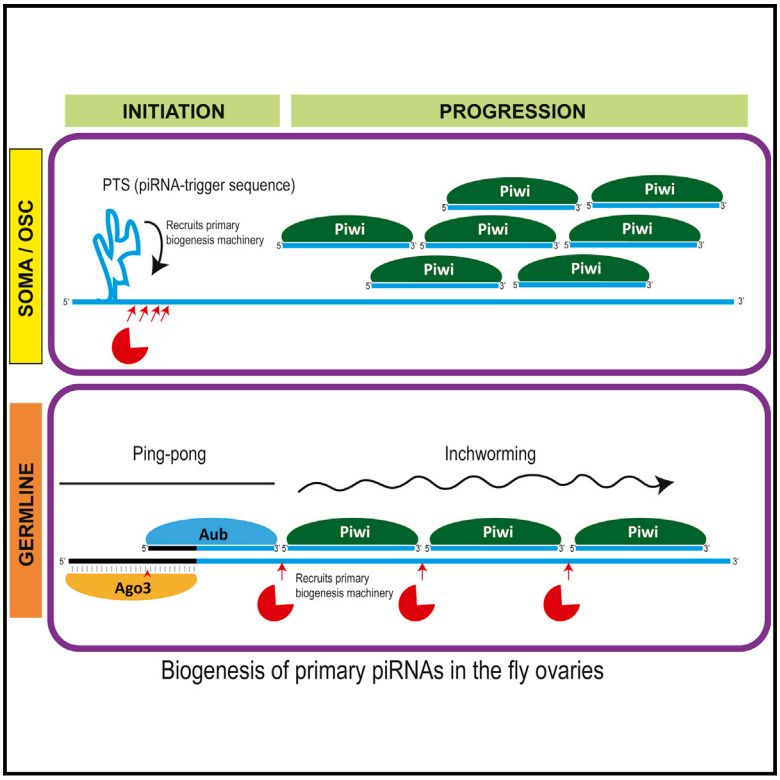
Article

Cell Reports

PIWI Slicing and RNA Elements in Precursors Instruct Directional Primary piRNA Biogenesis

Graphical Abstract



Biogenesis of primary piRNAs in the fly ovaries

Authors

David Homolka, Radha Raman Pandey, Coline Goriaux, ..., Ravi Sachidanandam, Marie-Odile Fauvarque, Ramesh S. Pillai

Correspondence

pillai@embl.fr

In Brief

Homolka et al. find that targeting a transcript with the PIWI endonuclease triggers piRNA production from downstream sequences. The presence of piRNA-trigger sequences is also sufficient to render any transcript a substrate for primary piRNA biogenesis.

Highlights

- A piRNA-trigger sequence from the 5′ end of a fly cluster drives primary processing
- Slicer activity of PIWI generates non-overlapping, contiguous primary piRNAs
- Cytoplasmic PIWI-triggered primary piRNAs are loaded into nuclear PIWI
- Primary piRNA biogenesis proceeds with a 5′–3′ directionality

Accession Numbers

GSE69102

Homolka et al., 2015, Cell Reports *12*, 418–428
July 21, 2015
http://dx.doi.org/10.1016/j.celrep.2015.06.030

OPEN ACCESS
CellPress

Cell Reports
Article

PIWI Slicing and RNA Elements in Precursors Instruct Directional Primary piRNA Biogenesis

David Homolka,[1,2,6] Radha Raman Pandey,[1,2,6] Coline Goriaux,[3] Emilie Brasset,[3] Chantal Vaury,[3] Ravi Sachidanandam,[4] Marie-Odile Fauvarque,[5] and Ramesh S. Pillai[1,2,*]

[1]European Molecular Biology Laboratory, Grenoble Outstation, University Grenoble Alpes-EMBL-CNRS, 71 Avenue des Martyrs, 38042 Grenoble Cedex 9, France
[2]Unit for Virus Host-Cell Interactions, University Grenoble Alpes-EMBL-CNRS, 71 Avenue des Martyrs, 38042 Grenoble Cedex 9, France
[3]Clermont Université, Université d'Auvergne, 28 Place Henri Dunant, 63000 Clermont-Ferrand, France
[4]Department of Oncological Sciences, Icahn School of Medicine at Sinai, One Gustave L. Levy Place, New York, NY 10029, USA
[5]iRTSV, CEA-Grenoble, 17, rue des Martyrs, 38054 Grenoble Cedex 9, France
[6]Co-first author
*Correspondence: pillai@embl.fr
http://dx.doi.org/10.1016/j.celrep.2015.06.030

SUMMARY

PIWI proteins and PIWI-interacting RNAs (piRNAs) mediate repression of transposons in the animal gonads. Primary processing converts long single-stranded RNAs into ~30-nt piRNAs, but their entry into the biogenesis pathway is unknown. Here, we demonstrate that an RNA element at the 5′ end of a piRNA cluster—which we termed piRNA trigger sequence (PTS)—can induce primary processing of any downstream sequence. We propose that such signals are triggers for the generation of the original pool of piRNAs. We also demonstrate that endonucleolytic cleavage of a transcript by a cytosolic PIWI results in its entry into primary processing, which triggers the generation of non-overlapping, contiguous primary piRNAs in the 3′ direction from the target transcript. These piRNAs are loaded into a nuclear PIWI, thereby linking cytoplasmic post-transcriptional silencing to nuclear transcriptional repression.

INTRODUCTION

PIWI-interacting RNAs (piRNAs) are a class of 24- to 30-nt small RNAs that are exclusively expressed in animal gonads and are loaded into the PIWI clade members of the Argonaute protein family (Ghildiyal and Zamore, 2009). Together, they have an essential role in the suppression of mobile elements called transposons in the germline genome, ensuring proper germ cell development and fertility (Luteijn and Ketting, 2013). One major question in the field is how piRNAs are made. A few hundred genomic regions (50–100 kilobases [kb]), called piRNA clusters, generate the majority of the small RNAs, while 3′ UTRs of numerous protein-coding transcripts and transposon sequences contribute the rest (Brennecke et al., 2007; Lau et al., 2009). All these are transcribed by RNA polymerase II (RNA pol II) into long single-stranded transcripts (Li et al., 2013) and exported to the cytoplasm for processing within perinuclear granules called nuage (Lim and Kai, 2007).

Primary processing refers to a pathway by which these long single-stranded transcripts are converted into thousands of overlapping small RNAs that display a clear preference for having a uridine (U) as the first residue (1U bias) (Brennecke et al., 2007). How these RNA pol II transcripts are distinguished from others as substrates and which nucleases generate the 5′ and 3′ ends are not known (Luteijn and Ketting, 2013). There is more clarity in secondary processing or the Ping-pong cycle (Brennecke et al., 2007; Gunawardane et al., 2007), where piRNA targets themselves become substrates for new piRNAs. Once a target is sliced by PIWI endonuclease activity, the downstream cleavage fragment is transferred to a new PIWI protein, where it matures as a secondary piRNA (Nishida et al., 2015; Xiol et al., 2014). Thus, PIWI slicing creates the 5′ end of the secondary piRNA; however, the mechanism of 3′ end formation is still unclear, although a 3′–5′ exonuclease activity (tentatively called Trimmer) is implicated (Kawaoka et al., 2011).

Here, we used artificial reporters with unique sequences to track piRNA biogenesis in insect cell lines and the *Drosophila* germline. Our results reveal two distinct mechanisms for recruiting transcripts into the primary processing pathway. First, reporter RNA carrying a piRNA-trigger sequence (PTS) from the 5′ end of a fly cluster transcript get processed in a directional manner into thousands of overlapping primary piRNAs. Second, cytosolic PIWI slicing of a target not only generates secondary piRNAs but also results in further processing of the transcript into non-overlapping primary piRNAs. These piRNAs are loaded into a nuclear PIWI, thus establishing a link between cytosolic silencing and transcriptional repression. Taken together, our results demonstrate that RNA elements in precursor transcripts and PIWI slicing serve to recruit the primary piRNA processing machinery.

