment sites were scanned for piRNAs. The piRNA abundance was finally normalized according to the span of the methylation enrichment regions from iDEAR.

Data availability

Strains, plasmids, and bioinformatics scripts are available upon request. Raw datasets have been submitted to the NCBI SRA database (accession number PRJNA662571). Supplementary materials are available at Figshare: https://doi.org/10.25387/g3. 13347677. Supplementary Table S1 provides an overview of the DamID sequencing reads and mapping rates. Supplementary Table S2 contains details about the distribution of GATC sites in the female *Drosophila* genome. Supplementary Table S3 summarizes the DNA oligomers used for cloning, DamID-seq, and PCR screening of transgenic flies. Supplementary Table S4 lists all methylated GATC sites and their sequencing reads in Dam-only, Dam-Piwi(WT), and Dam-Piwi(YK) transgenic flies. The genomic coordinates were converted to the *Drosophila melanogaster* genome version dm6 from dm3 by UCSC LiftOver. Supplementary Figure S1 shows the annotation of uniquely mapped reads in euchromatin and heterochromatin. Supplementary Figure S2 shows the annotation of multi-mapped reads in euchromatin and heterochromatin. Supplementary Figure S3 illustrates the number of modified GATC sites in repeat and nonrepeat sequences. Supplementary Figure S4 depicts methylation peaks normalized to the GATC density in genes and repeats. Reagents Table contains an overview of fly strains, plasmids, reagents, and software packages used in this work.

Supplementary material is available at figshare DOI: https://doi.org/10.25387/g3.13347677.

Results and discussion
Generating Dam-Piwi transgenic flies and optimizing Dam expression for Piwi mapping

To examine how extensively Piwi associates with specific chromatin sites (Huang *et al.* 2013; Le Thomas *et al.* 2013; Rozhkov *et al.* 2013; Post *et al.* 2014) and to avoid the challenges encountered in Piwi ChIP, we sought to use a method that met the following two criteria: (1) it does not rely on stable or strong protein-DNA binding and can record transient protein-DNA interactions in the genome; and (2) it does not require the use of antibodies, which introduce variability in ChIP efficiency and specificity. To fulfill these criteria, we used the DamID strategy,
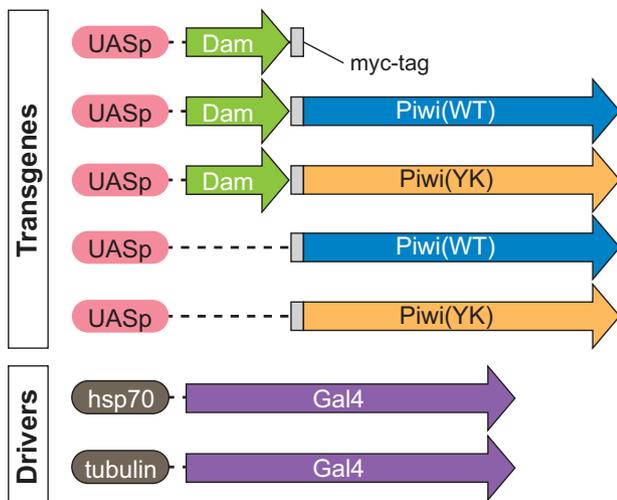
Figure 1 Schematic overview of DamID transgenes. *Dam*, *Dam-piwi*[(WT)], and *Dam-piwi*[(YK)] genes were expressed under a leaky *hsp70* promoter using the GAL4/UAS system. *myc-piwi*[(WT)] and *myc-piwi*[(YK)] genes were expressed using the same system but using two subsequent heat-shocks at 37°C. The Myc-tag comprises 10 aa. *Dam*, *myc*, *piwi*, and *gal4* sequences are depicted in scale, whereas the *UASp*, *hsp70*, and tubulin promoter sequences are not.

in which *Drosophila* Piwi is fused to the *E. coli* DNA adenine methyltransferase (Dam; Figure 1). If Piwi associates with a nascent transcript, it will position the methyltransferase close to the binding site to methylate the N^6 atom of the adenosine residue (m^6A) in nearby GATC sequences in the DNA (van Steensel and Henikoff 2000). A previous study reported that only very low levels of endogenous m^6A exist in the *Drosophila* genome (Zhang *et al.* 2015) and endogenous m^6A could not be detected in previous DamID studies. Thus, the endogenous background signal should not influence our DamID analysis, especially as we aimed to compare changes in patterns of Gm^6ATC sites between Dam-Piwi and Dam transgenic flies.

To perform Piwi-DamID, we generated the following transgenic flies expressing Dam-fusion proteins under the UASp promoter: (1) *dam-myc*; (2) *dam-myc-piwi*[(WT)] (binds to piRNAs); and (3) *dam-myc-piwi*[(YK)] (does not bind to piRNAs). Furthermore, we generated transgenic flies under the same *UASp* promoter expressing *myc-piwi*[(WT)] and *myc-piwi*[(YK)] to study the change of cellular Piwi localization upon piRNA binding (Figure 1). All transgenes contained a *myc* sequence to detect protein expression by Western blotting and immunofluorescence experiments. Wild-type Piwi can interact with piRNA that possibly guides the protein to its genomic loci, the Y551E and K555E mutant Piwi (Piwi[(YK)]) is unable to bind to piRNA (Le Thomas *et al.* 2013). Consequently, by comparing these two forms of Piwi, we would be able to specifically study the impact on piRNA-guided binding of Piwi to genomic sites.

We reasoned that the Dam protein alone should randomly methylate all accessible GATC sites in the genome. If Piwi is indeed associated with specific genomic targets, the attachment of Dam to full-length Piwi (Dam-Piwi) should lead to the enrichment of m^6A in GATC sites within a few kilobases of Piwi binding sites (van Steensel and Henikoff 2000; Vogel *et al.* 2007), shifting the distribution of methylation signals in the genome, yet not changing the overall level of methylation. Abolishing the piRNA-binding ability of Piwi would disrupt this enrichment.

For the DamID mapping to be successful, the Dam protein level must be strictly regulated because a strong methylation activity would provide a high level of nonspecific background

signal. Furthermore, overexpression of the Dam protein is embryonic lethal. To exemplify this, we expressed Dam-Myc, Dam-Myc-Piwi, and Myc-Piwi under a *tubulin: gal4* promoter using the GAL4/UAS system (Figure 2A). The expression of Dam alone or fused to wild-type or mutant Piwi resulted in embryonic lethality (Figure 2, B–D), whereas the expression of Myc-Piwi did not adversely affect embryonic survival (Figure 2E). Consequently, in order to provide only low levels of the Dam protein, we expressed *Dam*-fusion genes under the control of the *UASp* promoter that can be activated at a relatively low level upon binding of the yeast Gal4 protein. We used a *Drosophila* heat-shock promoter (*hsp70*) to regulate Gal4 expression while keeping the flies at 25°C. This way only leaky expression would be observed, resulting in very low levels of Dam-fusion proteins.

