Published online 20 December 2012 *Nucleic Acids Research, 2013, Vol. 41, No. 3* ***1581–1590***
doi:10.1093/nar/gks1275

The comprehensive epigenome map of piRNA clusters

Shinpei Kawaoka[1,*], Kahori Hara[1], Keisuke Shoji[1], Maki Kobayashi[2,3], Toru Shimada[1], Sumio Sugano[2], Yukihide Tomari[2,3], Yutaka Suzuki[2] and Susumu Katsuma[1,*]

[1]Department of Agricultural and Environmental Biology, Graduate School of Agricultural and Life Sciences, University of Tokyo, Yayoi 1-1-1, Bunkyo-ku, Tokyo 113-8657, [2]Department of Medical Genome Sciences, Graduate School of Frontier Sciences, University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639 and [3]Institute of Molecular and Cellular Biosciences, Yayoi 1-1-1, Bunkyo-ku, Tokyo 113-8657, Japan

Received March 16, 2012; Revised November 4, 2012; Accepted November 7, 2012

ABSTRACT

PIWI-interacting RNA (piRNA) clusters act as anti-transposon/retrovirus centers. Integration of selfish jumping elements into piRNA clusters generates de novo piRNAs, which in turn exert *trans*-silencing activity against these elements in animal gonads. To date, neither genome-wide chromatin modification states of piRNA clusters nor a mode for piRNA precursor transcription have been well understood. Here, to understand the chromatin landscape of piRNA clusters and how piRNA precursors are generated, we analyzed the transcriptome, transcription start sites (TSSs) and the chromatin landscape of the BmN4 cell line, which harbors the germ-line piRNA pathway. Notably, our epigenomic map demonstrated the highly euchromatic nature of unique piRNA clusters. RNA polymerase II was enriched for TSSs that transcribe piRNA precursors. piRNA precursors possessed 5′-cap structures as well as 3′-poly A-tails. Collectively, we envision that the euchromatic, opened nature of unique piRNA clusters or piRNA cluster-associated TSSs allows piRNA clusters to capture new insertions efficiently.

INTRODUCTION

Expanding invasion of selfish genetic elements called transposons threats the host genome especially in germ-line cells, as undesirable mutations caused by transposition in germ-line cells can be transmitted into the next generation (1–4). Likewise exogenous retroviruses can endanger a host genome. To adequately suppress these invasive elements, eukaryotic genomes encode anti-transposon/retrovirus loci, to which integration of exogenous elements triggers sequence-specific *trans*-silencing in the following generation. Previously, fly genetics identified fly anti-transposon/retrovirus loci such as *X-TAS* and *flamenco* (5–9).

These anti-transposon/retrovirus loci are now understood to be capable of generating PIWI-interacting RNAs (piRNAs) and hence are called piRNA clusters (1,3,4,10). piRNA clusters provide single-stranded piRNA precursors with yet unknown structures that are converted into mature piRNAs through a 3′–5′-exonucleolytic trimming reaction followed by 2′-*O*-methylation at the 3′-terminus (10–18). piRNAs are often complementary to transposons and thereby are capable of guiding PIWI subfamily proteins to degrade transposon RNAs (19–23).

Cytological analyses found that fly piRNA clusters are often heterochromatic; Brennecke *et al.* (10) identified 135 out of 142 piRNA clusters to be present in peri-centromeric and telomeric heterochromatin. Moreover, the piRNA-guided effector protein Piwi co-localized with H3K9me3 (histone H3 harboring trimethylated lysine 9), a representative heterochromatin mark, but not with H3K4me3, a characteristic euchromatin mark (24,25). Furthermore, Piwi binds the heterochromatin-associated protein HP1a (24), which is genetically required for *trans*-silencing against transposons in the fly gonads (26). HP1d/Rhino contributes to cluster transcription from a set of piRNA clusters such as *42AB* piRNA cluster (27). Recently, it has been shown that H3K9me3 is required for transcription of fly piRNA clusters (28). Loss of dSETDB1, a methylase responsible for establishing H3K9me3, disrupts piRNA production

*To whom correspondence should be addressed. Tel: +81 3 5841 8994; Fax: +81 3 5841 8993; Email: kawaokashinpei@gmail.com or skawaoka@cshl.edu
Correspondence may also be addressed to Susumu Katsuma. Tel: +81 3 5841 8994; Fax: +81 3 5841 8993; Email: katsuma@ss.ab.a.u-tokyo.ac.jp
Present address:
Shinpei Kawaoka, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724, USA.

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

and causes transposon de-silencing. In addition, the mutations in a set of piRNA pathway genes lead to chromatin opening followed by transposon de-silencing (29). In contrast to these reports showing that most piRNA clusters are heterochromatic, *3R-TAS* piRNA cluster has nucleosomes with euchromatin marks such as H3K4me3 (30). To date, genome-wide comprehensive mapping of chromatin modification states for piRNA clusters has not been described, and thus the exact nature of piRNA precursor transcription has not been well understood.

Chromatin immunoprecipitation followed by deep sequencing (ChIP-seq) or polymerase chain reaction (PCR) (ChIP-PCR) are the most straightforward methods to analyze chromatin modification states. However, these strategies cannot determine whether a modified histone is actually derived from a specific cell type when original samples are composed of different types of cells (31). Thus, it is important to use uniform cell populations. Testis and ovary samples used for piRNA studies usually contain substantial amount of somatic tissues. In contrast, the silkworm ovary-derived BmN4 cell line is an excellent model for studying the chromatin landscape of piRNA clusters, because BmN4 cells are uniform and harbor a fully functional, ping-pong capable germ-line piRNA pathway (32).

Here, to understand how piRNA clusters are epigenetically controlled and how piRNA precursors are born, we performed genome-wide analyses on the transcriptome, transcription start sites (TSSs) and chromatin modification states of BmN4 cells. We first determined which domains were actively transcribed in the BmN4 genome with the aid of RNA-seq analysis. By combining ChIP-seq with RNA-seq and TSS-seq, we next revealed which histone modifications define euchromatin and heterochromatin. Finally, we integrated and applied our findings to characterize the genomic nature of piRNA clusters. Our analyses uncovered the epigenomic states of piRNA clusters and the nature of piRNA precursor transcription.

MATERIALS AND METHODS

Cell culture

BmN4 cells were cultured at 27°C in IPL-41 medium (Applichem) supplemented with 10% fetal bovine serum. Stable cell lines expressing Flag-BmHP1a and Flag-BmHP1b were generated as described previously by using pIZ vector as a backbone (32).

RNA extraction

Total RNAs were extracted by using Trizol reagent (Invitrogen) or miRVana miRNA isolation kit (Ambion) according to the manufacture's instruction.

RNA-seq

Using 1 μg of total RNAs, RNA-seq library was constructed using mRNA Seq Sample Preparation Kit according to the manufacturer's instructions (Illumina). Briefly, RNA was subjected to polyA selection using Sera-Mag Magnetic Oligo-dT Beads. polyA+ RNA was partially degraded by incubating in fragmentation buffer at 94°C for 5 min. First strand cDNA was synthesized using random primer and SuperScript II (Invitrogen) and second strand cDNA was synthesized using RNaseH and DNA pol I (Illumina). Double-stranded cDNA was size fractionated by 6% polyacrylamide gel electrophoresis (PAGE) and cDNAs of 250–300 bp were recovered. Illumina GA sequencing adaptors were ligated to cDNA ends. cDNAs were amplified by 15 cycles of PCR reactions, using Phusion DNA Polymerase (Finnzymes). Thirty-six base pair single-end-read RNA-seq tags were generated using an Illumina GA sequencer according to the standard protocol.