RESULTS

BmAgo3 Slicing Generates Primary piRNAs in the *Bombyx* BmN4 Cells

The *Bombyx mori* (Silkworm) ovary-derived cell line BmN4 is a useful cell culture model for studying piRNA biogenesis. It

expresses two cytosolic PIWI proteins (Siwi and BmAgo3) that bind piRNAs with sequence signatures indicative of biogenesis via secondary processing (Kawaoka et al., 2009; Xiol et al., 2012). We have previously used BmN4 cells to produce artificial piRNAs (Xiol et al., 2014). Briefly, a reporter transcript was targeted by Siwi-loaded endogenous piRNAs, resulting in its cleavage by Siwi slicer activity and the transfer of one of the cleavage fragments as a new artificial secondary piRNA into BmAgo3. Notably, all reporter-derived piRNAs in BmAgo3 have their 5′ ends generated by Siwi slicing and so are of secondary origin (Figures S1A and S1B).

In this study, we made the reverse reporter construct: a sequence that is targeted by multiple independent BmAgo3-loaded endogenous piRNAs (Figure 1A; see the Supplemental Experimental Procedures). The reporter was expressed from a transfected plasmid in BmN4 cells, and 48 hr post-transfection, PIWI-bound small RNAs were analyzed by deep sequencing. Mapping of reads to the reporter sequence indicated the presence of endogenous antisense targeting piRNAs that are specifically enriched in BmAgo3 and map to the designated binding sites on the reporter (Figure 1A). Then, we examined the presence of reporter-derived artificial sequences (Figure 1B). As expected from the Ping-pong cycle (Kawaoka et al., 2009), abundant levels of artificial piRNAs are detected almost exclusively in Siwi complexes (Figure 1B). The majority of these are of secondary origin, as their 5′ ends overlap by 10 nt with the 5′ ends of the targeting BmAgo3-bound guides (peaks lying within the gray-shaded regions in Figure 1C). Thus, slicing by BmAgo3 generates 5′ ends of these Siwi-bound secondary piRNAs.

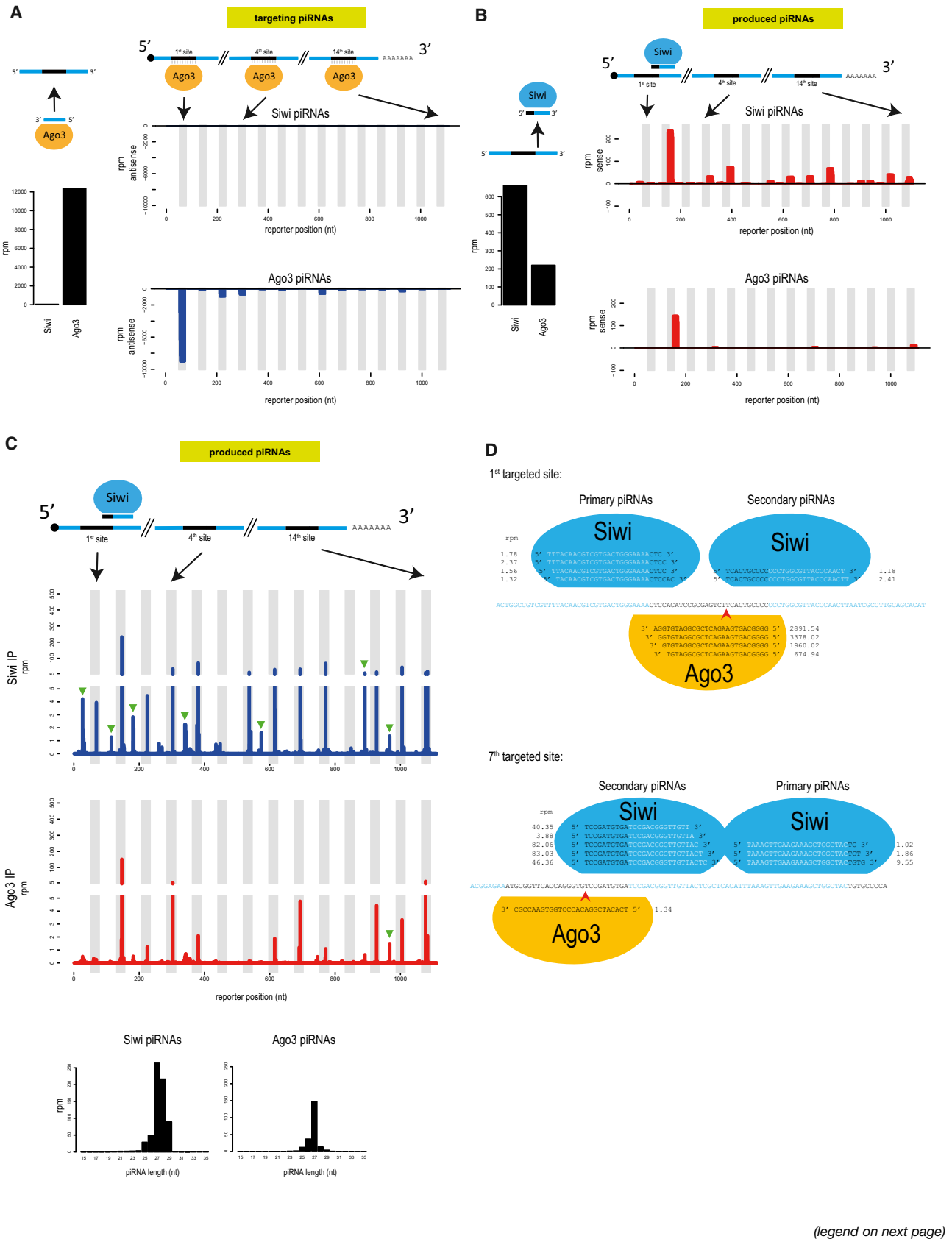
Unexpectedly, we also observed the presence of low levels of reads that do not show such a 10-nt overlap signature with any endogenous piRNA, indicating that the 5′ ends of these are not generated by slicing (see the green arrowheads in Figure 1C). The majority of such piRNAs are loaded into Siwi complexes, although some are also seen in BmAgo3 (Figure 1C). All these reads display a 1U bias, and we consider them to be true primary piRNAs. Curiously, their 5′ ends lie at a distance of ~30 nt from that of a secondary piRNA (Figure 1D). This suggests a mechanism for 5′ end formation that is non-random but linked to the initiation of secondary biogenesis on a transcript. Closer examination of one such primary piRNA read (near target site #7) reveals that its 5′ end is immediately downstream of the 3′ end of the secondary piRNA (Figure 1D). This suggests a scenario where generation of the 5′ end of the primary piRNA by an unknown endonuclease simultaneously generates the 3′ end of the secondary piRNA that is further shortened by trimming (Kawaoka et al., 2011). However, another primary piRNA (near target site #1) is generated from the upstream cleavage fragment (Figure 1D), suggesting that the primary processing machinery might be opportunistically using both slicer cleavage fragments. Similar primary piRNAs were generated from a second independent reporter targeted by BmAgo3 (Figures S1C–S1E). These experiments demonstrate that BmAgo3 slicing in *Bombyx* BmN4 cells not only generates the expected secondary piRNA loaded into Siwi but also unexpectedly triggers the biogenesis of primary piRNAs.