Piwi requires an associated piRNA to enter the nucleus (Le Thomas *et al.* 2013), which likely causes a conformational change in Piwi, exposing the nuclear localization signal at its N-terminus (Yashiro *et al.* 2018). Thus, preventing piRNAs from binding to Piwi should shift the majority of Piwi[(YK)] molecules from the nucleus to the cytoplasm. The few Piwi[(YK)] molecules that do enter the nucleus should not be able to recognize genomic targets via RNA complementarity as they lack associated piRNA. We expressed wild-type and mutant EGFP-Piwi in *Drosophila* ovaries under the *piwi* promoter and verified that the mutant Piwi[(YK)] protein is mostly restricted to the cytoplasm (Figure 3, A and B). Furthermore, we used two consecutive heat-shocks to express Myc-tagged Piwi from the GAL4/UAS constructs generated in this study in order to follow Piwi localization upon the disruption of piRNA binding (Figure 3, C and D). Both approaches showed that piRNA binding is necessary for Piwi to enter the nucleus where the DamID reaction is carried out.

To validate that the mutant Piwi protein could be used in DamID, we first drove Dam-Piwi[(YK)] expression via the GAL4/UAS system using a tubulin promoter. Expectedly, this resulted in embryonic lethality (Figure 2D), raising the possibility that some Dam-Piwi[(YK)] was able to interact with genomic DNA. This indicates that the fusion protein might have entered the nucleus and methylated genomic GATC sites without piRNA. However, we cannot rule out the possibility that a small fraction of Piwi[(YK)] might still be able to bind to piRNA that enables it to enter the nucleus and bind to chromatin independent of piRNA. Regardless, these results indicate that applying Piwi-DamID causes methylation of the *Drosophila* genome.