TSS-seq

Two hundred micrograms of the total RNA obtained was subjected to oligo-capping, with some modifications from the original protocol. Briefly, after successive treatment of the RNA with 2.5 U bacterial alkaline phosphatase (BAP) (TaKaRa) at 37°C for 1 h and 40 U tabacco acid pyrophosphatase (TAP) (Ambion) at 37°C for 1 h, the BAP–TAP-treated RNAs were ligated to 1.2 μg of the RNA oligonucleotide 5′-AAUGAUACGGCGACCACCG AGAUCUACACUCUUUCCCUACACGACGCUCUUC CGAUCUGG-3′ using 250 U T4 RNA ligase (TaKaRa) at 20°C for 3 h. After DNase I treatment (TaKaRa), polyA-containing RNA was selected using oligo-dT powder (Collaborative). The first strand cDNA was synthesized with 10 pmol random hexamer primer (5′-CA AGCAGAAGACGGCATACGANNNNNNC-3′) using SuperScript II (Invitrogen), with incubation at 12°C for 1 h and 42°C overnight. The template RNA was degraded by alkaline treatment. For PCR, 20% of the first strand cDNA was used as the PCR template. Gene Amp PCR kits (PerkinElmer) were used with the PCR primers 5′-AATGA TACGGCGACCACCGAG-3′ and 5′-CAAGCAGAAGA CGGCATACGA-3′ under the following reaction conditions: 15 cycles at 94°C for 1 min, 56°C for 1 min and 72°C for 2 min. The PCR fragments were size fractionated by 12% PAGE and the fraction that contained the 150–250 bp fragments was recovered. The quality and quantity of the obtained single-stranded first strand cDNAs were assessed using a BioAnalyzer (Agilent). One nanogram of the size-fractionated cDNA was used for sequencing reactions with the Illumina GA. The sequencing reactions were performed according to the manufacturer's instructions.

ChIP and ChIP-seq

BmN4 cells (2×10^6) were cultured in 100 mm treated cell culture dish (Corning) for 24 h at 27°C. Cells were fixed with 10% formalin/phosphate-buffered saline (PBS buffer; 137 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4 and 1.4 mM KH_2PO_4) for 15 min at room temperature. The reaction was quenched by the addition of glycine (final concentration: 125 mM). Cells were further washed twice with ice-cold PBS and collected by 1500 rpm centrifugation for 5 min at 4°C. Cells were lysed with ChIP-lysis buffer (1% sodium dodecyl sulfate (SDS), 10 mM ethylenediaminetetraacetic acid (EDTA), 50 mM Tris–HCl (pH 8.1) and protease inhibitor cocktail

(Roche)). After incubation on ice, homogenates were sonicated with Sonifier 250 (BRANSON) 10 s followed by incubation on ice 10 s (duty cycle = 50 and out put control = 7). This process was repeated 24 times to generate DNA of 200–500 bp. The resulting homogenates were cleared by 13 000 rpm centrifugation for 10 min at 4°C. Ten micrograms per microliters of cleared lysate was 10-fold diluted by ChIP dilution buffer (1% Triton X-100, 2 mM EDTA, 150 mM NaCl, 20 mM Tris–HCl (pH 8.1) and protease inhibitor cocktail). Lysate was then pre-incubated with protein A-Sepharose (GE Healthcare) and then cleared by 1500 rpm centrifugation 30 s at room temperature. Pre-cleared lysate (Input) was incubated with anti-pol II (Santa Cruz), anti-Mouse IgG (IgG-M; Santa Cruz), anti-H3K9me2 (Millipore), anti-H3K9me3 (Millipore or Abcam), anti-H3K9ac (Millipore), anti-H3K4me2 (Millipore), anti-H3K4me3 (Millipore or Abcam) and anti-rabbit IgG (IgG-R; Millipore) (1:1000) at 4°C for overnight. Lysate was then incubated with equilibrated Protein-A-Sepharose beads for 1 h at 4°C. Beads were washed sequentially with low salt buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris–HCl (pH 8.1) and 150 mM NaCl) (once), high salt buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris–HCl (pH 8.1) and 500 mM NaCl) (once), LiCl buffer (250 mM LiCl, 1% NP-40, 1% deoxycholate, 1 mM EDTA and 10 mM Tris–HCl (pH 8.1)) (once) and Tris-EDTA (TE) buffer (10 mM Tris–HCl and 1 mM EDTA) (twice). Immunoprecipitates were eluted with 50 µl of elution buffer (1% SDS and 0.1 M $NaHCO_3$) for 15 min at room temperature twice. Total 100 µl of elutant was further incubated with 4 µl of 5 M NaCl for 6 h at 65°C, 4 µl of 10 mg/ml RNase A for 2 h at 37°C and 2 µl of 20 mg/ml proteinase K for 2 h at 55°C. DNA was collected with phenol–chloroform followed by EtOH precipitation and dissolved with 150 µl of sterile nuclease-free water. Using the recovered DNA, the ChIP-seq samples were prepared for the Illumina GA according to the manufacturer's instructions. ChIP-seq experiments were performed with Millipore and Santa Cruz antibodies, and Abcam antibodies against H3K4me3 and H3K9me3 were used for validation of ChIP-seq data (Supplementary Figure S4).

Data analyses

We trimmed 3′-adaptor sequence from raw qseq data by in-house R program allowing two mismatches. Prior to mapping to the silkworm genome assembly build2, we sequenced a part of unassembled region in *Torimochi* to generate 'modified' silkworm genome assembly build2. Obtained sequences were mapped to the modified silkworm genome assembly build2 and silkworm predicted protein-coding genes by using Bowtie or Soap2 (33,34). Reads that were mapped to the genome perfectly and uniquely were further considered. We analyzed histone modifications and Pol II enrichment by using well-established softwares MACS (35), SICER (36) and BEDtools (37). MACS was used for H3K4me2, H3K4me3, H3K9ac and Poll, which show 'sharp peaks'. For H3K9me2 and H3K9me3, which show 'broader peaks', SICER was applied. In both cases, *P*-value cutoff (1.00e−05) was used to identify peaks. In the original pipeline to define piRNA clusters and the transcribed regions, each reads were normalized to total genome-mapping reads for comparison. Overlapping reads were merged by using BEDtools (37). Reads in each domain were normalized by total genome-mapping reads and defined as reads per million (RPM). piRNAs used in this study are described in (18). We focused only on unique mappers, corresponding to ∼50% of the total genome-mapping reads. When a piRNA overlapped with another piRNA, these were recognized as a single domain. piRNA domains located within 300-bp with each other were further considered as piRNA clusters. At the same time, we calculated the number of reads embedded within the clusters to infer an expression level of each piRNA cluster. For TSS analysis, overlaps occurred within ±500-bp from the TSSs were considered as the overlaps. For the other overlapping analyses, the overlaps >1 bp between two different peaks were recognized as actual overlaps. The statistical significances of overlaps between two peaks were evaluated using a Poisson distribution as described previously (38):

$$p(\chi, \lambda) = 1 - \sum_{t=0}^{x} \frac{e^{-\lambda}\lambda^{t}}{t!},$$

where $p(\chi, \lambda)$ is the probability of observed overlap, λ is the expected overlap between two populations and χ is the observed overlap. For constructing density plots, ChIP-seq and RNA-seq tags were mapped by using ELAND pipeline and average base concentrations were visualized using Genome studio according to the manufacturer's instruction (38).

ChIP-PCR

Quantitative ChIP-PCR analyses were performed with KAPA SYBR FAST qPCR kit Master Mix ABI PRISM (KAPA BIOSYSTEMS) and specific primers. Primers used in this study were described in Supplementary Table S1.

Rapid Amplification of cDNA Ends (RACE)

RACE experiments were performed by using GeneRacer Kit (Invitrogen) as described previously (18).

Data deposition

Deep sequencing data obtained in this study are available under the accession number of DRA000527 (DDBJ).