Cytoplasmic Slicing by Ago3 Triggers 30-nt Steps of Primary piRNAs Entering Nuclear Piwi

The piRNA pathway in the *Bombyx* BmN4 cell culture model mirrors the Ping-pong relationship seen between cytosolic *Drosophila* PIWI proteins Aubergine (Aub) and Ago3 in the fly ovarian germline (Brennecke et al., 2007; Kawaoka et al., 2009), so we wanted to determine whether PIWI slicing might trigger primary piRNAs in the fly germline. To this end, we prepared transgenic flies driving germline-specific expression of a noncoding reporter targeted by multiple independent endogenous piRNAs. The 5′ segment of the reporter was targeted by Aub-enriched piRNAs, while the 3′ segment was targeted by those present in Ago3 (Figure 2A; Supplemental Experimental Procedures). Working with two independent transgenic fly lines, we prepared deep-sequencing libraries from total ovarian small RNAs (15–28 nt) (Figure S2A). Their analysis demonstrates the expected targeting of the reporter transcript by endogenous antisense piRNAs (Figure 2A) and generation of reporter-derived (sense-oriented) piRNAs (Figure 2B). Many of these artificial piRNAs display the expected 10-nt overlap signature (peak at position −10), indicating that their 5′ ends are generated by PIWI slicing (Figures 2C and S2B–S2E). This was further confirmed by the sequencing of small RNAs from immunoprecipitated PIWI complexes (Figures 2E and S2A). Aub contained the targeting antisense piRNAs (Figure 2A), while Ago3 bound the reporter-derived secondary piRNAs generated by Aub slicing (Figure 2B), and vice versa. In fact, the majority of artificial piRNAs present in Aub (Figure 2F) and Ago3 (Figure S2D) are secondary piRNAs, as their 5′ ends are generated by reciprocal slicing.

Nuclear Piwi is the third PIWI clade member expressed in the fly germline, whose piRNAs we also examined (Figure 2E). Piwi libraries contained reporter-derived piRNAs distributed along the entire length of the reporter sequence, but with a majority arising from the segment targeted by Ago3 (Figure 2B). Intriguingly, the Piwi-bound reads did not map precisely within the region of the reporter targeted by endogenous piRNAs, as seen for artificial secondary piRNAs in Aub and Ago3 (gray-shaded region in Figure 2B). To examine their origins in detail, we first mapped their 5′ ends on the reporter sequence (Figure 2G). Only a minority of the 5′ ends of Piwi piRNAs are generated by direct Ago3 slicing (peak at position −10). Instead, a strong cluster of 5′ ends (at positions 17–23; site #1) is observed, which corresponds to a distance of ~30 nt downstream of the 5′ end of the Aub-bound secondary piRNA. Another cluster of 5′ ends (at positions 43–48; site #2) is detected a further ~30 nt downstream. The produced Piwi and Aub piRNAs exhibit strong 5′ U bias (Figure 2D). However, the starting nucleotides of Aub piRNAs are influenced by reporter design (as the selected Ago3 guides have a tenth position A). This is not the case for Piwi-loaded piRNAs, whose 5′ U bias is a consequence of the biogenesis process. Such 1U-containing Piwi-bound artificial piRNAs are true primary piRNAs, as their 5′ ends are not directly generated by Ago3 slicing but by an unknown nuclease that is recruited downstream after Ago3 slicer action on the reporter.

It is interesting that mapping of the mature 3′ ends of all the reporter-derived piRNAs (Figures 2F and 2G) suggests a piRNA processing mechanism where the same endonuclease that



(legend on next page)

generates the 5′ end of Piwi piRNAs also creates the 3′ end of the preceding piRNA. For example, endonucleolytic cleavages creating the 5′ ends of Piwi piRNAs at site #1 also creates the 3′ ends of the preceding Aub-bound piRNA, which undergoes further trimming to their mature size. Likewise, cleavages creating the 5′ ends of Piwi piRNAs at site #2 creates the 3′ ends of the preceding Piwi piRNA (with the 5′ end at site #1), but we are unable to infer the level of 3′ trimming involved in this case. Our reporter design (presence of multiple Ago3 cleavage sites and lack of sequences downstream) does not allow examination of further 5′/3′ end generations downstream, but it is expected to follow a pattern where Piwi is the recipient of these 1U-containing piRNAs (Figure 2D). Taken together, our data show that cytoplasmic slicing by Ago3 in the fly germline can not only trigger biogenesis of secondary piRNAs present in Aub but also launch waves of non-overlapping primary piRNAs from downstream sequences that load nuclear Piwi (Figure 2H). Note that Aub slicing also triggers generation of Piwi-bound primary piRNAs downstream (Figures S2D–S2G), but at much lower levels (Figure 2B).

A PTS Directs Primary Processing

In the aforementioned experiments, a transcript becomes a substrate for primary biogenesis when it is a target for PIWI slicing in the fly ovarian germline, and this mechanism involves multiple PIWI clade members. A much simpler piRNA pathway exists in the ovarian somatic follicle cells that surround the germline in the fly ovaries, where nuclear Piwi is the only PIWI clade member expressed (Li et al., 2009; Malone et al., 2009), so we wanted to examine how primary piRNAs are made in the soma, especially when Piwi slicer activity is shown to be dispensable for piRNA biogenesis (Darricarrère et al., 2013).

For these studies, we used the ovarian somatic cell (OSC) culture system, which is representative of the fly ovarian soma (Saito et al., 2009). The follicle cells express piRNAs from a variety of sources, including the major piRNA cluster *flamenco* (*flam*) and 3′ UTRs of select protein-coding transcripts (Brennecke et al., 2007; Saito et al., 2009). It is known that the *flam* transcript is alternatively spliced, with the exon 1 being present in every spliced variant detected in fly ovaries (Goriaux et al., 2014). To examine any potential importance of this sequence, we fused a fragment consisting of the *flam* promoter and the *flam* transcript region (1–718 nt) encompassing exon 1 to the firefly luciferase coding sequence. After transfection into OSC, endogenous Piwi complexes were isolated, and associated small RNAs were analyzed by deep sequencing (Figure 3A). Notably, piRNAs originating from the reporter sequence are detected, with sequences mapping to the very 5′ end of the transcript, i.e., from the *flam* element, extending into the luciferase sequence, stopping around the site of polyadenylation (Figure 3B). Since the same pattern is seen when expression was driven from fly *actin* promoter (Figure 3C), we conclude that the *flam* transcript fragment alone can trigger primary piRNA biogenesis from the reporter.

The piRNAs mapping to the *flam* element are likely to be a mixture of those arising from the endogenous *flam* cluster locus and those generated from the reporter transcript (Figure S3B). The pattern of piRNA production is not random but appears to be the consequence of a processive machinery that cleaves almost every nucleotide within the entire length of the transcript (Figure S3C). In fact, the piRNAs with 5′ ends starting in the *flam* fragment overflow seamlessly into the 5′ part of the luciferase region (Figure S3C). The artificial piRNAs so generated display a strong 1U bias, a signature feature of primary processing (Figure 3B). Hereinafter, we refer to the *flam* fragment (1–718 nt) as the PTS.

However, the situation is not that simple, as our control luciferase reporter expressed in the absence of the PTS also produced a low level of 1U-containing sequences in Piwi complexes (Figures 3B–3C). We do not consider these to be random degradation products, as they are not uniformly distributed over the entire reporter but start abruptly from the middle of the transcript and proceed until the 3′ end, leaving the 5′ end devoid of any such reads. Note that the 5′ end of the reporter is converted to piRNAs only in the presence of the upstream *flam* PTS. We do not find evidence for any chance complementarity with an endogenous piRNA being the trigger to initiate this processing (Figure S3D). Furthermore, insertion of a perfectly complementary binding site for a Piwi-bound piRNA within the luciferase reporter did not initiate primary processing in the OSC system (data not shown). We consider that the luciferase reporter sequence has an unknown signal that allows it to recruit the primary processing machinery at an internal location, perhaps very similar to that seen for protein-coding transcripts that enter the pathway (Robine et al., 2009).

PTS Can Function at an Internal Location in a Transcript

The *flam* PTS element when present in *cis*, was able to trigger piRNA production of the entire transcript region downstream.

