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 Lin1,2,*

1Yale Stem Cell Center, Yale University School of Medicine, New Haven, CT 06520-8073, USA
2Department 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; Sosheev 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-sequencing 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; Sienkna et al. 2012; Huang et al. 2013). Several other studies further illustrated that Piwi is guided to nascent transcripts by its associated piRNA...
To study how extensively Piwi is involved in epigenetic regulation and to what extent the regulation depends on piRNAs in the Drosophila ovariary, 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). However, piRNA-independent binding of Piwi to a small number of specific genomic sites has also been observed (Peng et al. 2016).

Fly stocks and generation of transgenic flies
Cultures were reared at 25°C on standard molasses agar medium. The fly stocks Bl2077 (w+; F[w+[mc]=Gal4-Hsp70.PB]), carrying the heat-shock-inducible gal4 transgene, BL3703 [w1118/+; Dp(1; Y); Y+]; CyO/Hub1 b1 So h1stu1; MKRS/TM6B, Tb1], a double balancer line, and BL5138 [v1 w7]; P[w-mC=ubiq-GAL4]L7/TM3, Sb1 Ser1], expressing ubiquitous Gal4 under the alphaTub84B promoter, were obtained from the Bloomington Drosophila Stock Center at Indiana University.

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; 35 cycles of 95°C, 30 s; 55°C, 30 s; 72°C, 30 s × 34; 72°C, 30 s; 12°C, infinity. PCR reactions were separated on a 1% agarose gel.

Immunofluorescence of Drosophila ovaries
Transgenic flies expressing EGF-Piwi(WT) and EGF-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% (v/v) p-formaldehyde 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 EGF-Piwi samples, the antibody-staining steps were omitted as the transgenes already carried a fluorescent marker. Samples were imaged using a Leica TCS SPS 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\(^{\text{WT}}\), and Dam-Piwi\(^{\text{YK}}\) 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 100μl of buffer A [buffer A: 10 mM Tris/HCl (pH 7.5), 60 mM NaCl, 10 mM EDTA, 0.2 mg/ml proteinase K] and ground with a sterile pestle.

400μl of buffer A and 500μl of buffer B [buffer B: 200 mM Tris/HCl (pH 7.5), 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 500μl 1× Tris/EDTA (TE) buffer. The sample was treated with 100μg/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 RNA at 37°C for 30 min. Following another phenol/chloroform extraction, the pellet was resuspended in 30μl 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), 2.5μg 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. 40pmol of adapters were added to the digested genomic DNA and ligated at 16°C for 2 h (T4 DNA ligase, NEB) before the ligation 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 50μl) was amplified by Q5 DNA polymerase (NEB, DamID_PCR): 98°C, 30s; (98°C, 10s, 55°C, 20s, 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 ddH₂O before removing the DamID adapters by DpnII digestion (10 units, NEB). Following a final round of column purification, 10ng of the eluted DNA fragments was used in Illumina library preparation (Illumina TrueSeq 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 1kb flanking region was applied for the annotation. For euchromatic vs heterochromatic annotation, chrXHet, chr2LHet, 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/g3.13347/677. 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/g3.13347/677.

**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,
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 N² atom of the adenosine residue (m²A) in nearby GATC sequences in the DNA (van Steensel and Henikoff 2000). A previous study reported that only very low levels of endogenous m²A exist in the Drosophila genome (Zhang, Yashiro, and Henikoff 2018), 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(WT) 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.

**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 promoter sequences are depicted in scale, whereas the UASp, hsp70, and tubulin promoter sequences are not.