RESULTS

Expression of piRNA pathway genes in BmN4 cell line

To define transcriptionally active or inactive domains in the BmN4 genome, and to know which genes are actively expressed in BmN4 cell line, we performed RNA-seq analysis. Our data showed that, besides silkworm *PIWI* genes *Siwi* and *BmAgo3* (39), orthologs of piRNA pathway genes—*Vasa* (40), *Zucchini*/*MitoPLD* (26,41–45), *Armitage*/*Mov10* (41,42,46–49), *Maelstrom* (50–52), *Tudor* family genes (*Tudor-domain containing 1*

(*Tdrd1*), *Spindle-E*/*Tdrd9*, *Vreteno*, *Yb* and *Kumo*/*Qin*) (46,47,53–60) and *Hen1* (11,13–15,61)—were expressed in BmN4 cells (Supplementary Figure S1A), supporting that the BmN4 cell line harbors the germ-line piRNA pathway (32).

The chromatin landscape of the piRNA-generating cell line

To define the chromatin landscape of BmN4 cells, we thoroughly analyzed five major histone modifications (Table 1 and Supplementary Table S2; see 'Materials and Methods' section) that are known markers for either euchromatin or heterochromatin. To infer the transcriptional state associated with a histone modification, we comparatively analyzed our RNA-seq and ChIP-seq data through calculating the number of histone modification peaks overlapping with RNA-seq tags (Figure 1A). H3K4me3 exhibited extensive overlap with RNA-seq tags (300-fold enrichment to the IgG-R peaks, $P < 0.01$ (see 'Materials and Methods' section)). The score for H3K4me2 and H3K9ac was relatively high (8.5-fold and 4-fold enrichment, respectively, $P < 0.01$). H3K9me2 peaks overlapped with RNA-seq tags to lesser extent (2.2-fold enrichment, $P = 0.08$). In contrast, H3K9me3 was barely overlapped with RNA-seq tags. Our data clearly indicate that actively transcribed loci positively correlated with H3K4me2, H3K4me3 and H3K9ac whereas H3K9me3 peaks were especially transcriptionally inactive, as deduced by RNA-seq and ChIP-seq, respectively.

To identify modifications that may function together, we compared the genomic distributions of the five histone modifications and calculated the overlap among them. We found that H3K4me2 frequently co-occurred with H3K4me3 and H3K9ac (Figure 1B), suggesting that these act in concert to define euchromatin. H3K4me3 peaks relatively more frequently overlapped with H3K4me2, but this was not statistically significant. H3K9ac peaks often shared same genomic regions with H3K4me2 and H3K4me3 peaks. H3K9me3 showed overlap most frequently with H3K9me2 and vice versa, but relatively hardly localized with H3K4me2, H3K4me3 and H3K9ac, supporting that H3K9me3 defines heterochromatin.

Next, we investigated enrichment of histone modifications and Pol II in the TSSs. As shown in Figure 1C, H3K4me3 and Pol II showed high and statistically significant enrichments in the TSSs (1100-fold and 50-fold enrichment, respectively, $P < 0.01$). Weaker enrichments were observed for H3K4me2 and H3K9ac (Figure 1C and Supplementary Figure S1B). H3K9me2 and H3K9me3 were depleted in the TSSs (Figure 1C). House-keeping gene *actin3* locus was a representative locus showing such features (Supplementary Figure S1C). Collectively, our data suggest that H3K4me3 is the most prominent euchromatin mark in BmN4 cell line. H3K4me3 with H3K4me2 and H3K9ac specify euchromatin domains whereas H3K9me2 and H3K9me3 represent heterochromatin domains in this cell line.

Chromatin modifications of piRNA clusters

Altogether our findings provide insights into the chromatin landscape of BmN4 cells. BmN4 cells are unique in that they comprise the germ-line piRNA pathway. To investigate the genomic character of piRNA-generating loci, we next examined the epigenomic features of piRNA clusters by investigating how piRNA clusters are defined by histone modification peaks. In the analysis, we focused on 965 unique piRNA clusters producing >100 RPM unique piRNAs. We found that unique piRNA clusters were overlapped most frequently with H3K4me3 peaks (15-fold enrichment, corresponding to 17.6% of 965 piRNA clusters, $P < 0.01$) followed by H3K4me2 (4-fold enrichment, corresponding to 3.8% of 965 piRNA clusters, $P < 0.01$), both representing euchromatin marks (Figure 2A). Unique piRNA clusters were hardly marked with H3K9me2 and H3K9me3. In addition, >70% of these piRNA clusters were positive for polyA-selected RNA-seq tags (Figure 2A). As we found that piRNA clusters were markedly enriched for histone modifications that mark euchromatin and concomitantly co-localized with RNA-seq tags, we suggest that at least a part of unique piRNA clusters retain the euchromatin features.

To further investigate the chromatin landscape of piRNA clusters, we focused on a unique piRNA cluster named *Torimochi*. *Torimochi* is a functional piRNA cluster in that it is capable of capturing insertions and generating de novo piRNAs (18). Visualization of ChIP-seq mapping data demonstrated that *Torimochi* is

Table 1. Characterization of peaks of five histone modifications, Pol II, RNA-seq and piRNA clusters

Name	Method	*P*-value cutoff	Number of peaks/ domains/clusters	Number of peaks not overlapping with IgG peaks	Note
IgG-R-M	MACS	1.00E−05	391	–	
H3K4me2	MACS	1.00E−05	1152	921	
H3K4me3	MACS	1.00E−05	6705	6593	
H3K9ac	MACS	1.00E−05	770	567	
IgG-R-S	SICER	1.00E−05	2122	–	
H3K9me2	SICER	1.00E−05	4735	3550	
H3K9me3	SICER	1.00E−05	1337	653	
IgG-M	MACS	1.00E−05	317	–	
Pol II	MACS	1.00E−05	3984	3242	
RNA-seq	BEDtools	–	84 148	–	>100-bp
piRNAs	BEDtools	–	965	–	>100 RPM

