Report

A Transgenerational Process Defines piRNA Biogenesis in Drosophila virilis

Graphical Abstract

Highlights

- piRNA cluster activity in D. virilis correlates with high levels of H3K9me3
- A transgenerational signal induces piRNA biogenesis in progeny
- An inherited maternal signal causes H3K9me3 deposition at piRNA-producing regions

Authors

Adrien Le Thomas, Georgi K. Marinov, Alexei A. Aravin

Correspondence

aaa@caltech.edu

In Brief

Small noncoding RNA molecules called piRNAs, encoded by specific genomic regions, silence the expression of selfish transposable elements in Drosophila germ cells. What defines a genomic region for piRNA production is unknown. In this study, Le Thomas et al. show that active piRNA biogenesis requires high levels of H3K9me3 that cover the piRNA-generating regions. Furthermore, piRNA generation in the progeny requires maternal deposition of an epigenetic signal, presumably a piRNA, that establishes the H3K9me3 mark.

Accession Numbers

GSE59965
A Transgenerational Process Defines piRNA Biogenesis in Drosophila virilis

Adrien Le Thomas,1,2 Georgi K. Marinov,1 and Alexei A. Aravin1,*
1California Institute of Technology, Division of Biology and Biological Engineering, 147-75, 1200 East California Boulevard, Pasadena, CA 91125, USA
2Ecole Doctorale Complexité du Vivant, Université Pierre et Marie Curie, 4 place Jussieu, 75005 Paris, France
*Correspondence: aaa@caltech.edu
http://dx.doi.org/10.1016/j.celrep.2014.08.013
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/3.0/).

SUMMARY

Piwi-interacting (pi)RNAs repress diverse transposable elements in germ cells of Metazoa and are essential for fertility in both invertebrates and vertebrates. The precursors of piRNAs are transcribed from distinct genomic regions, the so-called piRNA clusters; however, how piRNA clusters are differentiated from the rest of the genome is not known. To address this question, we studied piRNA biogenesis in two D. virilis strains that show differential ability to generate piRNAs from several genomic regions. We found that active piRNA biogenesis correlates with high levels of histone 3 lysine 9 trimethylation (H3K9me3) over genomic loci that give rise to piRNAs. Furthermore, piRNA biogenesis in the progeny requires the transgenerational inheritance of an epigenetic signal, presumably in the form of homologous piRNAs that are generated in the maternal germline and deposited into the oocyte. The inherited piRNAs enhance piRNA biogenesis through the installment of H3K9me3 on piRNA clusters.

INTRODUCTION

A distinct class of small RNAs called Piwi-interacting RNAs, or piRNAs, provides sequence specificity for the recognition and repression of a diverse set of invasive genetic elements in the germline of metazoans (Aravin et al., 2007; Siomi et al., 2011). piRNAs are loaded into members of the Piwi clade of Argonaute proteins. Piwi/piRNA complexes in Drosophila repress transposable elements through two different mechanisms. In the nucleus, piRNAs associated with PIWI are responsible for the deposition of H3K9me3 on homologous transposon targets (Le Thomas et al., 2013; Rozhkov et al., 2013; Sienski et al., 2012). In the cytoplasm, two other Piwi proteins, AUB and AGO3, are guided by the associated piRNAs to cleave and destroy homologous transposon transcripts using their endonuclease activity. Simultaneously, their associated piRNAs to cleave and destroy homologous transposon plasm, two other Piwi proteins, AUB and AGO3, are guided by the piRNAs associated with PIWI are responsible for the deposition of H3K9me3 on piRNA clusters.

RESULTS

To understand the determinants of piRNA cluster activity and identify the features that discriminate piRNA-generating regions
from other genomic loci, we took advantage of our previous finding of piRNA clusters that are differentially active in two D. virilis strains (Rozhkov et al., 2010). Three genomic regions designated as clusters #1, #2, and #3 are present in the genomes of both strains 9 and 160, yet piRNAs are only generated from these clusters in strain 160 (we will refer to strain 160 as the active or A strain) (Figures 1 and 2A). Strain 9 (the inactive or I strain) does not have general defects in the piRNA pathway as piRNAs are generated from other genomic regions in this strain. The reason for the differential ability to generate piRNAs from certain genomic regions in the two strains is unknown.