Dam-Piwi preferably methylates euchromatic regions

We next used biological triplicates of each Dam, Dam-Piwi[(WT)], and Dam-Piwi[(YK)] to map the Piwi-mediated methylation sites by DamID-seq (see Figure 4A for a schematic overview of sample processing). Methylated GATC sites in the genomic DNA from fly ovaries were detected by *DpnI* cleavage. DNA adapter sequences were ligated to the cleaved sites, followed by *DpnII* digestion of any unmethylated GATC site, but not hemimethylated ones. DNA that contained adapters on both sides therefore corresponded to genomic fragments in which two adjacent GATC sites had been methylated by Dam. PCR amplification resulted in DNA fragments of various sizes in Dam, Dam-Piwi[(WT)], and Dam-Piwi[(YK)], but not in w[1118] control flies (Figure 4B), indicating that the transgenic Dam protein caused m^6A formation in the *Drosophila* ovary and that any endogenous m^6A signals were negligible. Following the adapter removal, DNA libraries were prepared for Illumina 100 bp single-end sequencing. Overall, we obtained between 27 and 61 million reads for the individual
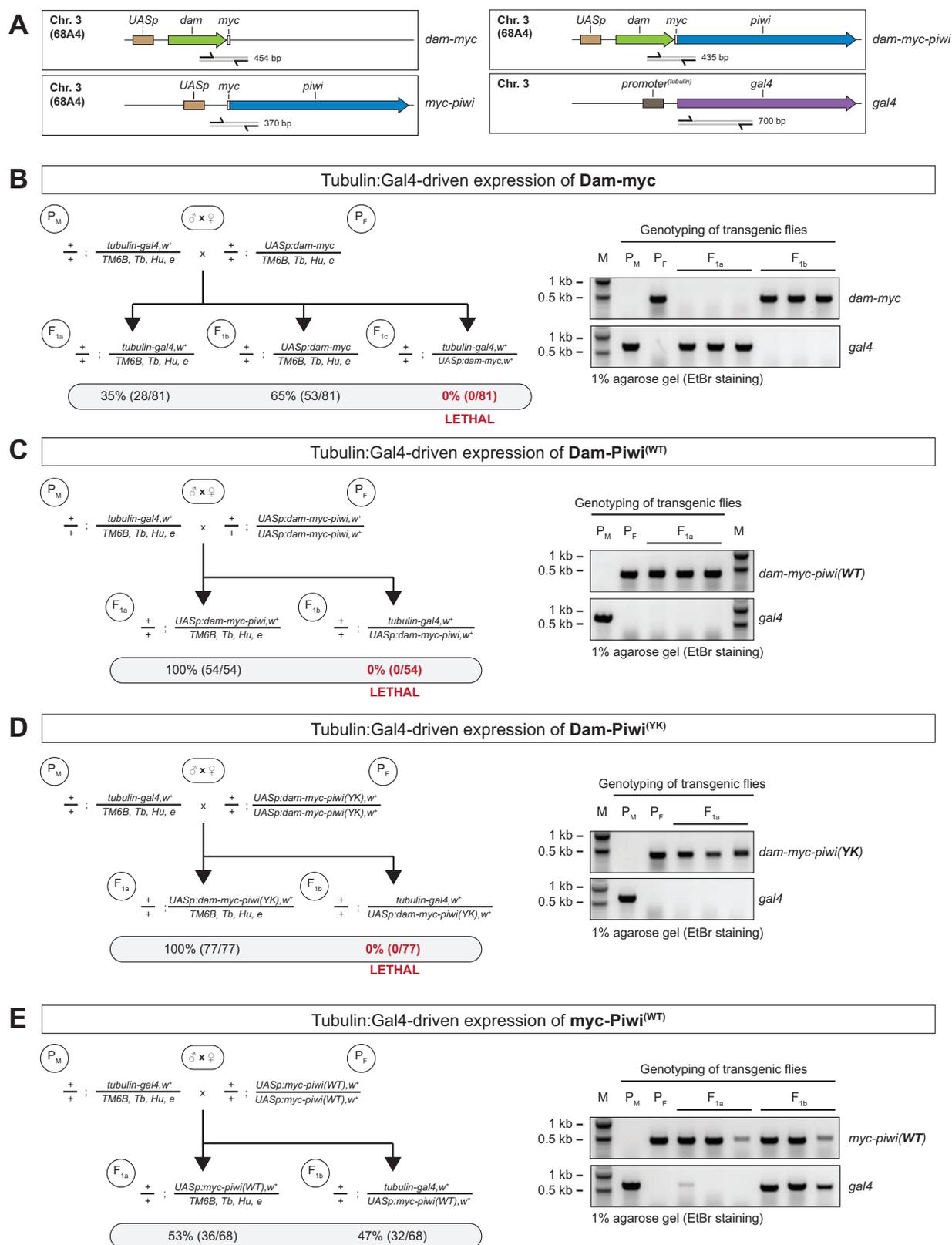
Figure 2 High concentration of Dam protein in *Drosophila* is embryonically lethal. (A) Schematic overview of transgenes containing *UASp: dam-myc*, *UASp: myc-piwi*(WT), *UASp: dam-myc-piwi*(WT), and *tubulin: gal4*. (B) Generation of *Dam-myc*-expressing flies under a *tubulin: gal4* promoter. Left panel: Crossing scheme of *tubulin: gal4* driver and *UASp: dam-myc* flies. Right panel: PCR genotyping of fly sublines using primers indicated in A (see also Supplementary Table S3). Tubulin-driven expression of *Dam-myc* was embryonically lethal. (C) Generation of *Dam-myc-piwi*(WT)-expressing flies under a tubulin promoter. Left panel: Crossing scheme of *tubulin: gal4* driver and *UASp: dam-myc-piwi*(WT) flies. Right panel: PCR genotyping of fly sublines using primers indicated in A. Tubulin-driven expression of *Dam-myc-piwi*(WT) was embryonically lethal. (D) Generation of *Dam-myc-piwi*(YK)-expressing flies under a *tubulin: gal4* promoter. Left panel: Crossing scheme of *tubulin: gal4* driver and *UASp: dam-myc-piwi*(YK) flies. Right panel: PCR genotyping of fly sublines using primers indicated in A. Tubulin-driven expression of *Dam-myc-piwi*(YK) was embryonically lethal. (E) Generation of *myc-Piwi*(WT)-expressing flies under a *tubulin: gal4* promoter. Left panel: Crossing scheme of *tubulin: gal4* driver and *UASp: myc-piwi*(WT) flies. Right panel: PCR genotyping of fly sublines using primers indicated in A.
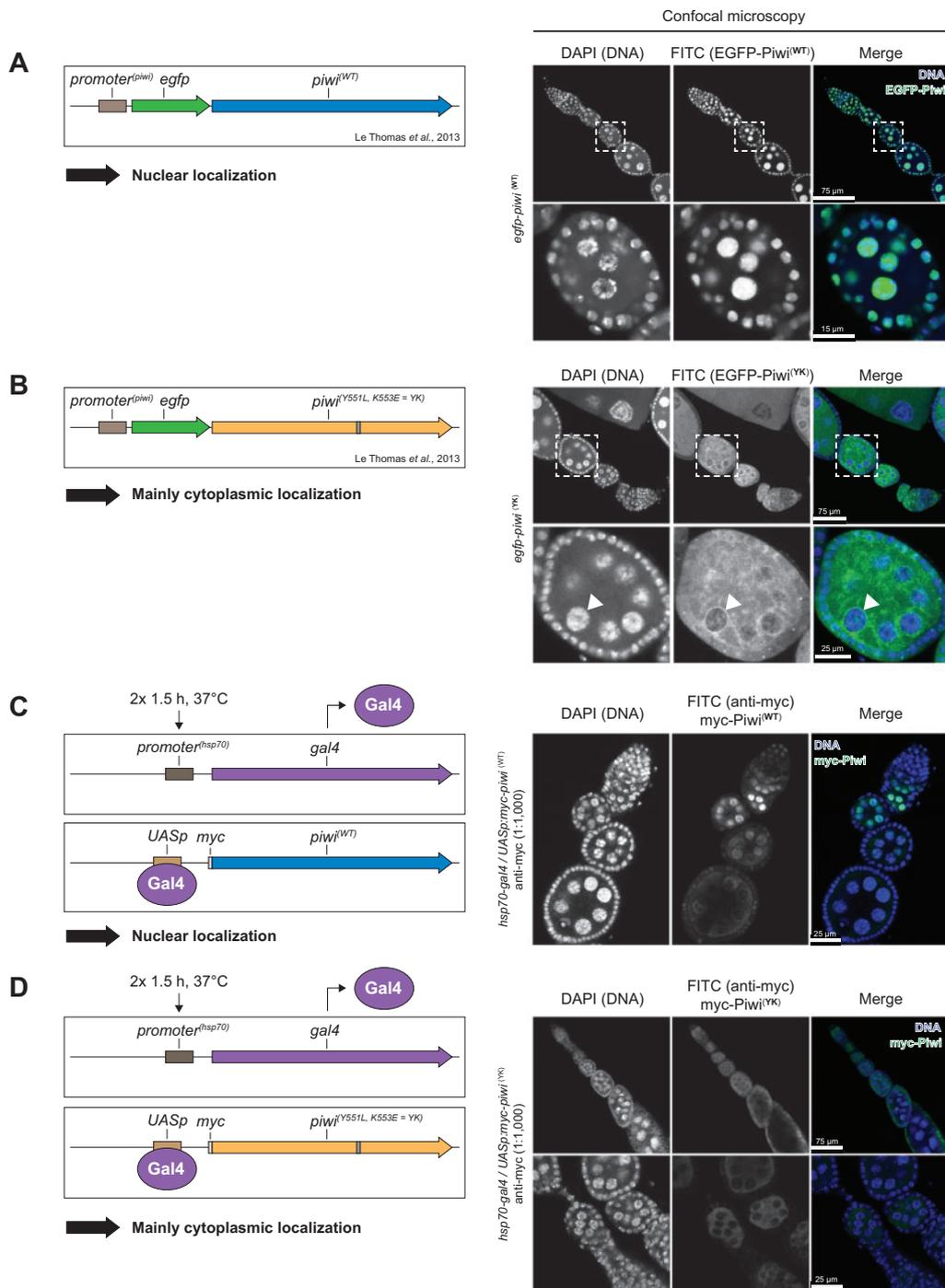
Figure 3 Disrupting piRNA binding causes cytoplasmic localization of Piwi. (A) Cellular localization of EGFP-tagged Piwi$^{(WT)}$. Left panel: Schematic overview of the transgene. Right panel: Confocal microscopy of a *Drosophila* ovariole expressing EGFP-Piwi$^{(WT)}$ under the *piwi* promoter. EGFP-Piwi$^{(WT)}$ is localized to the nuclei of somatic and germline cells. (B) Cellular localization of EGFP-tagged Piwi$^{(YK)}$. Left panel: Schematic overview of the transgene. Right panel: Confocal microscopy of a *Drosophila* ovariole expressing EGFP-Piwi$^{(YK)}$ under the *piwi* promoter. The YK mutation disrupts piRNA binding to Piwi and impairs the mutated Piwi protein from entering the nucleus. Arrowheads indicate the enrichment of Piwi$^{(YK)}$ at the outer nuclear membrane of nurse cell nuclei. (C) Cellular localization of Myc-tagged Piwi$^{(WT)}$. Left panel: Schematic overview of the transgene. A heat-shock promoter 70 (*hsp70*) is used to drive Gal4 expression which in turn causes the production of Myc-Piwi$^{(WT)}$. Right panel: Confocal microscopy of a *Drosophila* ovariole expressing Myc-Piwi$^{(WT)}$ following two rounds of heat-shock. Myc-Piwi$^{(WT)}$ is localized to the nucleus. (D) Cellular localization of Myc-tagged Piwi$^{(YK)}$. Left panel: Schematic overview of the transgene. A heat-shock promoter 70 (*hsp70*) promoter drives *myc-Piwi*$^{(YK)}$ expression. Right panel: Confocal microscopy of a *Drosophila* ovariole expressing *myc-piwi*$^{(YK)}$ following two rounds of heat-shock. Myc-Piwi$^{(YK)}$ is mainly localized to the cytoplasm.