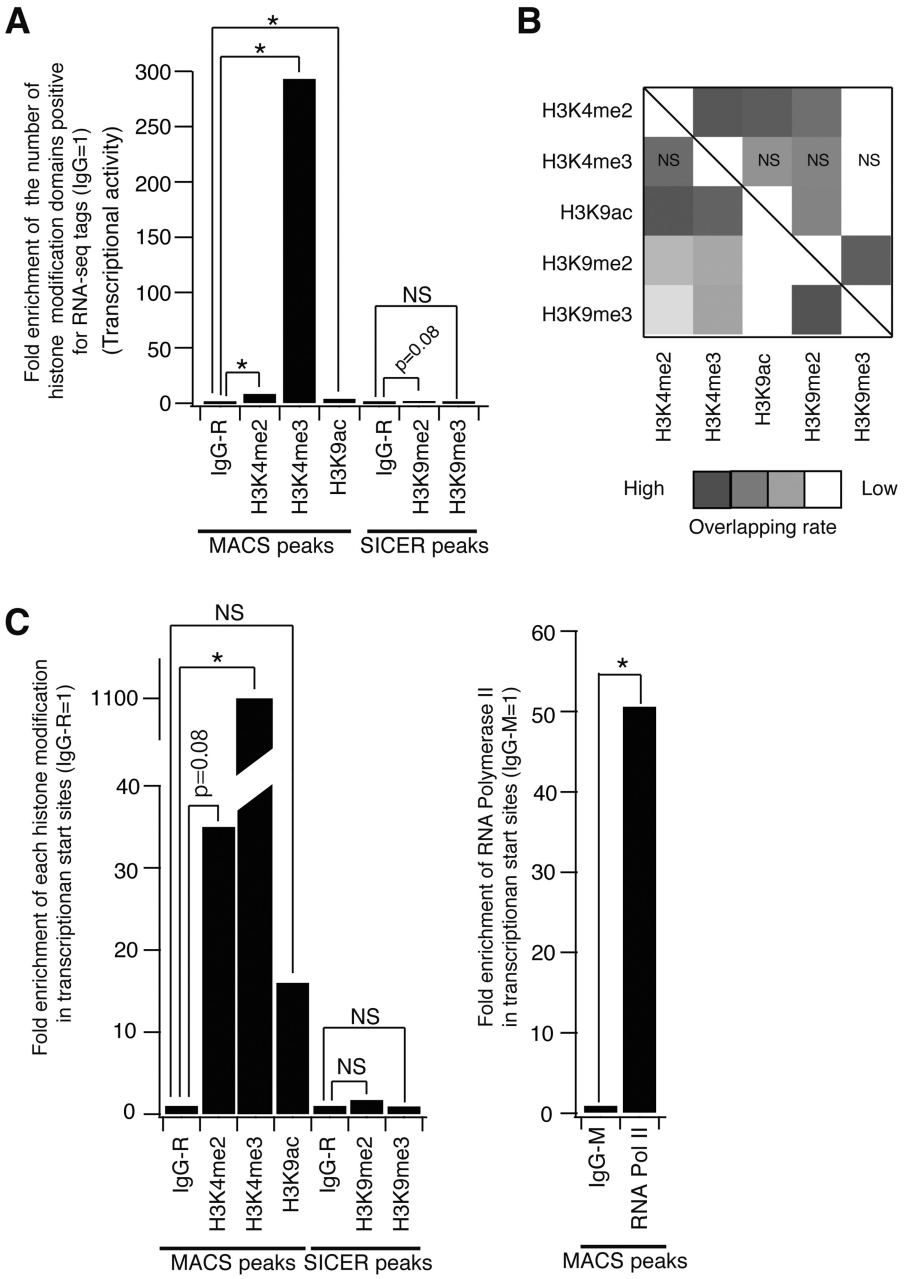

Figure 1. Histone code of the BmN4 cell genome. (**A**) Transcriptional activity in five histone modifications. Fold enrichments of the number of histone modification peaks overlapping with RNA-seq tags are shown. Every score was normalized to that in IgG-R (rabbit IgG) peaks. *P*-values were calculated using a Poisson distribution. $*P < 0.01$; NS, non-significant. (**B**) Heat map showing the overlaps among five histone modifications. Peaks not overlapping with IgG-R were considered. The statistical significance of each combination was evaluated using a Poisson distribution. NS, non-significant. Without NS, $P < 0.01$. (**C**) The chromatin landscape for whole-genome TSSs in the BmN4 genome. Every score was normalized to that in IgG-R peaks. Fold enrichment of RNA pol II peaks is also shown, where IgG-M (mouse IgG) peaks served as a control library. *P*-values were calculated using a Poisson distribution. $*P < 0.01$; NS, non-significant.

highly euchromatic—*Torimochi* was enriched for H3K4me2, H3K4me3 and H3K9ac while was depleted for H3K9me2 and H3K9me3 (Figure 2B). Quantitative ChIP-PCR analysis confirmed these data (Figure 2C). In addition, we confirmed these observations by using different, well-validated antibodies (Supplementary Figure S4). We also analyzed two silkworm heterochromatin proteins—BmHP1a and BmHP1b (Figure 2C and Supplementary Figures S2 and S3) (62). BmHP1a was enriched for heterochromatin region (*SART1* transposon) where H3K9me2 and H3K9me3 were enriched, but not for euchromatic region (Figure 2C and Supplementary Figure S2). BmHP1b did not show any enrichment for *SART1* region. In contrast, we found that the *Torimochi* locus entirely lacked these heterochromatin proteins (Figure 2C). In addition, we performed the same experiments against another piRNA cluster on Chr07 and obtained essentially the same results (Supplementary

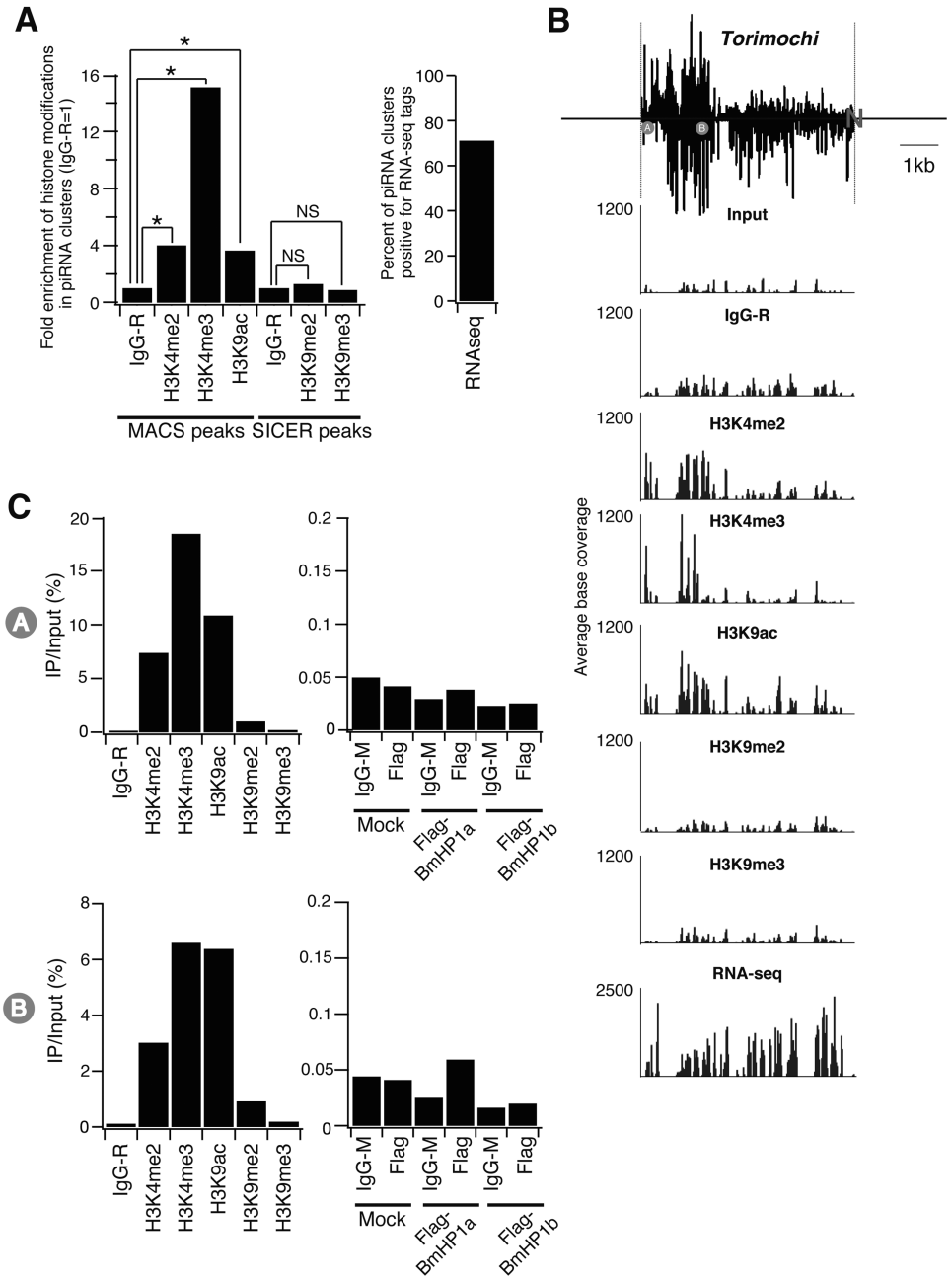

Figure 2. The chromatin landscape of piRNA clusters. (**A**) The chromatin landscape of piRNA clusters producing >100 RPM piRNAs. Every score was normalized to those of IgG-R (rabbit IgG) peaks. The percent of piRNA clusters positive for RNA-seq tags is also shown. *P*-values were calculated using a Poisson distribution. $*P < 0.01$; NS, non-significant. (**B**) The chromatin landscape of *Torimochi*, a representative piRNA cluster in the BmN4 genome. The density of piRNAs in the *Torimochi* locus was visualized by using the 'modified' silkworm genome where we sequenced a part of unassembled region in the *Torimochi* locus (see 'Materials and Methods' section). The remaining unassembled region is shown as N. Mapping patterns of ChIP-seq and RNA-seq tags were visualized by using Genome studio (Illumina). The primer annealing sites for Figure 2C are indicated. (**C**) ChIP-PCR analyses for the *Torimochi* locus. In addition to five histone modifications, DNA samples immunoprecipitated with the two silkworm heterochromatin proteins were analyzed.