**Figure 2** Schematic of experimental design. (A) wild-type EGFP-Piwi expression via 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(WT) 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 DpnII 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 wild-type control flies (Figure 4B), indicating that the transgenic Dam protein caused m²A formation in the Drosophila ovary and that any endogenous m²A 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 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.
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), 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: gal4 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.
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 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.
of Dam-Piwi\(^{\text{WT}}\) and Dam-Piwi\(^{\text{YK}}\) 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 \(P < 0.01\) were collected as m6A-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 m6A methylation signature (Figure 5A). In total, we identified 7448 m6A-enriched regions in Dam-Piwi\(^{\text{WT}}\) and 8282 regions in Dam-Piwi\(^{\text{YK}}\) (Figures 5B and 6A). When we compared the m6A-enriched regions between Dam-Piwi\(^{\text{WT}}\) and Dam-Piwi\(^{\text{YK}}\) samples, we found that 4535 (60.9%) were unique to Dam-Piwi\(^{\text{WT}}\) and 5340 (65.1%) were unique to Dam-Piwi\(^{\text{YK}}\). Furthermore, 2913 (39.1%) regions identified in Dam-Piwi\(^{\text{WT}}\) overlapped with regions in Dam-Piwi\(^{\text{YK}}\) 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%) m6A-enriched regions in the Ilyin et al. (2017) were also present in our Piwi\(^{\text{WT}}\)-DamID (\(P\)-value < 0.001) (Figure 5B). Moreover, we found 5170 m6A-enriched regions that were unique in our Dam-Piwi\(^{\text{WT}}\) flies. Our Dam-Piwi\(^{\text{YK}}\) sites also overlap but are sufficiently different from m6A-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
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) (Krywinski et al. 2009). The Dam-Piwi(WT) track is depicted on the outer, and 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 (orange). 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
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 corresponds 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\(^6\)A marks and therefore have Piwi binding close by, we mapped the identified m\(^6\)A-enriched regions with regard to genomic regions, intergenic regions, and functional non-protein-encoding RNA regions (tRNAs, rRNAs, and others; Figure 5E). The majority of m\(^6\)A signals was found in genes and repeats (including transposons). Unexpectedly, more m\(^6\)A 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 transposons. This is due to the fact that only 3% of the m\(^6\)A-enriched regions are distributed throughout the genome (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 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\(^6\)A-enriched regions in Dam-Piwi(WT) samples. Genomic regions with a statistically significant (P ≤ 0.01) increase in m\(^6\)A methylation according to the iDEAR algorithm (Gutierrez-Triana et al. 2016) were defined as “m\(^6\)A high” and regions devoid of m\(^6\)A signal were defined as “m\(^6\)A low”. We first identified “m\(^6\)A high” and “m\(^6\)A low” regions in both the Dam-Piwi(WT) 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\(^6\)A 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.

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\(^6\)A 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**

We thank Dr. Shangpin Guo for suggesting the use of the DamID method, Dr. Jie Yao for the Dam-vector, and Dr. Lynn Cooley for the UASp-Gateway cloning vector. We also thank Lauren Gonzalez, Tianshu Liu, Dr. Laurel Lorenz, Dr. Julie Park, Dr. Hongying Qi, Dr. Xionghuo Tang, and Dr. Yiyang Yang for valuable comments. N.N., N.L., and H.L. designed the research. N.N. planned and conducted the experiments. M.Z. carried out the DNA sequencing. N.L. performed the bioinformatic analysis. N.L., N.N., and H.L. analyzed the data. N.N. wrote the manuscript. N.N., H.L., and N.L. revised the manuscript. H.L. supervised the research.

**Funding**

The DamID work was supported by the NIH (DP1CA174418) and the Mathers Foundation to HL. The sequencing service was conducted at the Yale Stem Cell Center Genomics Core facility that was supported by the Connecticut Regenerative Medicine Research Fund and the Li Ka Shing Foundation.

**Conflicts of interest:** None declared.

**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 HP1α. 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 germ-line 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-Schultz DM, Nassif NA, Phillips 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, Mkhalева 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-Miyawaga 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. 2007. Discrete piRNA clusters identify discrete regulatory units that control germ-line stem cells and oogenesis in Drosophila. Genes Dev. 21:2463–2476.

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. Development. 131:150–161.

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.

Rozhkova 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 IncRNA 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