Figure 1. BmAgo3 Slicer Activity Triggers Primary Processing in the *Bombyx* BmN4 Cells

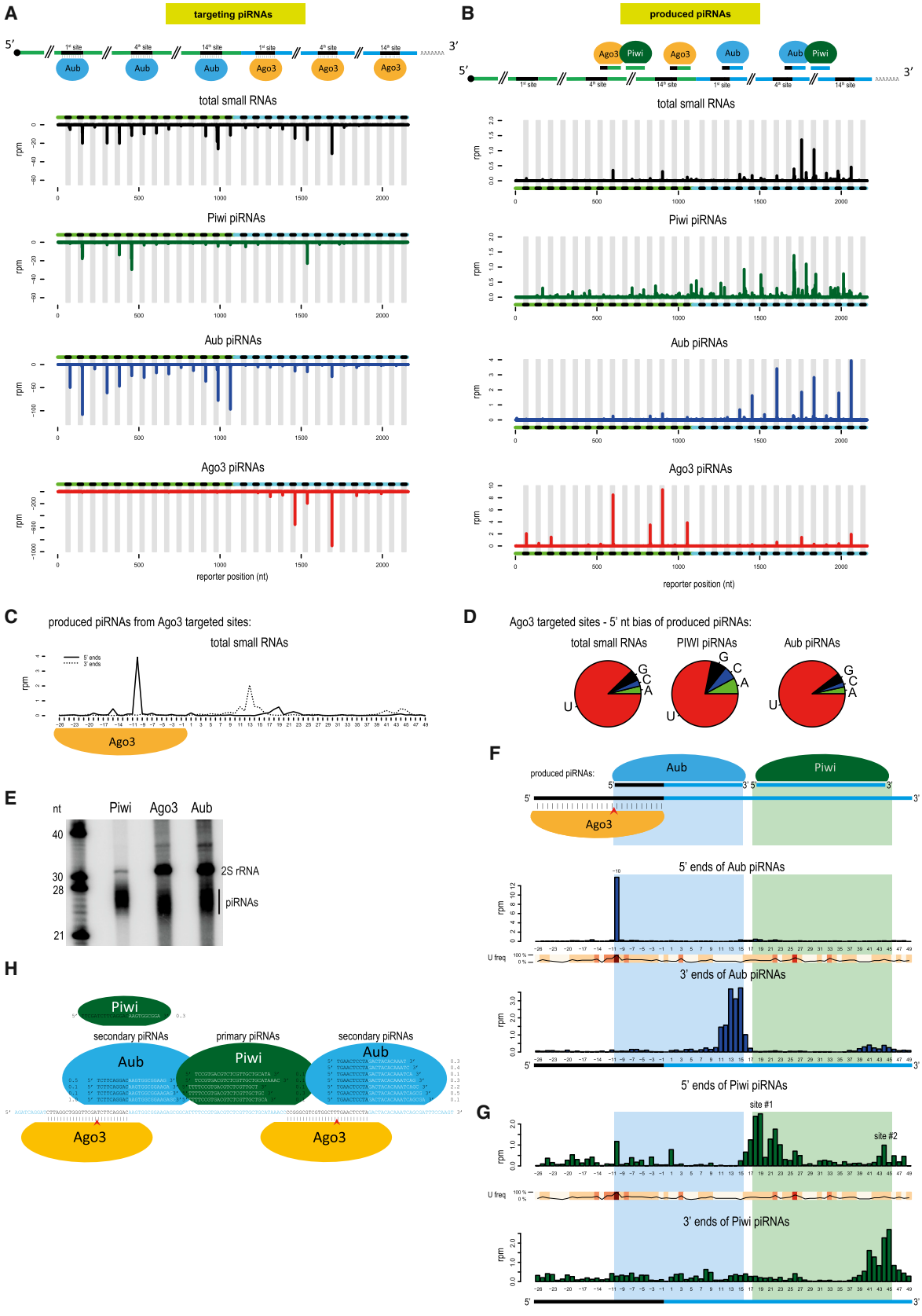
(A) Design of the artificial secondary piRNA precursor. The noncoding LacZ sequence has perfectly complementary binding sites for 14 piRNAs abundantly present in endogenous Ago3 complexes in BmN4 cells. The density plot (reads per million; rpm) shows Siwi and Ago3 bound piRNAs complementary to the artificial precursor. As designed, the transcript is targeted almost exclusively by Ago3 piRNAs. The shaded regions indicate the Ago3-piRNA-binding sites on the reporter.

(B) Reporter-derived piRNAs produced from the precursor transcript is shown. As expected, most of these secondary piRNAs originate from the regions targeted by Ago3 piRNAs and are preferentially loaded onto Siwi.

(C) The 5′ ends of produced artificial piRNAs are mapped along the precursor transcript. The y axis showing the piRNA abundance is scaled to show both low and high abundant reads. Notably, besides the typical secondary piRNAs that start within the region complementary to Ago3 guides (gray shaded region), other piRNAs are produced that do not show overlap with any targeting piRNAs (green arrowheads). However, such piRNAs are much less abundant. Length distributions of Siwi and Ago3 piRNAs are plotted at the bottom. IP, immunoprecipitation.

(D) Sequence-level detail of guiding and produced piRNAs around the first and seventh sites on the reporter transcript are shown. The transcript cleavage (red arrowhead) creates the 5′ end of secondary piRNAs, while primary piRNAs are produced outside the targeted regions.

See also Figure S1.



(legend on next page)

One possibility is that any transcript containing the PTS element recruits the machinery to its 5′ end, which then proceeds in a 5′–3′ direction, converting the entire transcript into piRNAs. Such a mechanism will have to involve decapping to remove the 5′ protective cap structure to gain access to the body of the transcript. To test this hypothesis, we moved the PTS to an internal location.

Given that the aforementioned luciferase transcript that was used has an inherent property to recruit the primary processing machinery at an internal location, we searched for ''clean'' reporter sequences. Noncoding EGFP and LacZ transcripts, when expressed on their own from the *actin* promoter, were not substrates for the piRNA pathway (Figure 4A). Therefore, we placed the PTS between the EGFP and LacZ sequences (Figure 4A). Analysis of Piwi-associated small RNAs reveals that the PTS can function when present in an internal location and in a directional manner, as the downstream LacZ sequences now become converted into 1U-containing primary piRNAs. One way to interpret this result is by considering that the PTS recruits an unknown endonuclease to an internal location, perhaps within the PTS element, to initiate piRNA production from any sequence downstream. Structured sequences such as tRNA, when placed downstream of the PTS, also become processed into overlapping, 1U-containing piRNAs (Figures 4B and S3C), indicating that the processing machinery harbors RNA helicase activities. We find that RNA elements embedded within transcripts direct their entry into primary processing, but the true identity of such a non-sequence conserved PTS remains to be determined.

Transcription Factor Binding Site in a Mouse Cluster Promoter Is Not Essential for Primary Processing

So far, our experiments point to a role for RNA elements within precursors in recruiting the insect primary processing machinery, and this can be uncoupled from transcription driven by piRNA cluster promoters. We wondered whether this is also true for mouse clusters that generate piRNAs via primary processing. Mouse male germ cells that enter the pachytene stage of meiosis express abundant levels of piRNAs from a set of 100 pachytene clusters that contain mostly unique, non-repeated sequences (Girard et al., 2006). Pachytene cluster transcripts are processed by primary biogenesis factors and do not depend on slicer activity of mouse PIWI proteins Mili and Miwi (Di Giacomo et al., 2013; Reuter et al., 2011; Zheng and Wang, 2012). The transcription factor A-Myb is shown to have a binding site proximal to the transcription start site in each of these clusters, and pachytene piRNA biogenesis is impacted in *A-Myb* mutant mice (Li et al., 2013). The bi-directional cluster on chromosome 17 (Chr17: 27425220–27504428) has a single A-Myb binding site (GACAGTTA) in the central promoter (Figure 5A). To test its importance, we made mutant mice (pi-Chr17$^{\Delta Myb}$) carrying a 228-base-pair (bp) deletion in the cluster promoter that removes the A-Myb binding site and the flanking sequences (Figure 5A). This resulted in an expected reduction in precursor transcript levels from the cluster, as determined by RNA sequencing (RNA-seq) analysis (Figures 5B and S4). However, precursor transcription is not entirely turned off, allowing us to question whether these transcripts are now processed into piRNAs.