Figure S3). Thus, as far as we investigated, unique piRNA clusters have a euchromatic nature in the BmN4 genome, and we did not detect any unique piRNA clusters with heterochromatin features.

Characterization of piRNA precursor transcription

The nature of piRNA precursor transcription and piRNA precursors has been poorly understood. To address this, we tried to identify TSSs embedded within piRNA clusters (piRNA cluster-associated TSSs). For this analysis, 965 piRNA clusters producing >100 RPM unique piRNAs were considered. As a result, we identified 407 piRNA cluster-associated TSSs generating 5′-capped RNAs, providing the first comprehensive evidence that many piRNA precursors are 5′-capped RNAs. This was confirmed by 5′-RACE experiments for one of the reliable

piRNA precursors *Torimochi* (Figure 3A). Moreover, 3′-RACE experiments followed by sequencing revealed that a *Torimochi* piRNA precursor was 3′-polyA-tailed (Figure 3A).

Finally, to know how piRNA precursors are transcribed, we determined the chromatin state around piRNA cluster-associated TSSs. We found that piRNA cluster-associated TSSs were enriched for euchromatin marks such as H3K4me3, and notably, RNA pol II (Figure 3B). *Torimochi*-TSSs were a representative pol II-enriched TSS (Figure 3C). An enrichment of pol II signal in the piRNA cluster-associated TSSs strongly indicated that at least a part of piRNA precursors is transcribed by RNA pol II.

DISCUSSION

Here, we comprehensively explored the chromatin landscape of BmN4 cells, which is characterized by expression of germ-line piRNAs, piRNA pathway genes and germ-cell marker genes (Supplementary Figure S1A) (32).

In BmN4 cells, euchromatin is specified by histone modifications H3K4me2, H3K4me3 and H3K9ac, whereas H3K9me2 and H3K9me3 define heterochromatin (Figure 1). Heterochromatin also contains silkworm heterochromatin protein BmHP1a (Supplementary Figure S2A and B). In addition to H3K4me2 and H3K4me3, pol II was enriched in TSS regions (Figure 1C and Supplementary Figure S1B). These features agree well with similar observations in other organisms, such as

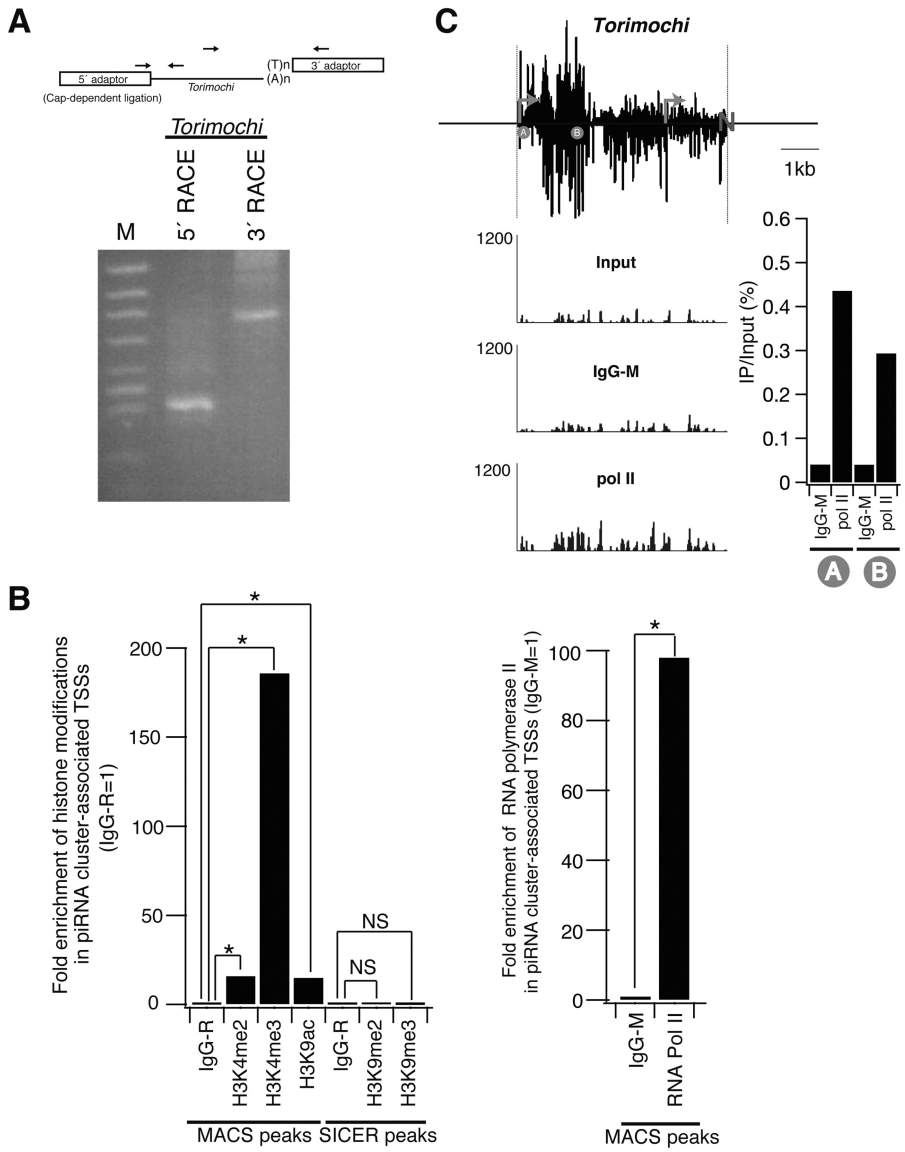

Figure 3. The nature of piRNA precursor transcription. (**A**) 5′- and 3′-RACE for *Torimochi* piRNA precursor. The amplicons were cloned and sequenced to validate that 5′-end of the precursor was G-capped and that 3′-end of the precursor was poly-A-tailed. (**B**) The chromatin landscape of piRNA-cluster-associated TSSs. IgG-R peaks (rabbit IgG) served as a control for five histone modifications and IgG-M peaks (mouse IgG) served as a control for RNA pol II. *P*-values were calculated using a Poisson distribution. $*P < 0.01$; NS, non-significant. (**C**) Enrichment of RNA pol II in the *Torimochi*-associated TSSs. The arrows indicate TSSs in the *Torimochi* locus. Mapping patterns of ChIP-seq tags were visualized by using Genome studio (Illumina).

flies, worms and humans (38,63–69). Collectively, our study demonstrates a conserved histone modification rule in BmN4 cells.

We applied our gained insights to better understand piRNA-generating loci, and in turn piRNA biogenesis. Our analyses demonstrated that unique piRNA clusters are exclusively modified with euchromatin marks. piRNA cluster-associated TSSs were enriched for euchromatic marks and pol II (Figure 3). With the aid of TSS-seq, polyA-RNA-seq, and RACE analyses, we concluded that at least a part of piRNA precursors is transcribed by pol II and originate from 5′-capped and 3′-polyA-tailed RNAs. This suggests that the nature of piRNA precursors resembles that of normal protein-coding genes or miRNA genes (70). This is in line with that piRNA maturation is thought to occur in the cytoplasm. Taken together, a number of unique piRNA clusters and associated TSSs are highly euchromatic and allow for pol II-based transcription of piRNA precursors that bear canonical mRNA features.

Torimochi was a representative euchromatic piRNA cluster (Figure 2B and C). We previously showed that integration of GFP transgene into *Torimochi* generated de novo GFP piRNAs, establishing piRNA-based silencing against GFP sequence (18). In that study, we utilized *piggyBac*-based transposition system to integrate a GFP transgene into the BmN4 genome. Curiously, among eight lines where transposition successfully occurred, three lines showed *Torimochi*-based GFP silencing. Our findings indicated that the euchromatic (opened) nature of *Torimochi* allows this locus to capture new insertions efficiently, making *Torimochi* a hot spot for transposition.