libraries with a mapping rate of 59–67% to the *Drosophila* genome while allowing unique and multiple mapping (see Supplementary Table S1 for a detailed overview). On average 90% of the mapped reads could be assigned to unique genomic loci, whereas 10% could be mapped to more than one location in the genome. To avoid ambiguities that result from sequencing reads that could be mapped to several sites in the genome, we only used uniquely mapped reads in our downstream analysis to identify m⁶A-enriched sites using a previously described algorithm (iDEAR) (Gutierrez-Triana *et al.* 2016). The individual biological triplicates
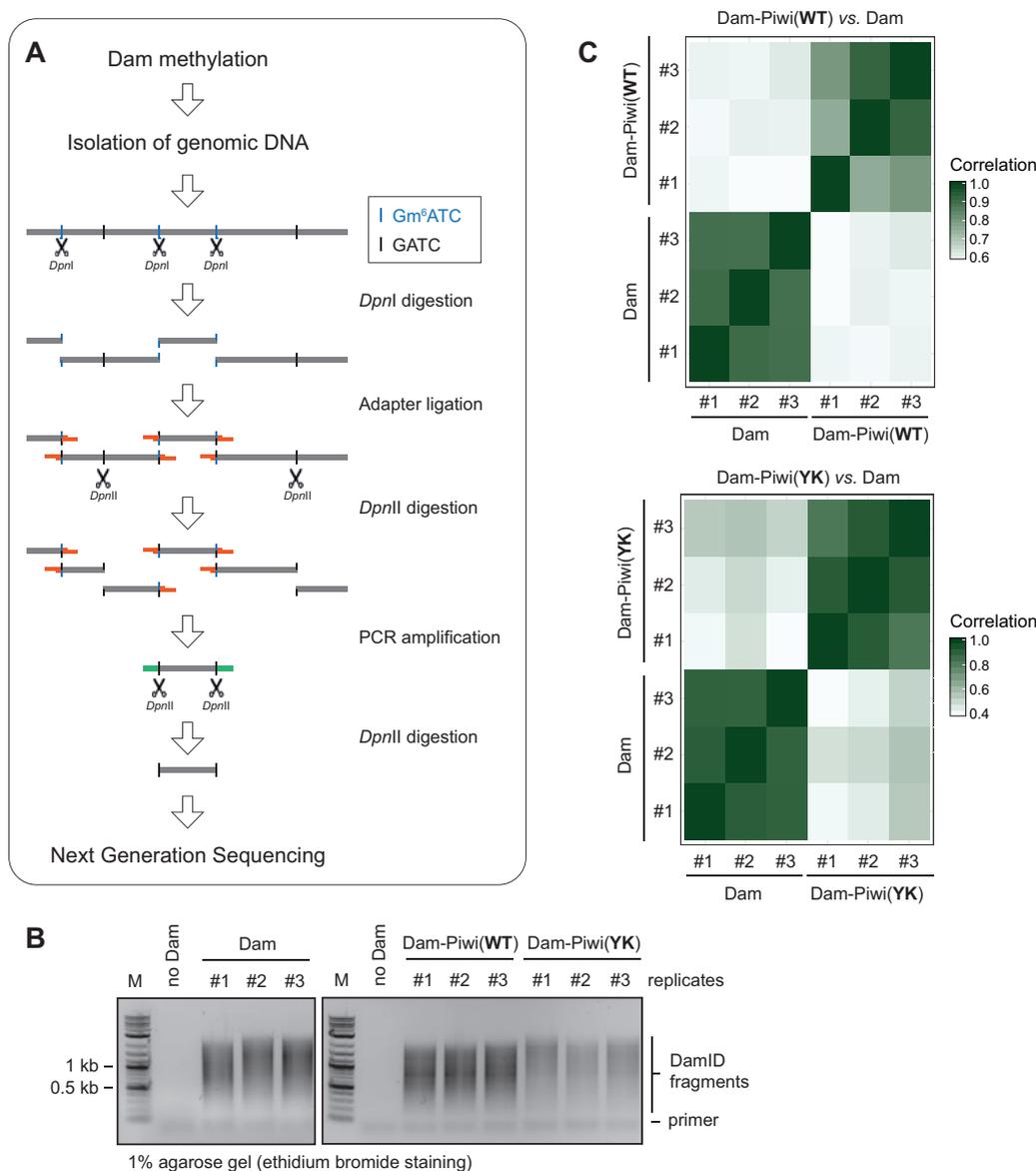
Figure 4 Piwi-DamID-seq. (A) Experimental outline of DamID-seq. (B) PCR amplification of GATC fragments. (C) Quality control analysis of Piwi-DamID samples using the iDEAR algorithm (Gutierrez-Triana *et al.* 2016).

of Dam-Piwi$^{(WT)}$ and Dam-Piwi$^{(YK)}$ strongly correlated with each other, but were distinct from the Dam only control (Figure 4C), indicating that our DamID-seq data were reproducible. The Piwi wild type and mutant sequencing results were compared to the Dam-only library, and regions identified by iDEAR with a *P*-value of $P \leq 0.01$ were collected as m^{6}A-enriched sites.

We reasoned that when we disrupted the Piwi-piRNA interaction in our DamID approach, genomic sites that are usually targeted by wild type Piwi should lose their m^{6}A methylation signature (Figure 5A). In total, we identified 7448 m^{6}A-enriched regions in Dam-Piwi$^{(WT)}$ and 8282 regions in Dam-Piwi$^{(YK)}$ (Figures 5B and 6A). When we compared the m^{6}A-enriched regions between Dam-Piwi$^{(WT)}$ and Dam-Piwi$^{(YK)}$ samples, we found that 4535 (60.9%) were unique to Dam-Piwi$^{(WT)}$ and 5340 (65.1%) were unique to Dam-Piwi$^{(YK)}$. Furthermore, 2913 (39.1%) regions identified in Dam-Piwi$^{(WT)}$ overlapped with regions in Dam-Piwi$^{(YK)}$ and 2942 (35.5%) *vice versa*, as individual regions could overlap with either one or more regions of the compared

data sets. Therefore, at least 4535 regions in the genome appeared to be targeted specifically by the Piwi-piRNA complex.

Recently, a study applied Piwi-DamID in the *Drosophila* ovary (Ilyin *et al.* 2017). As this study also covers whole ovarian tissue during the same developmental stage, we included its Piwi-DamID data for comparison. Following this comparison, 2089 out of 3941 (53%) m^{6}A-enriched regions in the Ilyin *et al.* (2017) were also present in our Piwi$^{(WT)}$-DamID (*P*-value < 0.001) (Figure 5B). Moreover, we found 5170 m^{6}A-enriched regions that were unique in our Dam-Piwi$^{(WT)}$ flies. Our Dam-Piwi$^{(YK)}$ sites also overlap but are sufficiently different from m^{6}A-enriched regions in the Ilyin *et al.* flies (Figure 5B).