We isolated Miwi and Mili from testes lysates prepared from heterozygous and homozygous pi-Chr17$^{\Delta Myb}$ animals, and we analyzed associated piRNAs by deep sequencing. Genome mapping indicated unchanged production of piRNAs in the homozygous mutant testes from all the top 50 pachytene clusters, except the bidirectional Chr17 cluster lacking the A-Myb site (Figures 5B, S4A, and S4B). Notably, the piRNA levels from the Chr17 cluster are reduced but not completely abrogated, with major deficiencies seen only in the antisense strand of the cluster region (Figure 5C). The relative levels of the piRNAs produced correlated with that of the precursor transcript detected in the homozygous mutant animals (Figures 5C and S4B). This indicates that, although the A-Myb binding site may contribute to enhancing Chr17 cluster transcription, it does

Figure 2. Cytoplasmic Ago3 Slicing in the Fly Germline Loads Nuclear Piwi with Primary piRNAs

(A) Design of the *Drosophila* reporter transcript with binding sites for Aub and Ago3 piRNAs (shaded areas). The artificial transcript contains the non-coding GFP part with 14 sites complementary to endogenous Aub piRNAs followed by the non-coding LacZ part having 14 sites complementary to endogenous Ago3 piRNAs. The 5′ ends of targeting piRNAs present in total small RNAs and PIWI immunoprecipitations are mapped along the transcript. As designed, these map to the binding sites on the transcript.

(B) piRNA-guided cleavage of the transcript leads to production of reporter-derived piRNAs. Mapping of reporter-derived reads present in total small RNA libraries and PIWI complexes is shown. As expected, artificial piRNAs from region targeted (shaded areas) by Ago3 enter Aub, and vice versa. In Piwi complexes, most of the artificial piRNAs arise from the LacZ part of the transcript targeted by Ago3.

(C) Mapping of the 5′ and 3′ ends of produced artificial piRNAs relative to the 5′ end (nucleotide position −1) of the targeting Ago3 piRNA. Read counts from total small RNA libraries were aggregated for the 14 individual Ago3-targeted sites. The strong 5′-end peak at −10 nt indicates generation of most artificial piRNAs by Ago3 slicing. Notice that another small 5′-end peak lies immediately downstream of the 3′-end peak at +13 nt.

(D) Produced piRNAs show strong 5′ U bias.

(E) 5′-end labeling of piRNA populations isolated from Piwi, Aub, and Ago3 complexes from fly ovaries.

(F) Mapping the 5′ and 3′ ends of Aub and Piwi piRNAs derived from the reporter relative to the 5′ end (position −1) of the targeting Ago3 piRNA. Read counts were aggregated from 14 individual Ago3-targeted sites. The 5′ end of Aub piRNA is created by Ago3 slicing and, therefore, lies at the −10 nt position. The 3′ ends (+13 to +15 nt) of Aub piRNAs reveal variable lengths of Aub piRNAs, potentially as a consequence of trimming.

(G) The 5′ ends of Piwi piRNAs start at site #1 (+17 to +23 nt), immediately downstream the Aub piRNAs, and end at positions +41 to +45. Another smaller peak for 5′ ends of the subsequent Piwi piRNA is observed at positions at site #2 (+43 to +48 nt).

(H) Detailed sequence-level view of the reporter sequence showing two Ago3-targeted sites and the individual reporter-derived artificial piRNAs formed in their vicinity. The transcript cleavage by Ago3 (red arrowhead) creates the 5′ end of secondary piRNAs, which are loaded into Aub. However, other piRNAs are also produced downstream from these piRNAs, which are exclusively loaded onto Piwi, and its 5′ end formation is not as a result of Ago3 slicing.

See also Figure S2.

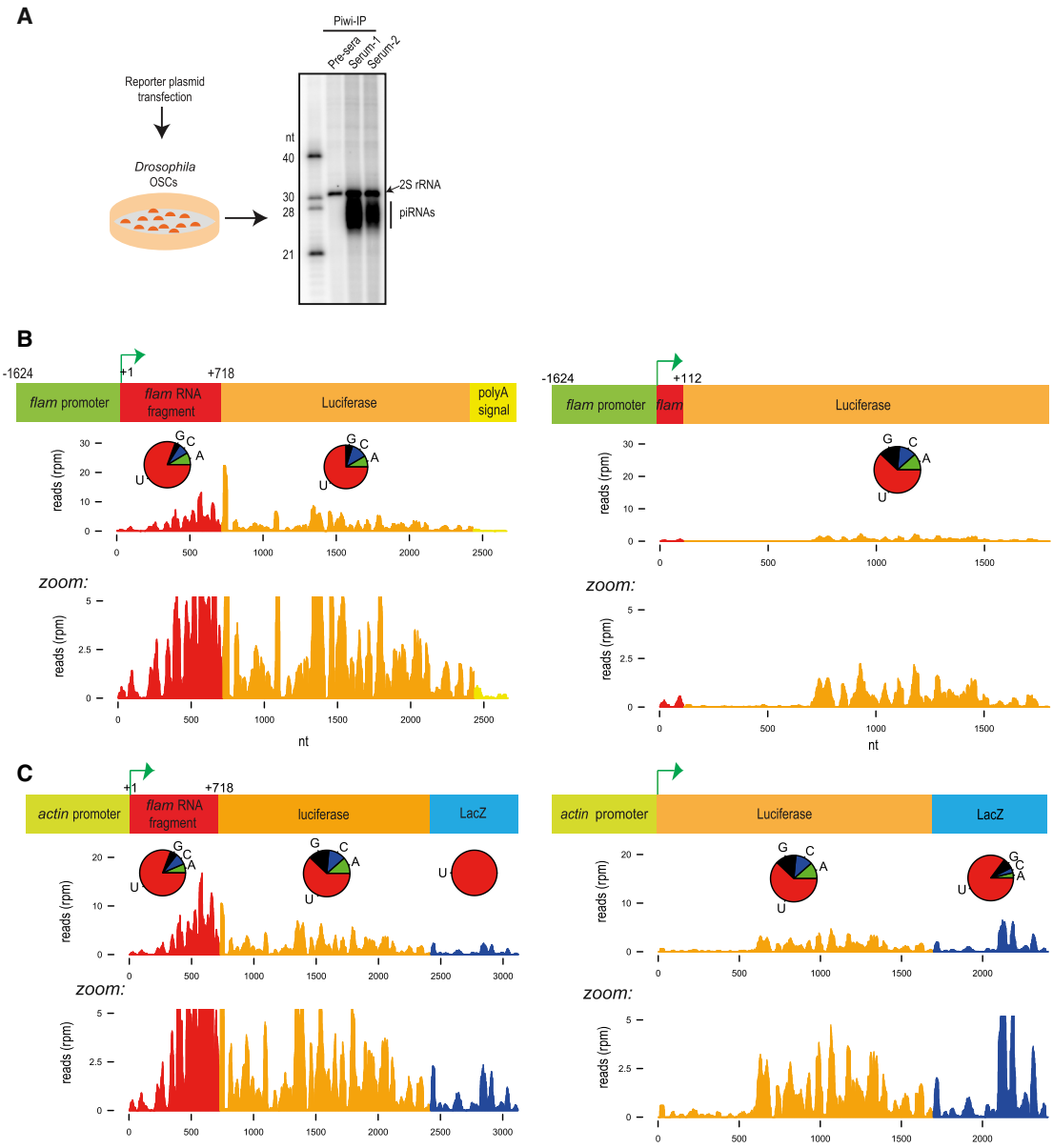


Figure 3. A PTS Element in the *flamenco* Transcript Can Direct Generation of Primary piRNAs

(A) Investigation of piRNA biogenesis in the *Drosophila melanogaster* OSCs. After transfection of the desired plasmids into the cells, piRNAs present in the PIWI clade member Piwi were examined by deep sequencing. IP, immunoprecipitation.