Currently known fly piRNA clusters exhibit heterochromatic features and are present in constitutive heterochromatin (10,71). Two fly heterochromatin proteins HP1a and HP1d/Rhino have been implicated in the piRNA pathway, and deposition of H3K9me3, which is catalyzed by dSETDB1 protein, is required for piRNA precursor transcription (24,27,28). Silkworm telomeres are constitutive heterochromatin, which are composed of large arrays of telomere-specific transposons and thus are absent in the silkworm draft genome sequence (72). Abundant piRNAs are expressed from silkworm telomere-specific transposons such as *SART1* (32,72,73), suggesting that silkworm telomeres are large piRNA clusters. Even though we cannot correctly address chromatin states for silkworm telomeres due to their repetitive nature, we nonetheless performed ChIP-PCR analyses against silkworm telomeres (Supplementary Figure S2). We found that telomeric piRNA clusters were marked with heterochromatin marks H3K9me2, H3K9me3 and heterochromatin protein BmHP1a. Thus, piRNA clusters in constitutive heterochromatin showed heterochromatin features in the silkworm as is the case for fly heterochromatic piRNA clusters. Interestingly, telomeric piRNA clusters associated with euchromatin marks and pol II, too, implying a multifaceted chromatin landscape in telomeric piRNA clusters (Supplementary Figure S2). *SART1* piRNA precursors were 5′-capped and 3′-polyA-tailed (data not shown). These results indicated that piRNA clusters are divided into at least two groups: unique piRNA clusters that are exclusively marked with euchromatin marks and piRNA clusters with both euchromatic and heterochromatic features in constitutive heterochromatin. We envision that biased localization of fly piRNA clusters in constitutive heterochromatin may explain why euchromatic piRNA clusters appear to be rare in the fly genome.

ACCESSION NUMBERS

Deep sequencing data obtained in this study are available under the accession number of DRA000527 (DDBJ).

SUPPLEMENTARY DATA

Supplementary data are available at NAR Online: Supplementary Tables 1 and 2 and Supplementary Figures 1–4.

ACKNOWLEDGEMENTS

The authors thank T. Horiuchi and M. Kawamoto for their technical assistance. They also thank P.B. Kwak, J.G. Betancur and the members of our laboratories for the critical reading on the article.

FUNDING

The Program for Promotion of Basic and Applied Researches for Innovations in Bio-oriented Industry (to Su.K.); Special Coordination Funds for Promoting Science and Technology from the Ministry of Education, Culture, Sports, Science and Technology of the Japanese Government (MEXT) [22115502 to Su.K. and 17018007 to T.S., in part and the Professional Program for Agricultural Bioinformatics]; National Bio-Resource Project 'Silkworm' of MEXT and Grant-in-Aid for Scientific Research on Innovative Areas (Functional machinery for non-coding RNAs) (to Su.K. and Y.T.). Recipient of a fellowship from the Japan Society for the Promotion of Science (to Sh.K.). Funding for open access charge: The Program for Promotion of Basic and Applied Researches for Innovations in Bio-oriented Industry (to Su.K.).

Conflict of interest statement. None declared.

REFERENCES

1. Malone,C.D. and Hannon,G.J. (2009) Molecular evolution of piRNA and transposon control pathways in Drosophila. *Cold Spring Harb. Symp. Quant. Biol.*, **74**, 225–234.
2. Slotkin,R.K. and Martienssen,R. (2007) Transposable elements and the epigenetic regulation of the genome. *Nat. Rev. Genet.*, **8**, 272–285.
3. Ghildiyal,M. and Zamore,P.D. (2009) Small silencing RNAs: an expanding universe. *Nat. Rev. Genet.*, **10**, 94–108.
4. Klattenhoff,C. and Theurkauf,W. (2008) Biogenesis and germline functions of piRNAs. *Development*, **135**, 3–9.
5. Ronsseray,S., Josse,T., Boivin,A. and Anxolabehere,D. (2003) Telomeric transgenes and trans-silencing in Drosophila. *Genetica*, **117**, 327–335.
6. Ronsseray,S., Lehmann,M. and Anxolabehere,D. (1991) The maternally inherited regulation of P elements in Drosophila