Next, we examined whether Piwi binding shows any preference between euchromatin and heterochromatin. The female *Drosophila* genome is 139.4 Mb in size, of which 120.4 Mb (86.3%) are euchromatic and 18.9 Mb (13.7%) are heterochromatic regions (*Drosophila* genome release 5, FlyBase; Supplementary Table S2). In the Dam-only library, more than 98% of the unique reads fall
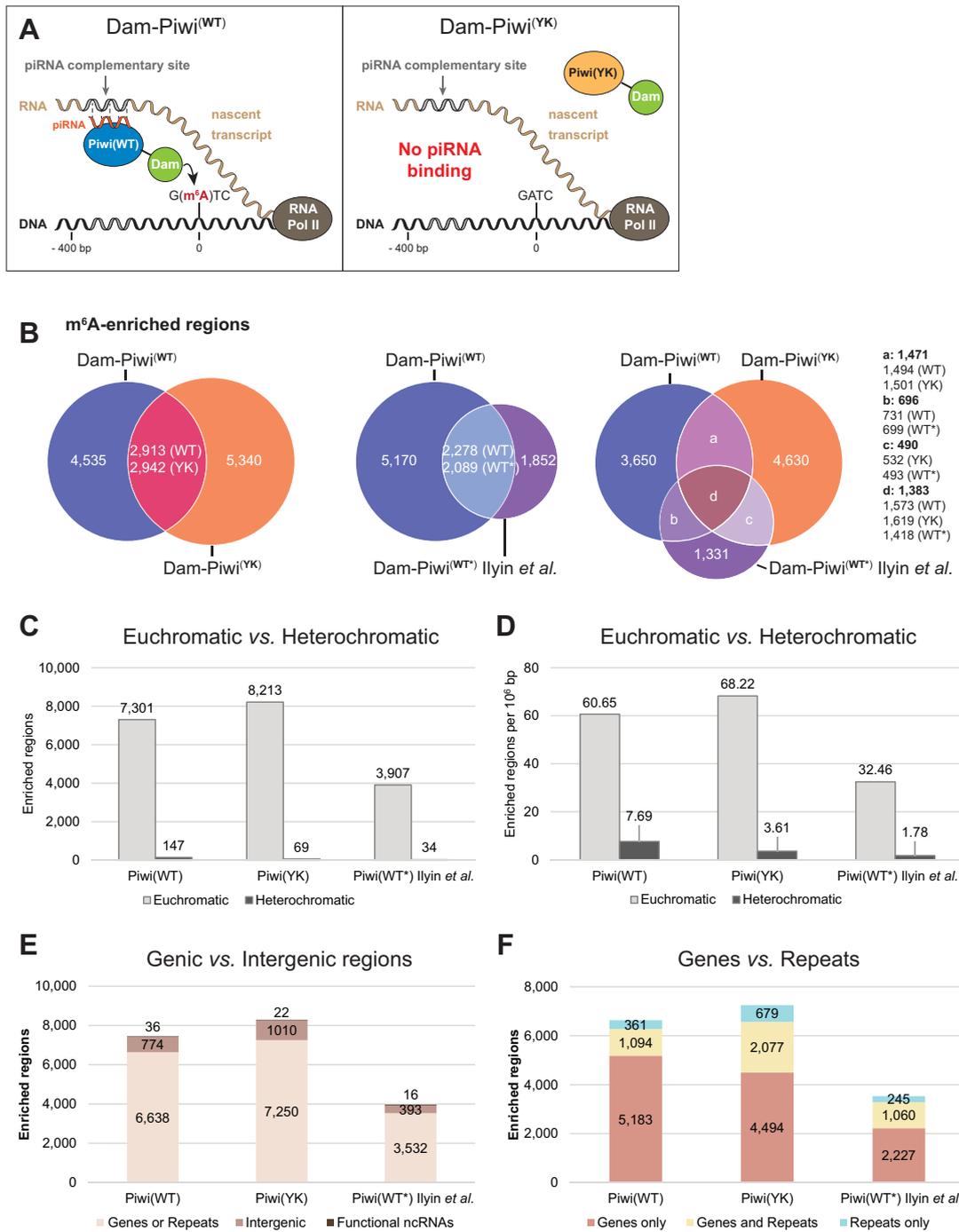
Figure 5 Piwi predominantly interacts with euchromatic regions in the *Drosophila* genome. (A) Schematic overview of genomic interaction of wild type Piwi and Piwi YK mutant deficient in piRNA binding. (B) Overlap of m⁶A-enriched regions in Dam-Piwi(WT) and Dam-Piwi(YK)-expressing ovaries, Dam-Piwi(WT) from this and a previous study (Ilyin *et al.* 2017), and all three samples. (C) Distribution of enriched methylation regions in euchromatin (light gray) and heterochromatin (dark gray). (D) Distribution of enriched methylation regions in euchromatin (light gray) and heterochromatin (dark gray) normalized to enriched regions per 1 million base pairs. (E) Distribution of enriched methylation regions in genic/repeats *vs.* intergenic regions. Genes and Repeats (light brown; including transposons): genomic regions encoding protein, including 5'UTR, introns, exons, 3'UTR, and 1 kb up/downstream of UTRs. Intergenic regions (medium brown): sequences outside of genic/repeats regions. Functional nonprotein-encoding RNAs (dark brown): tRNA, rRNA, and other noncoding RNA encoding regions. (F) Distribution of enriched methylation regions in genes only (red), repeats only (blue), and regions covering both genes and repeats (yellow).

into euchromatin (Supplementary Figure S1A). We normalized the reads against the total euchromatin and heterochromatin lengths. The density of DamID methylation in euchromatin is at least four times of that in heterochromatin (Supplementary Figure S1B). A similar bias for euchromatin was observed for Dam-Piwi(WT) and Dam-Piwi(YK) (Supplementary Figure S1). Reads