(B) Mapping of reads from cells transfected with constructs expressing the luciferase sequence from the *flamenco* promoter and additionally carrying the indicated length of the *flamenco* transcript (in red). Note that the luciferase sequence alone has an inherent property to trigger low levels of piRNAs from an internal location. The pie charts show the 5′-nt preference of the produced piRNAs.

(C) Mapping of reads from cells expressing the indicated reporters whose expression was driven by the fly *actin* promoter.

See also Figures S3 and S5A.

not determine the cytoplasmic processing fate of the resulting transcript. We also report that pi-Chr17$^{\Delta Myb}$ homozygous mice of both sexes are viable and fertile. This indicates that, at least for this mouse cluster that we examined, nuclear history from transcription by the A-Myb transcription factor has no role in cytosolic processing of cluster transcripts by the primary biogenesis machinery.

DISCUSSION

Recruitment of a Transcript into the Primary Processing Pathway

How piRNAs are made is a long-standing question in the field. Here, we addressed this mainly in the *Drosophila* ovarian environment by using reporter transcripts expressed in the OSC

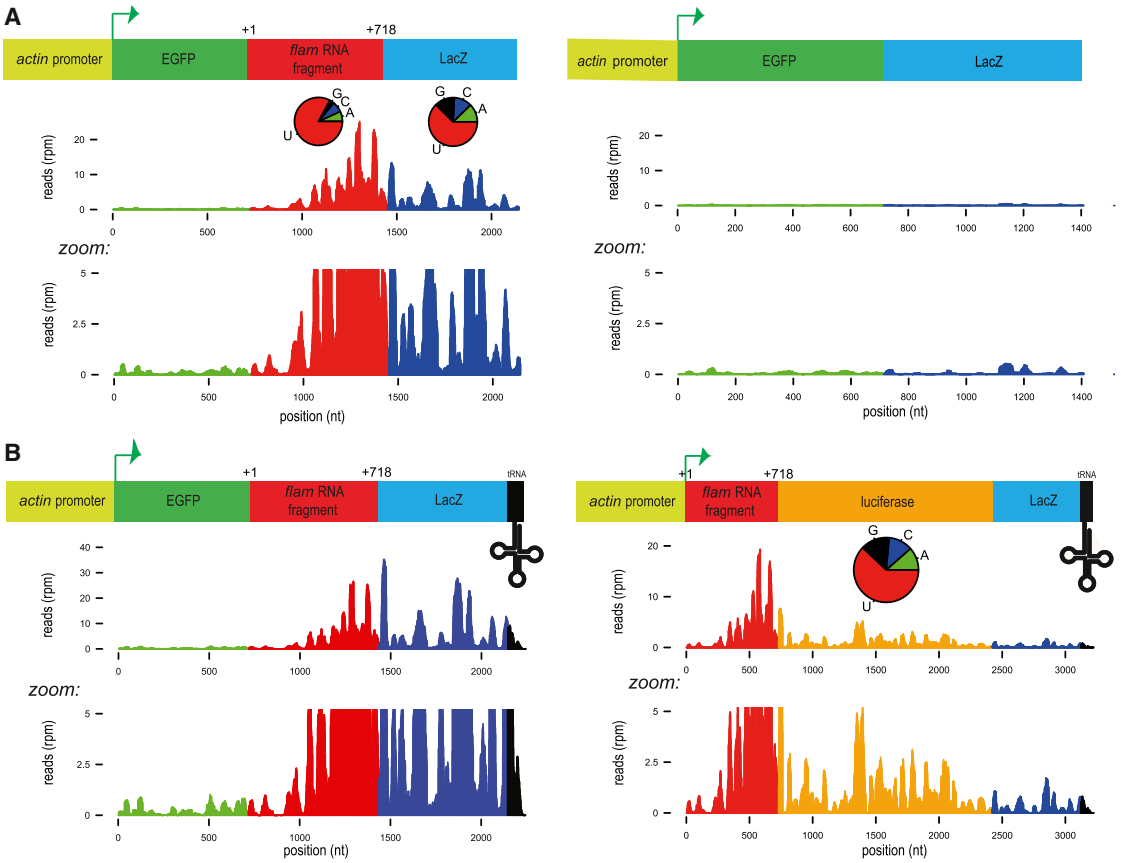


Figure 4. PTS Triggers piRNA Biogenesis from Downstream Sequences

(A) The piRNAs are produced from LacZ sequences present downstream of the *flam* PTS element. No such piRNAs are detected when the PTS element is absent. The pie charts show the 5′-nt preference of the produced piRNAs.

(B) piRNAs are produced from the *Wolbachia* tRNA sequence that is inserted downstream of *flam* PTS fragment.

See also Figures S3C, S5C, S6A, and S6B.

and the fly germline. It is known that piRNA precursors are pol II transcripts, long single-stranded RNAs with a 5′ cap and 3′ poly(A) tail, and are hard to distinguish from other cellular transcripts. Our studies now reveal the existence of two distinct mechanisms by which RNA elements can recruit the primary piRNA processing machinery to initiate processing (Figure 6). In the fly germline, slicing by Ago3 results in the generation of the expected secondary piRNA in Aub (Figure 2). Additionally, it also triggers production of primary piRNAs from downstream sequences in ~30-nt steps, which are loaded into nuclear Piwi. Notably, Aub slicing did not lead to robust primary piRNA biogenesis, likely reflecting mechanistic differences in the protein complexes mediating the ultimate fate of the cleavage fragments. Thus, PIWI slicing identifies a transcript as a precursor for primary processing.

In the OSC system, we showed that exon 1 of the fly piRNA cluster transcript *flam* is sufficient to identify reporter transcripts as precursors for primary processing. We termed this RNA element the PTS. This sequence is conserved in other drosophilids, as a similar sequence in the *flam* locus from *D. simulans* is functional in promoting primary processing in the *D. melanogaster* OSC system (Figures S5A and S5B). However, the situation is more complex, as our experiments reveal that the luciferase reporter transcript itself can recruit the processing machinery to an internal location (Figure 3B). This is also the case with a 6-kb fragment of the human Malat1 RNA, as it very efficiently became processed into piRNAs in the OSC (Figure S5D). Thus, either the PTS could be a structural motif, or it could contain a small, as-yet unrecognized motif that binds a piRNA biogenesis factor (Figures S6A and S6B) (Vourekas et al., 2015).

Inchworm Model of Primary Processing

How does the primary biogenesis machinery progress along the precursor? The answer to this is obtained by examination of primary processing initiated by Ago3 slicing in the fly germline (Figure 2). After Ago3-mediated slicing of the target RNA, the fragment with the 5′ monophosphate is loaded into Aub. This event also recruits the primary processing machinery to the downstream sequences. An unknown activity within the machinery then cleaves the transcript a few nucleotides (~5 nt) downstream of the Aub footprint (Figure 2F). This cleavage event