melanogaster can be elicited by two P copies at cytological site 1A on the X chromosome. *Genetics*, **129**, 501–512.
7. Ronsseray,S., Lehmann,M., Nouaud,D. and Anxolabehere,D. (1996) The regulatory properties of autonomous subtelomeric P elements are sensitive to a Suppressor of variegation in Drosophila melanogaster. *Genetics*, **143**, 1663–1674.
8. Ronsseray,S., Marin,L., Lehmann,M. and Anxolabehere,D. (1998) Repression of hybrid dysgenesis in Drosophila melanogaster by combinations of telomeric P-element reporters and naturally occurring P elements. *Genetics*, **149**, 1857–1866.
9. Desset,S., Meignin,C., Dastugue,B. and Vaury,C. (2003) COM, a heterochromatic locus governing the control of independent endogenous retroviruses from Drosophila melanogaster. *Genetics*, **164**, 501–509.
10. Brennecke,J., Aravin,A.A., Stark,A., Dus,M., Kellis,M., Sachidanandam,R. and Hannon,G.J. (2007) Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. *Cell*, **128**, 1089–1103.
11. Horwich,M.D., Li,C., Matranga,C., Vagin,V., Farley,G., Wang,P. and Zamore,P.D. (2007) The Drosophila RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC. *Curr. Biol.*, **17**, 1265–1272.
12. Kawaoka,S., Izumi,N., Katsuma,S. and Tomari,Y. (2011) 3′-end formation of PIWI-interacting RNAs in vitro. *Mol. Cell*, **43**, 1015–1022.
13. Kirino,Y. and Mourelatos,Z. (2007) Mouse Piwi-interacting RNAs are 2′-O-methylated at their 3′ termini. *Nat. Struct. Mol. Biol.*, **14**, 347–348.
14. Ohara,T., Sakaguchi,Y., Suzuki,T., Ueda,H., Miyauchi,K. and Suzuki,T. (2007) The 3′ termini of mouse Piwi-interacting RNAs are 2′-O-methylated. *Nat. Struct. Mol. Biol.*, **14**, 349–350.
15. Saito,K., Sakaguchi,Y., Suzuki,T., Suzuki,T., Siomi,H. and Siomi,M.C. (2007) Pimet, the Drosophila homolog of HEN1, mediates 2′-O-methylation of Piwi-interacting RNAs at their 3′ ends. *Genes Dev.*, **21**, 1603–1608.
16. Houwing,S., Kamminga,L.M., Berezikov,E., Cronembold,D., Girard,A., van den Elst,H., Filippov,D.V., Blaser,H., Raz,E., Moens,C.B. *et al.* (2007) A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish. *Cell*, **129**, 69–82.
17. Vagin,V.V., Sigova,A., Li,C., Seitz,H., Gvozdev,V. and Zamore,P.D. (2006) A distinct small RNA pathway silences selfish genetic elements in the germline. *Science*, **313**, 320–324.
18. Kawaoka,S., Mitsutake,H., Kiuchi,T., Kobayashi,M., Yoshikawa,M., Suzuki,Y., Sugano,S., Shimada,T., Kobayashi,J., Tomari,Y. *et al.* (2012) A role for transcription from a piRNA cluster in de novo piRNA production. *RNA*, **18**, 265–273.
19. Gunawardane,L.S., Saito,K., Nishida,K.M., Miyoshi,K., Kawamura,Y., Nagami,T., Siomi,H. and Siomi,M.C. (2007) A slicer-mediated mechanism for repeat-associated siRNA 5′ end formation in Drosophila. *Science*, **315**, 1587–1590.
20. Nishida,K.M., Saito,K., Mori,T., Kawamura,Y., Nagami-Okada,T., Inagaki,S., Siomi,H. and Siomi,M.C. (2007) Gene silencing mechanisms mediated by Aubergine piRNA complexes in Drosophila male gonad. *RNA*, **13**, 1911–1922.
21. Saito,K., Nishida,K.M., Mori,T., Kawamura,Y., Miyoshi,K., Nagami,T., Siomi,H. and Siomi,M.C. (2006) Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the Drosophila genome. *Genes Dev.*, **20**, 2214–2222.
22. De Fazio,S., Bartonicek,N., Di Giacomo,M., Abreu-Goodger,C., Sankar,A., Funaya,C., Antony,C., Moreira,P.N., Enright,A.J. and O′Carroll,D. (2011) The endonuclease activity of Mili fuels piRNA amplification that silences LINE1 elements. *Nature*, **480**, 259–263.
23. Reuter,M., Berninger,P., Chuma,S., Shah,H., Hosokawa,M., Funaya,C., Antony,C., Sachidanandam,R. and Pillai,R.S. (2011) Miwi catalysis is required for piRNA amplification-independent LINE1 transposon silencing. *Nature*, **480**, 264–267.
24. Brower-Toland,B., Findley,S.D., Jiang,L., Liu,L., Yin,H., Dus,M., Zhou,P., Elgin,S.C. and Lin,H. (2007) Drosophila PIWI associates with chromatin and interacts directly with HP1a. *Genes Dev.*, **21**, 2300–2311.
25. Lau,N.C., Robine,N., Martin,R., Chung,W.J., Niki,Y., Berezikov,E. and Lai,E.C. (2009) Abundant primary piRNAs, endo-siRNAs, and microRNAs in a Drosophila ovary cell line. *Genome Res.*, **19**, 1776–1785.
26. Todeschini,A.L., Teysset,L., Delmarre,V. and Ronsseray,S. (2010) The epigenetic trans-silencing effect in Drosophila involves maternally-transmitted small RNAs whose production depends on the piRNA pathway and HP1. *PLoS One*, **5**, e11032.
27. Klattenhoff,C., Xi,H., Li,C., Lee,S., Xu,J., Khurana,J.S., Zhang,F., Schultz,N., Koppetsch,B.S., Nowosielska,A. *et al.* (2009) The Drosophila HP1 homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters. *Cell*, **138**, 1137–1149.
28. Rangan,P., Malone,C.D., Navarro,C., Newbold,S.P., Hayes,P.S., Sachidanandam,R., Hannon,G.J. and Lehmann,R. (2011) piRNA production requires heterochromatin formation in Drosophila. *Curr. Biol.*, **21**, 1373–1379.
29. Klenov,M.S., Lavrov,S.A., Stolyarenko,A.D., Ryazansky,S.S., Aravin,A.A., Tuschl,T. and Gvozdev,V.A. (2007) Repeat-associated siRNAs cause chromatin silencing of retrotransposons in the Drosophila melanogaster germline. *Nucleic Acids Res.*, **35**, 5430–5438.
30. Yin,H. and Lin,H. (2007) An epigenetic activation role of Piwi and a Piwi-associated piRNA in Drosophila melanogaster. *Nature*, **450**, 304–308.
31. Zhang,Z. and Pugh,B.F. (2011) High-resolution genome-wide mapping of the primary structure of chromatin. *Cell*, **144**, 175–186.
32. Kawaoka,S., Hayashi,N., Suzuki,Y., Abe,H., Sugano,S., Tomari,Y., Shimada,T. and Katsuma,S. (2009) The Bombyx ovary-derived cell line endogenously expresses PIWI/PIWI-interacting RNA complexes. *RNA*, **15**, 1258–1264.
33. Li,R., Yu,C., Li,Y., Lam,T.W., Yiu,S.M., Kristiansen,K. and Wang,J. (2009) SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics*, **25**, 1966–1967.
34. Langmead,B., Trapnell,C., Pop,M. and Salzberg,S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.*, **10**, R25.
35. Zhang,Y., Liu,T., Meyer,C.A., Eeckhoute,J., Johnson,D.S., Bernstein,B.E., Nusbaum,C., Myers,R.M., Brown,M., Li,W. *et al.* (2008) Model-based analysis of ChIP-seq (MACS). *Genome Biol.*, **9**, R137.
36. Zhang,C., Schones,D.E., Zeng,C., Cui,K., Zhao,K. and Peng,W. (2009) A clustering approach for identification of enriched domains from histone modification ChIP-seq data. *Bioinformatics*, **25**, 1952–1958.
37. Quinlan,A.R. and Hall,I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**, 841–842.
38. Yamashita,T., Sathira,N.P., Kanai,A., Tanimoto,K., Arauchi,T., Tanaka,Y., Hashimoto,S., Sugano,S., Nakai,K. and Suzuki,Y. (2011) Genome-wide characterization of transcription start sites in humans by integrative transcriptome analysis. *Genome Res.*, **21**, 775–789.
39. Kawaoka,S., Minami,K., Katsuma,S., Mita,K. and Shimada,T. (2008) Developmentally synchronized expression of two Bombyx mori Piwi subfamily genes, SIWI and BmAGO3 in germ-line cells. *Biochem. Biophys. Res. Commun.*, **367**, 755–760.
40. Kuramochi-Miyagawa,S., Watanabe,T., Gotoh,K., Takamatsu,K., Chuma,S., Kojima-Kita,K., Shiromoto,Y., Asada,N., Toyoda,A., Fujiyama,A. *et al.* (2010) MVH in piRNA processing and gene silencing of retrotransposons. *Genes Dev.*, **24**, 887–892.
41. Haase,A.D., Fenoglio,S., Muerdter,F., Guzzardo,P.M., Czech,B., Pappin,D.J., Chen,C., Gordon,A. and Hannon,G.J. (2010) Probing the initiation and effector phases of the somatic piRNA pathway in Drosophila. *Genes Dev.*, **24**, 2499–2504.
42. Olivieri,D., Sykora,M.M., Sachidanandam,R., Mechtler,K. and Brennecke,J. (2010) An in vivo RNAi assay identifies major genetic and cellular requirements for primary piRNA biogenesis in Drosophila. *EMBO J.*, **29**, 3301–3317.
43. Pane,A., Wehr,K. and Schupbach,T. (2007) zucchini and squash encode two putative nucleases required for rasiRNA production in the Drosophila germline. *Dev. Cell*, **12**, 851–862.