that could be mapped to multiple sites in the genome were analyzed similarly. There was a strong bias toward heterochromatin for multi-mapping reads (Supplementary Figure S2). However, this is only 10% of all mappable reads, so excluding this bias in the 10% reads will not change the overall conclusion.
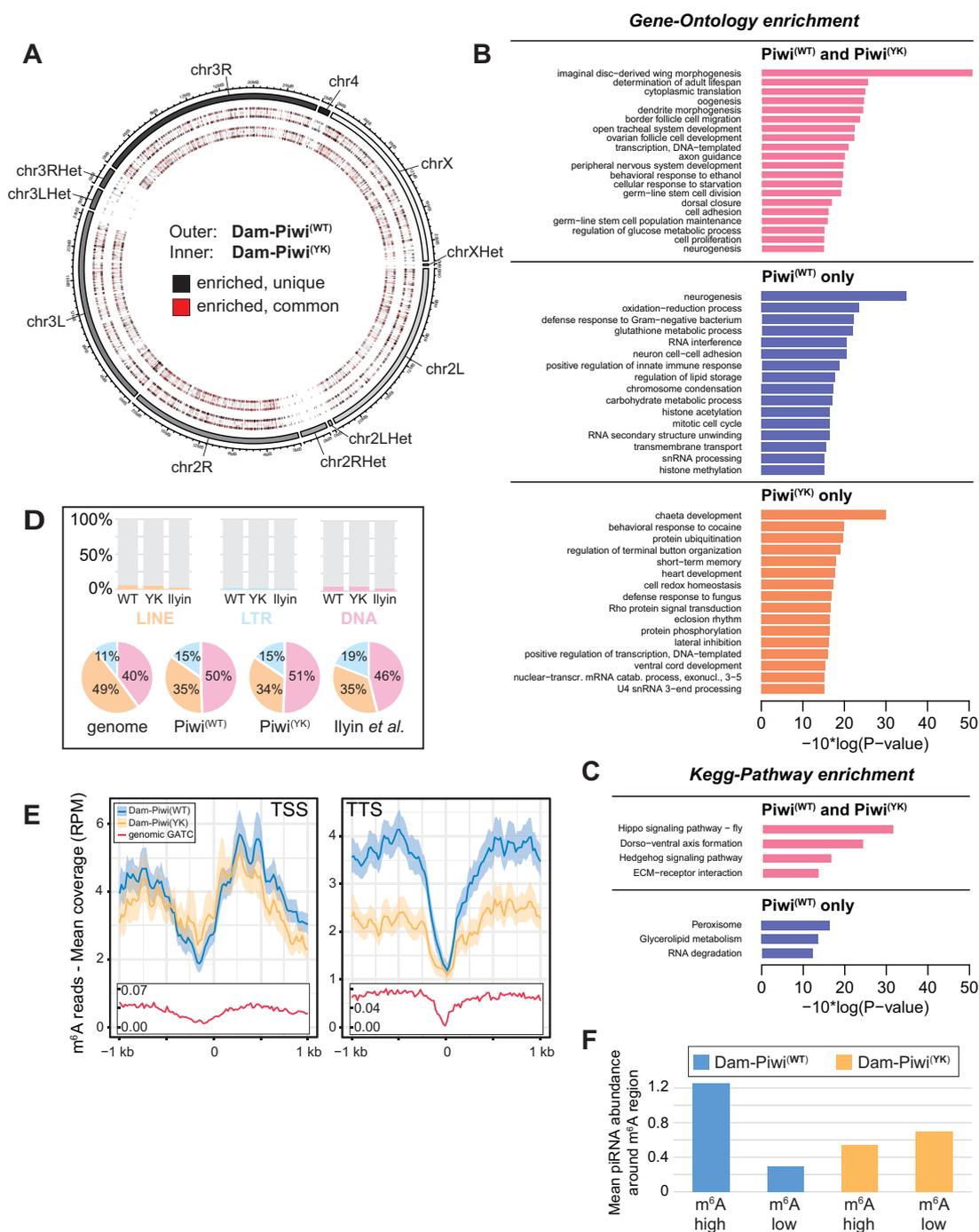
Figure 6 Genome-wide distribution of enriched methylation regions and their correlation to piRNA-target sites. (A) Circos plot of genome-wide distribution of m^{6}A-enriched regions in the *Drosophila* genome (version dm3) (Krzywinski *et al.* 2009). The Dam-Piwi$^{(WT)}$ track is depicted on the outer, the Dam-Piwi$^{(YK)}$ track on the inner circle. m^{6}A-enriched regions that are unique to each track are shown in black, common ones in red. (B) GO term enrichment of genes within 1 kb of methylation-enriched regions: common in Dam-Piwi$^{(WT)}$ and Dam-Piwi$^{(YK)}$ transgenic flies (pink); only in Dam-Piwi$^{(WT)}$ but not in Dam-Piwi$^{(YK)}$ transgenic flies (purple); only in Dam-Piwi$^{(YK)}$ but not in Dam-Piwi$^{(WT)}$ transgenic flies (orange). (C) KEGG pathway enrichment of genes within 1 kb of methylation-enriched regions: common in Dam-Piwi$^{(WT)}$ and Dam-Piwi$^{(YK)}$ transgenic flies (pink); only in Dam-Piwi$^{(WT)}$ but not in Dam-Piwi$^{(YK)}$ transgenic flies (purple). There was no KEGG pathway enrichment of genes that were observed only in Dam-Piwi$^{(YK)}$ but not in Dam-Piwi$^{(WT)}$ transgenic flies. (D) Enrichment of m^{6}A methylation in transposons. (E) Distribution of enriched m^{6}A methylation at the transcription start site (TSS) and the transcription termination site (TTS) of protein-coding genes in Dam-Piwi$^{(WT)}$ and Dam-Piwi$^{(YK)}$. (F) Mean piRNA abundance around genomic regions with increased (m^{6}A high) or decreased (m^{6}A low) m^{6}A methylation in Dam-Piwi$^{(WT)}$ and Dam-Piwi$^{(YK)}$.

As the distribution of GATC sites in the genome is discontinuous, we analyzed whether the modified GATC sites were enriched in simple repeats, transposon repeats, satellite repeats, low-complexity repeats, RNA repeats, or rRNA repeats. We found that 3.2–4.2% of euchromatic GATC sites fall in repeats, whereas 70.2–72.2% of heterochromatic GATC sites fall in repeats, particularly transposon repeats (Supplementary Figure S3). Overall, GATC sites are distributed in very similar densities between euchromatin (2.77 GATC sites/kb) and heterochromatin (2.64 GATC sites/kb). If these GATC sites were equally accessible to the Dam and

Dam-Piwi fusion proteins, we would expect similar densities of m^6A-enriched GATC sites between them. Our Piwi-DamID analysis, however, showed that almost all identified m^6A sites in Piwi[(WT)] and Piwi[(YK)] occurred in euchromatic regions (Figure 5C). Normalized to the number of base pairs in each region, there was a 7.6-fold increase of euchromatic to heterochromatic m^6A-enriched sites in Dam-Piwi[(WT)] and an 18.3-fold increase in Dam-Piwi[(YK)] (Figure 5D). This shows that the identified binding sites of both wild-type Piwi and the YK mutant are enriched within euchromatic regions of the *Drosophila* genome, possibly reflecting that euchromatin is more transcriptionally active or accessible to Piwi.

Piwi heavily associates with protein-coding genes

We next investigated what types of euchromatic sequences are targeted by Piwi. In the *Drosophila* genome, 69% of the DNA correspond to regions that harbor either protein-coding genes or transposable elements. The remaining 31% do not contain any known genetic elements. To identify what genomic regions are enriched with m^6A marks and therefore have Piwi binding close by, we mapped the identified m^6A-enriched regions with regard to genic regions, intergenic regions, and functional nonprotein-encoding RNA regions (tRNAs, rRNAs, and others; Figure 5E). The majority of m^6A signals was found in genes and repeats (including transposons). Unexpectedly, more m^6A signals were detected in protein-coding genes than in repeats (Figure 5F). This is surprising as Piwi has so far been mainly described to regulate transposable elements in the germline (Toth *et al.* 2016; Yamashiro and Siomi 2017). Normalization of enriched m^6A signal to the GATC density in genes, genes and repeats, and repeats resulted in an enrichment of methylation signal in the repeats (Supplementary Figure S4).