Recent studies revealed that two chromatin factors, SETDB1, a methyltransferase responsible for installation of H3K9me3, and the HP1 paralog Rhino that binds this mark are required for piRNA biogenesis in D. melanogaster (Klattenhoff et al., 2009; Rangan et al., 2011). We hypothesized that a difference in the chromatin state of the genomic regions in strains A and I might explain their differential ability to produce piRNAs. To determine whether chromatin state plays a role in the differential activity of the piRNA clusters, we profiled H3K9me3 in ovaries from flies of both strains. For each strain, chromatin immunoprecipitation sequencing (ChIP-seq) libraries from two independent biological samples were generated, sequenced, and analyzed. Genomic regions that generate piRNAs exclusively in strain A (clusters #1–3) had significantly higher levels of H3K9me3 in this strain compared to strain I (Figures 1, 2A, 2B, and S1A). In contrast, cluster #4, which produced a similar amount of piRNAs in both strains, had comparable levels of H3K9me3 (Figures 2A, 2B, and S1A). The H3K9me3 profile on piRNA clusters in strain A closely parallels the profile of piRNA generation: both signals drop at the same genomic position (Figure 1). Therefore, the activity of native piRNA-generating regions in D. virilis correlates with high levels of the H3K9me3 mark for all three differentially expressed clusters.

To understand how H3K9me3 affects the activity of piRNA-generating loci, we used ChIP-seq to profile genome-wide RNA polymerase II occupancy and RNA-seq to measure transcript levels in ovaries from strains A and I. We found that high H3K9me3 levels on differentially expressed clusters in strain A correlated with high occupancy of RNA polymerase II on these regions (Figures 2C and S1B). Furthermore, the piRNA clusters gave rise to more precursor transcripts in strain A than in strain I, indicating that the enrichment of Pol II at these sites represents polymerase engaged in active transcription (Figures 2D and S2). Together, these results suggest that high levels of H3K9me3 on piRNA clusters correlate with the generation of precursor transcripts for piRNA processing.

To further understand the properties of genomic regions that generate piRNAs, we studied the activity of piRNA clusters in the progeny of the cross between strains A and I. The cross between the two strains can be performed in two different directions: the progeny of the cross between strain A females and strain I males was designated as MD (for maternal deposition) as the A chromosomes and corresponding piRNAs are inherited from the mothers (Figure 3A). In the progeny of the opposite cross, designated as NMD (for no maternal deposition), the A chromosomes are inherited from the fathers, and no cognate piRNAs are inherited. Importantly, the genotypes of MD and NMD progeny are absolutely identical.
Profiling of piRNAs in ovaries of MD and NMD flies showed that similar amounts of piRNAs were generated from cluster #4, which is equally expressed in both parental strains (Figure 3B). Surprisingly, ~3-fold more piRNAs were generated from the three differentially active clusters in MD compared to NMD progeny. Because the genotypes of MD and NMD progeny are identical, this result indicates that the transgenerational inheritance of an epigenetic signal that is transmitted from the mothers to their progeny enhances the ability to generate piRNAs.

One mechanism of piRNA biogenesis involves the ping-pong amplification loop. In the ping-pong loop, piRNAs guide cleavage of complementary targets leading to the generation of new piRNAs from the cleaved product (Brennecke et al., 2007; Gunawardane et al., 2007; Olovnikov and Kalmykova, 2013). These secondary piRNAs belong to complementary piRNA pairs. In the MD progeny, the majority of piRNA sequences coming from clusters 1–3 can form pairs with a 10 nt overlap between their 5’ ends (so called ping-pong pairs). In contrast, only a small fraction of piRNAs form ping-pong pairs in the NMD progeny. It must be noted, however, that, because of limited sequencing depth, this fraction underestimates the real proportion of secondary piRNAs, especially for the piRNA pairs that produce few reads of one partner. In an attempt to account for the different number of sequences derived from differentially expressed clusters in MD and NMD progenies, we sampled the same number of reads from each library and calculated the fraction of reads in ping-pong pairs a thousand times. A similar proportion of secondary piRNAs was detected in strain A and in the MD progeny; however, the fraction of piRNAs forming ping-pong pairs in clusters #1, #2, and #3 was dramatically reduced in the NMD progeny, compared to the fraction in the MD progeny (Figure 3D). A similar fraction of piRNAs mapping to cluster #4 formed ping-pong pairs in strain A and both the MD and the NMD progeny, indicating that the ping-pong machinery was intact in both progenies. These data suggest that an epigenetic signal inherited by MD progeny from their mothers eventually boosts the biogenesis of secondary piRNAs.