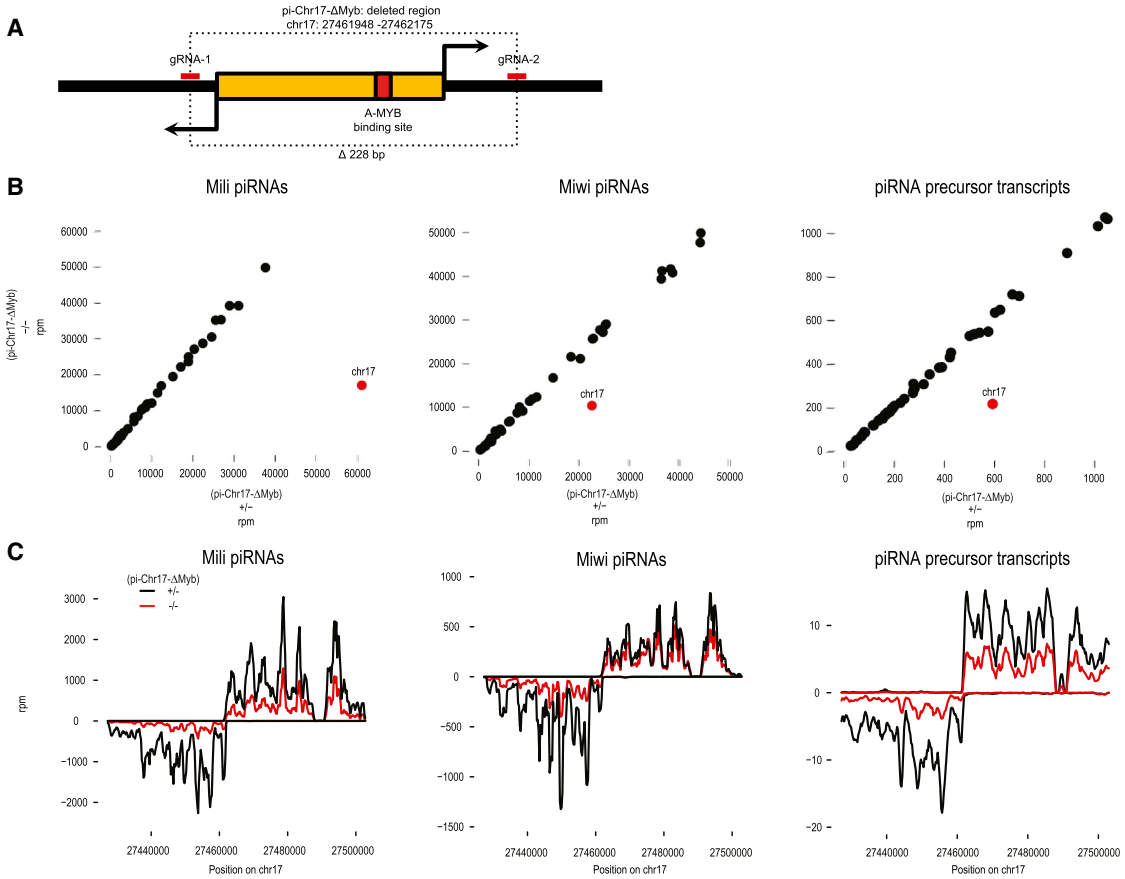


Figure 5. Nuclear Transcription History Does Not Define Mouse Cluster Transcripts for Processing

(A) The promoter of the chr17 bi-directional mouse pachytene piRNA cluster encompassing the A-MYB binding site and 5′ starts of piRNA precursor transcripts was deleted by the Cas9 endonuclease, using the indicated guide RNAs (gRNAs).

(B) Abundance of Mili- and Miwi-associated piRNAs was compared between pi-Chr17$^{\Delta Myb+/-}$ and pi-Chr17$^{\Delta Myb-/-}$ mice for the top 50 piRNA producing clusters. Only piRNA production from the chr17 cluster harboring the deletion was affected. The decreased piRNA production correlates with the low amount of piRNA precursor transcripts.

(C) Mili- and Miwi-associated piRNAs were mapped along the Chr17 cluster. The deletion of the promoter element significantly affects piRNA production, especially from the minus strand, which correlates with decreased abundance of piRNA precursor transcripts.

See also Figure S4.

simultaneously generates the 3′ end of the Aub-bound piRNA, which needs shortening by 3′–5′ exonucleolytic trimming, and also the 5′ end of a new downstream Piwi-bound piRNA. This process is likely to be repeated until the entire downstream transcript region is consumed so that generated primary piRNAs are loaded into Piwi. Thus, the core mechanism of primary processing is the movement of the processing machinery like an inchworm along the transcript to cleave and simultaneously generate 5′ and 3′ ends of new primary piRNAs (Figure 6). It moves in steps of ~30 nt that are dictated, in part, by the footprint of the PIWI protein. Processivity of the reaction is likely provided by the grip the machinery has on the downstream transcript sequence.

Why, then, are the piRNAs generated from the PTS-containing transcripts tiled at almost single-nucleotide resolution (Figures S6C–S6F)? One possible explanation could be that the PTS might allow multiple initiation sites on the transcript, giving rise to overlapping 30-nt steps of primary piRNAs (Figure 6). Moreover, the cleaved products are not precisely spaced, leading a loss of phase and an appearance of piRNAs initiating from every position on the transcript (Figure 2G). This could also suggest that the processing machinery might not be using the PIWI footprint as the only guide for identifying cleavage sites but might also be searching for U residues. One caveat is that only 1U-containing piRNAs are stably associated with PIWI proteins and enriched, due to the nucleotide preference of the MID domain (Cora et al., 2014), and thus represented in our sequencing libraries. Verification of this hypothesis awaits the identification of such a U-specific nuclease activity within the primary processing machinery (Luteijn and Ketting, 2013).

Slicer-Triggered Primary Processing Links Cytoplasmic Silencing to Nuclear Repression

In the *Drosophila* ovarian germline cells, post-transcriptional silencing by piRNAs results in amplification of specific piRNAs

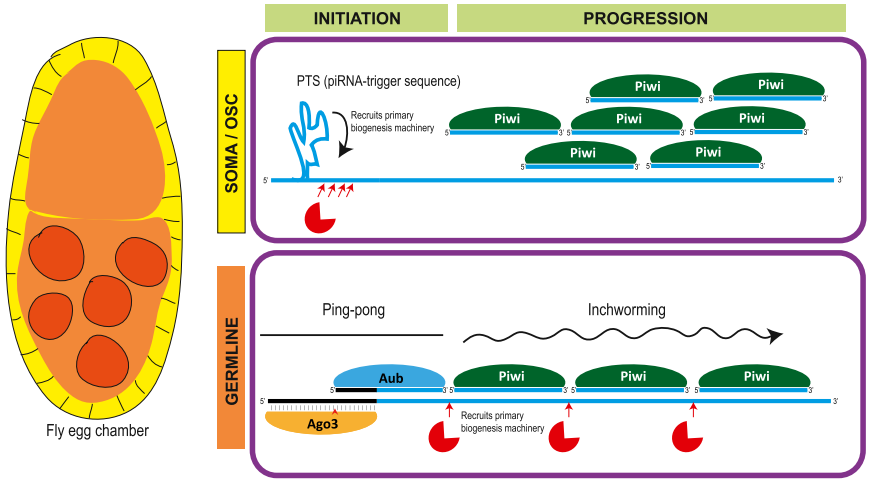


Figure 6. Model for Primary Processing in the Fly Ovarian System

The *Drosophila* egg chamber is composed of the germline (orange) that is surrounded by a single layer of protective somatic follicle cells (yellow). Nuclear Piwi is the only PIWI clade member expressed in the soma or its cell-culture model (OSC). The germline has Piwi and two cytosolic members: Aub and Ago3. In the soma, the PTS present at the 5′ end of a fly cluster transcript is sufficient to drive any heterologous transcript into primary processing. In the fly germline, Aub and Ago3 are slicers that participate in the ping-pong cycle to generate secondary piRNAs. Ago3 slicing additionally recruits the primary processing machinery to generate piRNAs in ~30-nt steps from downstream sequences to load nuclear Piwi. See also Figures S6C–S6F.

by the cytoplasmic PIWI proteins Aub and Ago3 (Brennecke et al., 2007; Gunawardane et al., 2007). The slicer-trigger initiation mechanism described here for primary piRNA biogenesis is a new component of this adaptive arm of the piRNA pathway in insects. It serves to increase the repertoire of piRNA sequences, resulting in an enhancement of the silencing potential of the pathway (Figure 2). Genetic studies in flies also support the link between the ping-pong cycle and the loading of primary piRNAs into nuclear Piwi, as several biogenesis factors have overlapping roles in both processes (Malone et al., 2009). Additionally, the slicer-trigger pathway for primary biogenesis offers an explanation for several previous observations that link cytoplasmic piRNA biogenesis to nuclear outcomes in the *Drosophila* ovarian germline and early embryos (de Vanssay et al., 2012; Le Thomas et al., 2014). Finally, can this slicer-triggered primary piRNA biogenesis pathway be conserved? It is already known that biogenesis of piRNAs associating with nuclear mouse PIWI protein Miwi2 is dependent on the cytoplasmic PIWI protein Mili and its slicer activity (Aravin et al., 2008; De Fazio et al., 2011), suggesting a potential conservation of this piRNA biogenesis mechanism.