Regions with enriched m^6A methylation in Dam-Piwi[(WT)] and Dam-Piwi[(YK)] sample are distributed throughout the genome (Figure 6A), indicating regulated Piwi binding to many genomic sites. Specifically, m^6A-enriched regions identified by both our Dam-Piwi[(WT)] and Dam-Piwi[(YK)] DamID analysis are enriched with genes involved in development, such as imaginal disc-derived wing morphogenesis, ovarian follicle cell development, and open tracheal system development (Figure 6B, upper panel). KEGG pathway analysis also indicated that genes in these m^6A-enriched regions are involved in developmental pathways such as dorso-ventral axis formation and the Hippo and Hedgehog signaling pathways (Figure 6C). While Piwi[(WT)]-specific m^6A-containing regions (*i.e.*, in our Dam-Piwi[(WT)] mapping) are enriched in genes involved in neurogenesis and metabolism (Figure 6B, middle panel), and glycerolipid metabolism and RNA degradation as shown by KEGG pathway analysis (Figure 6C), Piwi[(YK)]-specific m^6A-containing regions were enriched in a variety of functions (Figure 6B, lower panel), yet no KEGG-pathway enrichment was found.

As Piwi is known to repress transposons in the germline, we were interested in whether any m^6A enrichment could be identified around transposons which would indicate co-transcriptional binding of Piwi to these transposon RNAs. In particular, we focused on LINE, LTR, and DNA transposons. Only a very small portion of transposon regions (± 1 kb) showed increased m^6A levels (LINE: ~6%, LTR: ~4%, DNA: ~7%) in Dam-Piwi wild type as well as YK mutant flies (Figure 6D, bar plots). The distribution of these transposons was very similar to their overall occurrence in the genome (Figure 6D, pie charts). However, DNA transposons were slightly over-represented among transposons that obtained m^6A methylation. Highly similar results were observed for Dam-

Piwi[(WT)] and Dam-Piwi[(YK)] (Figure 6D, pie charts). In summary, Piwi was mainly associated with protein-coding genes instead of transposon sequences. This indicates that Piwi might play a broad role in the transcriptional regulation of protein-coding genes.

To further characterize Piwi binding to protein-coding genes, we examined the overall distribution of m^6A signals along transcripts of coding genes (Figure 6E). In particular, we studied m^6A deposition in Dam-Piwi[(WT)] (blue signal) and Dam-Piwi[(YK)] mutants (yellow signal) around the transcription start site (TSS, ±1 kb) as well as the transcription termination site (TTS, ±1 kb). In addition, we plotted the overall distribution of possible methylation sites (GATC motifs) along the same regions (red signal). Although the methylation signal around the TSS was similar in both Dam-Piwi[(WT)] and mutant flies, we found an increase of m^6A signal with ±1 kb of the TTS in the Dam-Piwi[(WT)] flies, indicating that Piwi might be guided to these regions by its associated piRNA.

Piwi appears to be guided to specific genomic sites by its associated piRNAs

To investigate if Piwi was indeed guided by its associated piRNA to piRNA-complementary genomic sites, we examined whether there was an increased number of piRNA binding sites around m^6A-enriched regions in Dam-Piwi[(WT)] samples. Genomic regions with a statistically significant ($P \leq 0.01$) increase in m^6A methylation according to the iDEAR algorithm (Gutierrez-Triana *et al.* 2016) were defined as "m^6A high" and regions devoid of m^6A signal were defined as "m^6A low". We first identified "m^6A high" and "m^6A low" regions in both the Dam-Piwi[(WT)] (7448 m^6A high, 3213 m^6A low) and Piwi[(YK)] mutant (8282 m^6A high, 8434 m^6A low) datasets. Next, we selected a known set of piRNA sequences from *Drosophila* that have been identified in a previous study with perfect match to the genomic sequence (Yin and Lin 2007; 16,049 piRNA species, 74,013 reads) and allocated the piRNA reads within a distance of ±400 bp of the m^6A-enriched region. Average sizes of the m^6A-enriched regions were 540 bp in Piwi[(WT)] and 823 bp in Piwi[(YK)] mutants. As some piRNAs could be mapped to several sites in the genome, we normalized the piRNA abundance at each m^6A region with respect to their genome-wide occurrence. In Dam-Piwi[(WT)] flies we identified 1.25 piRNA reads in "m^6A high" and 0.29 piRNA reads in "m^6A low" regions (Figure 6F). When we disrupted piRNA binding to Piwi by using the Dam-Piwi[(YK)] mutant, the number of piRNA reads in "m^6A high" regions decreased to 0.54. In "m^6A low" regions the number of reads increased to 0.7. Hence, these data support the notion that the Piwi protein is guided to genomic binding sites by its associated piRNAs. Piwi proteins devoid of piRNAs lose their ability to bind to original sites and instead interact with new targets in the genome. Consequently, even though some of the Dam-Piwi[(YK)] was able to enter the nucleus, it could not find the sites corresponding to its associated piRNA, but instead interact randomly with further sites. Our findings do not exclude additional Piwi binding to genomic sites by other means such as via direct or indirect binding of other cellular factors.

Conclusion

In this study, we mapped Piwi-binding sites in the genome of *Drosophila* ovarian germline and somatic cells using the DamID-seq method. This method captured thousands of Piwi-binding sites, with >90% located in euchromatin. Furthermore, we found an enrichment of Piwi-dependent m^6A sites close to piRNA-

complementary sites. Disruption of the Piwi-piRNA interaction significantly reduced the binding of Piwi around the transcription termination site. These findings support the hypothesis that Piwi is guided to specific genomic sites by its associated piRNA. Remarkably, most Piwi binding sites detected by DamID are in protein-coding genes, especially at the 5' and 3' ends. This is a surprising finding, as previous studies have suggested that the primary role of Piwi in the genome is the repression of transposons. Our study may therefore point to a possibly new and major function of Piwi in regulating gene expression supporting the growing body of evidence showing that Piwi regulates genes as well as transposons.

Acknowledgments

Funding

Literature cited

Aravin A, Gaidatzis D, Pfeffer S, Lagos-Quintana M, Landgraf P, *et al.* 2006. A novel class of small RNAs bind to MILI protein in mouse testes. Nature. 442:203–207.

Bartel DP. 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell. 116:281–297.

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

Brower-Toland B, Findley SD, Jiang L, Liu L, Yin H, *et al.* 2007. *Drosophila* PIWI associates with chromatin and interacts directly with HP1a. Genes Dev. 21:2300–2311.

Cox DN, Chao A, Baker J, Chang L, Qiao D, *et al.* 1998. A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal. Genes Dev. 12:3715–3727.