Next, we studied the chromatin state of differentially active clusters in MD and NMD progenies. We compared the levels of H3K9me3 using separate ChIP-qPCR and ChIP-seq experiments performed on two independent biological samples of MD and NMD ovaries (Figures 4A and S3). In both progenies, H3K9me3 levels over the clusters were intermediate between the strong and weak enrichment seen in strain A and strain I, respectively. Importantly, both methods show that H3K9me3 enrichment was higher in ovaries of MD compared to NMD flies, indicating that the maternally supplied epigenetic signal affects chromatin of piRNA clusters in the progeny (Figures 4A and S3). The identical genomes of the MD and the NMD progenies each contain two alleles of differentially active piRNA clusters that have different H3K9me3 levels in their parents, the A and the I strains. To understand the impact of the maternally supplied epigenetic signal on individual alleles, it is critical to differentiate sequences derived from the A and I chromosomes. We therefore carried out whole-genome sequencing of the genomes of both MD and NMD progenies. Because the genotypes of MD and the NMD progenies each contain two alleles of differentially active piRNA clusters that have different H3K9me3 levels in their parents, the A and the I strains. To understand the impact of the maternally supplied epigenetic signal on individual alleles, it is critical to differentiate sequences derived from the A and I chromosomes. We therefore carried out whole-genome sequencing of the genomes of both MD and NMD progenies.
strains. We generated 82 times and 71 times coverage for strains A and I, respectively, and identified SNPs for each strain relative to the D. virilis reference genome (see the Experimental Procedures for details). We found 326,026 and 1,086,963 SNPs in the genomes of A and I, respectively, of which, 169,192 SNPs were shared. A total of 1,074,605 SNPs differed between the genomes of the two strains (average density of 5.19 such SNPs per Kb) allowing us to determine the allelic origin of piRNAs and ChIP-seq reads derived from polymorphic genomic regions. Importantly, we identified such SNPs in each of the differentially expressed piRNA clusters (Table S1). This allowed us to unambiguously map a fraction of the ChIP-seq and piRNA reads to each individual allele (Figures 4B–4D and S4).

We found that the two alleles of differentially expressed piRNA clusters maintained their distinct chromatin states in the hybrid progeny. Levels of H3K9me3 were higher on the A alleles compared to the I alleles in both the MD and the NMD ovaries (Figures 4B and S4). However, MD progenies had higher levels of H3K9me3 mark on both the A and I alleles. This result indicates that maternally inherited epigenetic signal in MD flies (Figures 4C and 4D). MD progeny correlated with the induction of piRNA generation from previously inactive genomic loci (Figures 4C and 4D). Therefore, transgenerational inheritance of an epigenetic signal correlates with deposition of the H3K9me3 mark and induction of piRNA generation from previously inactive genomic loci (Figure 4E).

DISCUSSION

Our results reveal an essential role for a transgenerationally inherited epigenetic factor in generation of piRNAs in germ cells of D. virilis. This maternally inherited factor is required to maintain the high level of piRNA generation from active piRNA clusters. Furthermore, it is able to activate previously inactive loci or H chromosomes. These findings parallel the study by de Vanssay and colleagues who showed that a maternally inherited factor produced by a transgenic piRNA locus is able to activate piRNA generation from a previously inactive homologous locus in D. melanogaster (de Vanssay et al., 2012). Multiple lines of evidence point to piRNAs themselves as the carriers of the epigenetic signal that triggers piRNA generation from homologous loci in the progeny. First, in Drosophila both piRNAs and Piwi proteins are inherited from the maternal germline to the early embryos (Brennecke et al., 2008). Second, piRNAs can serve as sequence-specific guides to identify and activate homologous loci. Recent studies have shown that piRNAs and the nuclear Piwi protein trigger installation of the H3K9me3 mark on homologous targets providing a possible mechanism by which inherited piRNAs could lead to chromatin changes (Le Thomas et al., 2013; Rozhkov et al., 2013; Sienski et al., 2012; Wang and Elgin, 2011). Finally, the study of a similar phenomenon in D. melanogaster showed that the epigenetic signal produced by the transgenic piRNA locus does not require inheritance of the locus itself (de Vanssay et al., 2012). This result indicates that the epigenetic signal has a nonchromosomal nature and
eliminates the possibility that it is carried by a chromatin mark linked to the active locus. Our results, together with these previous studies, strongly support the role of inherited piRNAs as a transgenerationally inherited epigenetic signal that activates piRNA generation from homologous loci in the progeny.