44. Saito,K., Inagaki,S., Mituyama,T., Kawamura,Y., Ono,Y., Sakota,E., Kotani,H., Asai,K., Siomi,H. and Siomi,M.C. (2009) A regulatory circuit for piwi by the large Maf gene traffic jam in Drosophila. *Nature*, **461**, 1296–1299.
45. Watanabe,T., Chuma,S., Yamamoto,Y., Kuramochi-Miyagawa,S., Totoki,Y., Toyoda,A., Hoki,Y., Fujiyama,A., Shibata,T., Sado,T. *et al.* (2011) MITOPLD is a mitochondrial protein essential for nuage formation and piRNA biogenesis in the mouse germline. *Dev. Cell*, **20**, 364–375.
46. Qi,H., Watanabe,T., Ku,H.Y., Liu,N., Zhong,M. and Lin,H. (2011) The Yb body, a major site for Piwi-associated RNA biogenesis and a gateway for Piwi expression and transport to the nucleus in somatic cells. *J. Biol. Chem.*, **286**, 3789–3797.
47. Saito,K., Ishizu,H., Komai,M., Kotani,H., Kawamura,Y., Nishida,K.M., Siomi,H. and Siomi,M.C. (2010) Roles for the Yb body components Armitage and Yb in primary piRNA biogenesis in Drosophila. *Genes Dev.*, **24**, 2493–2498.
48. Frost,R.J., Hamra,F.K., Richardson,J.A., Qi,X., Bassel-Duby,R. and Olson,E.N. (2010) MOV10L1 is necessary for protection of spermatocytes against retrotransposons by Piwi-interacting RNAs. *Proc. Natl Acad. Sci. USA*, **107**, 11847–11852.
49. Zheng,K., Xiol,J., Reuter,M., Eckardt,S., Leu,N.A., McLaughlin,K.J., Stark,A., Sachidanandam,R., Pillai,R.S. and Wang,P.J. (2010) Mouse MOV10L1 associates with Piwi proteins and is an essential component of the Piwi-interacting RNA (piRNA) pathway. *Proc. Natl Acad. Sci. USA*, **107**, 11841–11846.
50. Aravin,A.A., van der Heijden,G.W., Castaneda,J., Vagin,V.V., Hannon,G.J. and Bortvin,A. (2009) Cytoplasmic compartmentalization of the fetal piRNA pathway in mice. *PLoS Genet.*, **5**, e1000764.
51. Lim,A.K. and Kai,T. (2007) Unique germ-line organelle, nuage, functions to repress selfish genetic elements in Drosophila melanogaster. *Proc. Natl Acad. Sci. USA*, **104**, 6714–6719.
52. Soper,S.F., van der Heijden,G.W., Hardiman,T.C., Goodheart,M., Martin,S.L., de Boer,P. and Bortvin,A. (2008) Mouse maelstrom, a component of nuage, is essential for spermatogenesis and transposon repression in meiosis. *Dev. Cell*, **15**, 285–297.
53. Shoji,M., Tanaka,T., Hosokawa,M., Reuter,M., Stark,A., Kato,Y., Kondoh,G., Okawa,K., Chujo,T., Suzuki,T. *et al.* (2009) The TDRD9-MIWI2 complex is essential for piRNA-mediated retrotransposon silencing in the mouse male germline. *Dev. Cell*, **17**, 775–787.
54. Vagin,V.V., Wohlschlegel,J., Qu,J., Jonsson,Z., Huang,X., Chuma,S., Girard,A., Sachidanandam,R., Hannon,G.J. and Aravin,A.A. (2009) Proteomic analysis of murine Piwi proteins reveals a role for arginine methylation in specifying interaction with Tudor family members. *Genes Dev.*, **23**, 1749–1762.
55. Zhang,Z., Xu,J., Koppetsch,B.S., Wang,J., Tipping,C., Ma,S., Weng,Z., Theurkauf,W.E. and Zamore,P.D. (2011) Heterotypic piRNA Ping-Pong requires qin, a protein with both E3 ligase and Tudor domains. *Mol. Cell*, **44**, 572–584.
56. Reuter,M., Chuma,S., Tanaka,T., Franz,T., Stark,A. and Pillai,R.S. (2009) Loss of the Mili-interacting Tudor domain-containing protein-1 activates transposons and alters the Mili-associated small RNA profile. *Nat. Struct. Mol. Biol.*, **16**, 639–646.
57. Wang,J., Saxe,J.P., Tanaka,T., Chuma,S. and Lin,H. (2009) Mili interacts with tudor domain-containing protein 1 in regulating spermatogenesis. *Curr. Biol.*, **19**, 640–644.
58. Handler,D., Olivieri,D., Novatchkova,M., Gruber,F.S., Meixner,K., Mechtler,K., Stark,A., Sachidanandam,R. and Brennecke,J. (2011) A systematic analysis of Drosophila TUDOR domain-containing proteins identifies Vreteno and the Tdrd12 family as essential primary piRNA pathway factors. *EMBO J.*, **30**, 3977–3993.
59. Zamparini,A.L., Davis,M.Y., Malone,C.D., Vieira,E., Zavadil,J., Sachidanandam,R., Hannon,G.J. and Lehmann,R. (2011) Vreteno, a gonad-specific protein, is essential for germline development and primary piRNA biogenesis in Drosophila. *Development*, **138**, 4039–4050.
60. Huang,H.Y., Houwing,S., Kaaij,L.J., Meppelink,A., Redl,S., Gauci,S., Vos,H., Draper,B.W., Moens,C.B., Burgering,B.M. *et al.* (2011) Tdrd1 acts as a molecular scaffold for Piwi proteins and piRNA targets in zebrafish. *EMBO J.*, **30**, 3298–3308.
61. Kamminga,L.M., Luteijn,M.J., den Broeder,M.J., Redl,S., Kaaij,L.J., Roovers,E.F., Ladurner,P., Berezikov,E. and Ketting,R.F. (2010) Hen1 is required for oocyte development and piRNA stability in zebrafish. *EMBO J.*, **29**, 3688–3700.
62. Mitsunobu,H., Izumi,M., Mon,H., Tatsuke,T., Lee,J.M. and Kusakabe,T. (2011) Molecular characterization of heterochromatin proteins 1a and 1b from the silkworm, Bombyx mori. *Insect Mol. Biol.*, **21**, 9–20.
63. Barski,A., Cuddapah,S., Cui,K., Roh,T.Y., Schones,D.E., Wang,Z., Wei,G., Chepelev,I. and Zhao,K. (2007) High-resolution profiling of histone methylations in the human genome. *Cell*, **129**, 823–837.
64. Filion,G.J., van Bemmel,J.G., Braunschweig,U., Talhout,W., Kind,J., Ward,L.D., Brugman,W., de Castro,I.J., Kerkhoven,R.M., Bussemaker,H.J. *et al.* (2010) Systematic protein location mapping reveals five principal chromatin types in Drosophila cells. *Cell*, **143**, 212–224.
65. Gerstein,M.B., Lu,Z.J., Van Nostrand,E.L., Cheng,C., Arshinoff,B.I., Liu,T., Yip,K.Y., Robilotto,R., Rechtsteiner,A., Ikegami,K. *et al.* (2010) Integrative analysis of the Caenorhabditis elegans genome by the modENCODE project. *Science*, **330**, 1775–1787.
66. Kharchenko,P.V., Alekseyenko,A.A., Schwartz,Y.B., Minoda,A., Riddle,N.C., Ernst,J., Sabo,P.J., Larschan,E., Gorchakov,A.A., Gu,T. *et al.* (2011) Comprehensive analysis of the chromatin landscape in Drosophila melanogaster. *Nature*, **471**, 480–485.
67. Negre,N., Brown,C.D., Ma,L., Bristow,C.A., Miller,S.W., Wagner,U., Kheradpour,P., Eaton,M.L., Loriaux,P., Sealfon,R. *et al.* (2011) A cis-regulatory map of the Drosophila genome. *Nature*, **471**, 527–531.
68. Roy,S., Ernst,J., Kharchenko,P.V., Kheradpour,P., Negre,N., Eaton,M.L., Landolin,J.M., Bristow,C.A., Ma,L., Lin,M.F. *et al.* (2010) Identification of functional elements and regulatory circuits by Drosophila modENCODE. *Science*, **330**, 1787–1797.
69. Riddle,N.C., Minoda,A., Kharchenko,P.V., Alekseyenko,A.A., Schwartz,Y.B., Tolstorukov,M.Y., Gorchakov,A.A., Jaffe,J.D., Kennedy,C., Linder-Basso,D. *et al.* (2011) Plasticity in patterns of histone modifications and chromosomal proteins in Drosophila heterochromatin. *Genome Res.*, **21**, 147–163.
70. Lee,Y., Kim,M., Han,J., Yeom,K.H., Lee,S., Baek,S.H. and Kim,V.N. (2004) MicroRNA genes are transcribed by RNA polymerase II. *EMBO J.*, **23**, 4051–4060.
71. Moshkovich,N. and Lei,E.P. (2010) HP1 recruitment in the absence of argonaute proteins in Drosophila. *PLoS Genet.*, **6**, e1000880.
72. Fujiwara,H., Osanai,M., Matsumoto,T. and Kojima,K.K. (2005) Telomere-specific non-LTR retrotransposons and telomere maintenance in the silkworm, Bombyx mori. *Chromosome Res.*, **13**, 455–467.
73. Kawaoka,S., Hayashi,N., Katsuma,S., Kishino,H., Kohara,Y., Mita,K. and Shimada,T. (2008) Bombyx small RNAs: genomic defense system against transposons in the silkworm, Bombyx mori. *Insect Biochem. Mol. Biol.*, **38**, 1058–1065.