Cox DN, Chao A, Lin H. 2000. piwi encodes a nucleoplasmic factor whose activity modulates the number and division rate of germline stem cells. Development. 127:503–514.

Darricarrere N, Liu N, Watanabe T, Lin H. 2013. Function of Piwi, a nuclear Piwi/Argonaute protein, is independent of its slicer activity. Proc Natl Acad Sci USA. 110:1297–1302.

Gangaraju VK, Yin H, Weiner MM, Wang J, Huang XA, *et al.* 2011. Drosophila Piwi functions in Hsp90-mediated suppression of phenotypic variation. Nat Genet. 43:153–158.

Girard A, Sachidanandam R, Hannon GJ, Carmell MA. 2006. A germline-specific class of small RNAs binds mammalian Piwi proteins. Nature. 442:199–202.

Gloor GB, Preston CR, Johnson-Schlitz DM, Nassif NA, Phillis RW, *et al.* 1993. Type I repressors of P element mobility. Genetics. 135:81–95.

Gonzalez J, Qi H, Liu N, Lin H. 2015. Piwi is a key regulator of both somatic and germline stem cells in the Drosophila testis. Cell Rep. 12:150–161.

Grewal SI. 2010. RNAi-dependent formation of heterochromatin and its diverse functions. Curr Opin Genet Dev. 20:134–141.

Grivna ST, Beyret E, Wang Z, Lin H. 2006. A novel class of small RNAs in mouse spermatogenic cells. Genes Dev. 20:1709–1714.

Gutierrez-Triana JA, Mateo JL, Ibberson D, Ryu S, Wittbrodt J. 2016. iDamIDseq and iDEAR: an improved method and computational pipeline to profile chromatin-binding proteins. Development. 143:4272–4278.

Holoch D, Moazed D. 2015. RNA-mediated epigenetic regulation of gene expression. Nat Rev Genet. 16:71–84.

Huang XA, Yin H, Sweeney S, Raha D, Snyder M, *et al.* 2013. A major epigenetic programming mechanism guided by piRNAs. Dev Cell. 24:502–516.

Ilyin AA, Ryazansky SS, Doronin SA, Olenkina OM, Mikhaleva EA, *et al.* 2017. Piwi interacts with chromatin at nuclear pores and promiscuously binds nuclear transcripts in Drosophila ovarian somatic cells. Nucleic Acids Res. 45:7666–7680.

Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, *et al.* 2009. Circos: an information aesthetic for comparative genomics. Genome Res. 19:1639–1645.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods. 9:357–359.

Lau NC, Seto AG, Kim J, Kuramochi-Miyagawa S, Nakano T, *et al.* 2006. Characterization of the piRNA complex from rat testes. Science. 313:363–367.

Le Thomas A, Rogers AK, Webster A, Marinov GK, Liao SE, *et al.* 2013. Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. Genes Dev. 27:390–399.

Lin H, Chen M, Kundaje A, Valouev A, Yin H, *et al.* 2015. Reassessment of Piwi binding to the genome and Piwi impact on RNA polymerase II distribution. Dev Cell. 32:772–774.

Lin H, Spradling AC. 1997. A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the Drosophila ovary. Development. 124:2463–2476.

Lin H, Yin H. 2008. A novel epigenetic mechanism in Drosophila somatic cells mediated by Piwi and piRNAs. Cold Spring Harb Symp Quant Biol. 73:273–281.

Marinov GK, Wang J, Handler D, Wold BJ, Weng Z, *et al.* 2015. Pitfalls of mapping high-throughput sequencing data to repetitive sequences: Piwi's genomic targets still not identified. Dev Cell. 32:765–771.

Pal-Bhadra M, Leibovitch BA, Gandhi SG, Chikka MR, Bhadra U, *et al.* 2004. Heterochromatic silencing and HP1 localization in Drosophila are dependent on the RNAi machinery. Science. 303:669–672.

Peng JC, Valouev A, Liu N, Lin H. 2016. Piwi maintains germline stem cells and oogenesis in *Drosophila* through negative regulation of polycomb group proteins. Nat Genet. 48:283–291.

Post C, Clark JP, Sytnikova YA, Chirn GW, Lau NC. 2014. The capacity of target silencing by *Drosophila* PIWI and piRNAs. RNA. 20:1977–1986.

Rozhkov NV, Hammell M, Hannon GJ. 2013. Multiple roles for Piwi in silencing *Drosophila* transposons. Genes Dev. 27:400–412.

Saito K. 2013. The epigenetic regulation of transposable elements by PIWI-interacting RNAs in Drosophila. Genes Genet Syst. 88:9–17.

Saito K, Inagaki S, Mituyama T, Kawamura Y, Ono Y, *et al.* 2009. A regulatory circuit for piwi by the large Maf gene traffic jam in Drosophila. Nature. 461:1296–1299.

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

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

Soshnev AA, Josefowicz SZ, Allis CD. 2016. Greater than the sum of parts: complexity of the dynamic epigenome. Mol Cell. 62:681–694.

Sytnikova YA, Rahman R, Chirn GW, Clark JP, Lau NC. 2014. Transposable element dynamics and PIWI regulation impacts lncRNA and gene expression diversity in *Drosophila* ovarian cell cultures. Genome Res. 24:1977–1990.

Thomson T, Lin H. 2009. The biogenesis and function of PIWI proteins and piRNAs: progress and prospect. Annu Rev Cell Dev Biol. 25:355–376.

Toth KF, Pezic D, Stuwe E, Webster A. 2016. The piRNA pathway guards the germline genome against transposable elements. Adv Exp Med Biol. 886:51–77.

van Steensel B, Henikoff S. 2000. Identification of *in vivo* DNA targets of chromatin proteins using tethered dam methyltransferase. Nat Biotechnol. 18:424–428.

Vogel MJ, Peric-Hupkes D, van Steensel B. 2007. Detection of in vivo protein-DNA interactions using DamID in mammalian cells. Nat Protoc. 2:1467–1478.

Watanabe T, Takeda A, Tsukiyama T, Mise K, Okuno T, *et al.* 2006. Identification and characterization of two novel classes of small RNAs in the mouse germline: retrotransposon-derived siRNAs in oocytes and germline small RNAs in testes. Genes Dev. 20: 1732–1743.

Yamashiro H, Siomi MC. 2017. PIWI-interacting RNA in Drosophila: biogenesis, transposon regulation, and beyond. Chem Rev. 118: 4404-4421.

Yashiro R, Murota Y, Nishida KM, Yamashiro H, Fujii K, *et al.* 2018. Piwi nuclear localization and its regulatory mechanism in Drosophila ovarian somatic cells. Cell Rep. 23:3647–3657.

Yin H, Lin H. 2007. An epigenetic activation role of Piwi and a Piwi-associated piRNA in *Drosophila melanogaster*. Nature. 450: 304–308.

Zhang G, Huang H, Liu D, Cheng Y, Liu X, *et al.* 2015. N6-methyladenine DNA modification in Drosophila. Cell. 161: 893–906.

Communicating editor: B. J. Andrews