Transgenerationally inherited piRNAs activate piRNA generation in a nuclear process that is linked to the deposition of H3K9me3 on homologous genomic regions. We found that acquisition of the H3K9me3 mark by genomic regions that did not previously produce piRNAs correlates with initiation of piRNA biogenesis (Figure 4). In contrast, absence of inherited piRNAs leads to a decrease in H3K9me3 levels on homologous regions and a concomitant decrease in the corresponding piRNAs (Figure 4). These results suggest that modification of the chromatin structure of genomic regions homologous to inherited piRNAs induces piRNA biogenesis in the progeny. In agreement with our results, previous studies have shown that the biogenesis of piRNAs in D. melanogaster requires the activity of Eggless/SETDB1, one of the methyl-transferases carrying out the trimethylation of H3K9 (Rangan et al., 2011) and of Rhino, the HP1 homolog, that is enriched over piRNA clusters (Klattenhoff et al., 2009; Zhang et al., 2014). Rhino has a chromodomain that is similar to the chromodomain of HP1 and binds the H3K9me3 mark (Le Thomas et al., 2014; Mohn et al., 2014).

The exact molecular mechanism by which the H3K9me3 mark is linked to piRNA biogenesis in the nucleus remains to be elucidated; however, our results suggest that, counterintuitively, high levels of H3K9me3 correlate with increased transcription in the context of piRNA clusters. Indeed, we found that high levels of H3K9me3 on differentially expressed piRNA clusters in D. virilis correlated with increased Pol II occupancy and with the generation of more precursor transcripts (Figures 1 and 2). Therefore,
we propose that changes in chromatin state associated with high levels of H3K9me3 lead to the recruitment of nonconventional readers of this mark such as Rhino, which subsequently affect transcription in these regions, providing more precursors for piRNA biogenesis. Overall, our data revealed that the determination of a piRNA producing loci in \textit{D. virilis} is mediated by the process of the transgenerational epigenetic inheritance.

**EXPERIMENTAL PROCEDURES**

**Drosophila Stocks**

\textit{D. virilis} strain 160 (A) and strain 9 (I) were a generous gift from M. Evgenyev.

**ChIP and RNA-Seq**

ChIP experiments were carried out using commercially available antibodies anti-H3K9me3 (ab8898) and anti-RNA Pol II (ab5408). Ovaries were fixed for 10 min at room temperature using 1% paraformaldehyde (PFA) followed by 5 min quenching by directly adding glycine (final concentration 25 mM). A detailed protocol is provided in the Supplemental Experimental Procedures. Quantitative ChIP-qPCR experiments were performed in at least two biological replicates with two technical replicates each. Error bars represent the SEM. Primer pairs used in the qPCR experiments are presented in Table S3. Chip-seq and RNA-seq library construction and sequencing were carried out using standard protocols and sequenced on the Illumina HiSeq 2000 (50 bp reads). Publicly available data sets for piRNAs were extracted from the GEO Short Read Archive GSE22067 (Rozhkov et al., 2010).

**Analysis of Ping-Pong Processing**

To determine the fraction of piRNAs that participate in ping-pong pairs, we counted uniquely mapped piRNA reads that map to opposite strands of each other, overlap, and have a 10 bp distance between their 5' ends. To account for different number of piRNAs in each sample, we carried out the analysis by sampling reads to ~25% of the read counts in the library with the fewest reads in a region and repeating this 1,000 times.

**D. virilis Genome Sequencing and Allele-Specific Mapping**

Genomic libraries were generated from each \textit{D. virilis} strain and sequenced (as 2 \times 100 reads) on an Illumina HiSeq 2000 platform. Sequencing coverage was 25% of the read counts in the library with the fewest reads in a region and repeating this 1,000 times.

**D. virilis Genome Sequencing and Allele-Specific Mapping**

ACCESSION NUMBERS

High-throughput sequencing data for ChiP-seq and RNA-seq experiments are available through the Gene Expression Omnibus (accession number GSE59965).

**SUPPLEMENTAL INFORMATION**

Supplemental Information includes Supplemental Experimental Procedures, four figures, and four tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2014.08.013.

**AUTHOR CONTRIBUTIONS**

A.L.T. and A.A.A. designed the experiments, A.L.T. performed the experiments, G.K.M. analyzed the deep sequencing data, and A.L.T, G.K.M., and A.A.A. interpreted the data and wrote the manuscript.

**ACKNOWLEDGMENTS**

We thank Katalin Fejes Toth and members of the Aravin lab for discussion. This work was supported by grants from the NIH (R01 GM097363 and DP2 OD007371A) and by the Damon Runyon-Rachleff, the Searle Scholar, and the Packard Fellowship Awards to A.A.A.

**REFERENCES**