EXPERIMENTAL PROCEDURES

Clones and Constructs

For expression in the *Bombyx mori* ovary-derived cell line BmN4 (Kawaoka et al., 2009), constructs were prepared in the pBEMBL vector that drives expression from an *OpIE2* promoter (from *Orgyia pseudotsugata* multicapsid nucleopolyhedrosis virus) (Xiol et al., 2012). Reporters targeted by endogenous piRNAs were prepared as described previously (Xiol et al., 2014). Expression in the *Drosophila melanogaster* OSCs (Saito et al., 2009) was driven using either the promoter of the fly piRNA cluster *flamenco* (Goriaux et al., 2014) or that of the *actin* gene. See the Supplemental Experimental Procedures for detailed sequence information on all constructs used.

Fly Experiments

A reporter based on noncoding (all ATGs were mutated) EGFP and LacZ backbone sequences and containing perfectly complementary binding sites for endogenous piRNAs present in fly ovaries was chemically synthesized (Shanghai ShineGene Molecular Biotech). The reporter sequence was cloned into a vector providing the *vasa* promoter and 5′ UTR sequences to drive germline-specific expression in the fly ovaries. This was used to generate fly transgenes carrying a single-copy insertion at a targeted location in the genome (Genetics Services).

Cell Culture and Immunoprecipitations

The culture and transfection of *Bombyx mori* ovary-derived cell line BmN4 (Kawaoka et al., 2009) were as previously described (Xiol et al., 2012). Cells were collected 48 hr post-transfection and used for immunoprecipitations. The culture and transfection of *Drosophila melanogaster* OSCs were as described previously (Saito et al., 2009).

Rabbit polyclonal antibodies to *Drosophila* PIWI proteins were as follows: Piwi (two rabbits: GJKO and GJLD) raised against an antigen (42–178 amino acids [aa]), Aub (1–200 aa), and Ago3 (1–200 aa). Rabbit polyclonal antibodies to detect *Bombyx* Siwi and Ago3 were previously described (Xiol et al., 2012). Antibodies to mouse Mili (mouse monoclonal) and Miwi (rabbit polyclonal) were previously described (Reuter et al., 2011).

All mouse experiments were conducted in line with national French and European regulations.

Small RNA Libraries and Bioinformatics

Unless otherwise indicated, all piRNA libraries described in this study come from small RNAs isolated by PIWI immunoprecipitations from transfected cells or transgenic fly ovaries or from mouse testes. Additional total small RNA libraries were prepared from transgenic fly ovaries. Library preparation and sequence analysis details can be found in the Supplemental Experimental Procedures.

ACCESSION NUMBERS

Deep-sequencing data generated in this study were deposited to the NCBI GEO under accession number GEO: GSE69102 and are described in Table S1.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, and one table and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2015.06.030.

AUTHOR CONTRIBUTIONS

D.H. analyzed all the bioinformatics data, prepared all the figures, and conducted all the experiments for Figures 1 and 2. R.R.P. conducted all the experiments for Figures 3, 4, and 5. C.G., E.B., and C.V. provided the plasmid with the *flam*-luciferase reporter. R.S. mapped mouse deep-sequencing data to the genome and provided coordinates/read-count information. M.-O.F. performed fly crosses and provided expert advice to D.H. for fly dissections.

R.S.P coordinated the study and wrote the article with inputs from D.H. and R.R.P.

ACKNOWLEDGMENTS

We thank Kate M. O'Connor-Giles (University of Wisconsin) for her kind gift of plasmid. We thank the following European Molecular Biology Laboratory (EMBL) Core Facilities: GeneCore for deep sequencing and Protein Expression and Purification for antibody production. We are grateful for fellowships from the EMBL Interdisciplinary Postdoc Programme under Marie Curie COFUND Actions (R.R.P.) and the European Molecular Biology Organization for post-doctoral fellowships (R.R.P and D.H). This work was supported by grants from the European Union (ERC Starting Grant "pisilence"), Fondation Recherche Médicale (DEP20131128529), and Agence National de la Recherche (GuidedMethylation) to R.S.P. Work in the R.S.P. lab is supported by the EMBL.

Received: March 12, 2015
Revised: May 12, 2015
Accepted: June 5, 2015
Published: July 9, 2015

REFERENCES

Aravin, A.A., Sachidanandam, R., Bourc'his, D., Schaefer, C., Pezic, D., Toth, K.F., Bestor, T., and Hannon, G.J. (2008). A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. Mol. Cell *31*, 785–799.

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.

Cora, E., Pandey, R.R., Xiol, J., Taylor, J., Sachidanandam, R., McCarthy, A.A., and Pillai, R.S. (2014). The MID-PIWI module of Piwi proteins specifies nucleotide- and strand-biases of piRNAs. RNA *20*, 773–781.

Darricarrère, N., Liu, N., Watanabe, T., and 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.

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.

de Vanssay, A., Bougé, A.L., Boivin, A., Hermant, C., Teysset, L., Delmarre, V., Antoniewski, C., and Ronsseray, S. (2012). Paramutation in Drosophila linked to emergence of a piRNA-producing locus. Nature *490*, 112–115.

Di Giacomo, M., Comazzetto, S., Saini, H., De Fazio, S., Carrieri, C., Morgan, M., Vasiliauskaite, L., Benes, V., Enright, A.J., and O'Carroll, D. (2013). Multiple epigenetic mechanisms and the piRNA pathway enforce LINE1 silencing during adult spermatogenesis. Mol. Cell *50*, 601–608.

Ghildiyal, M., and Zamore, P.D. (2009). Small silencing RNAs: an expanding universe. Nat. Rev. Genet. *10*, 94–108.

Girard, A., Sachidanandam, R., Hannon, G.J., and Carmell, M.A. (2006). A germline-specific class of small RNAs binds mammalian Piwi proteins. Nature *442*, 199–202.

Goriaux, C., Desset, S., Renaud, Y., Vaury, C., and Brasset, E. (2014). Transcriptional properties and splicing of the flamenco piRNA cluster. EMBO Rep. *15*, 411–418.

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.

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.

Kawaoka, S., Izumi, N., Katsuma, S., and Tomari, Y. (2011). 3′ end formation of PIWI-interacting RNAs in vitro. Mol. Cell *43*, 1015–1022.

Lau, N.C., Robine, N., Martin, R., Chung, W.J., Niki, Y., Berezikov, E., and Lai, E.C. (2009). Abundant primary piRNAs, endo-siRNAs, and microRNAs in a Drosophila ovary cell line. Genome Res. *19*, 1776–1785.

Le Thomas, A., Stuwe, E., Li, S., Du, J., Marinov, G., Rozhkov, N., Chen, Y.C., Luo, Y., Sachidanandam, R., Toth, K.F., et al. (2014). Transgenerationally inherited piRNAs trigger piRNA biogenesis by changing the chromatin of piRNA clusters and inducing precursor processing. Genes Dev. *28*, 1667–1680.

Li, C., Vagin, V.V., Lee, S., Xu, J., Ma, S., Xi, H., Seitz, H., Horwich, M.D., Syrzycka, M., Honda, B.M., et al. (2009). Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies. Cell *137*, 509–521.

Li, X.Z., Roy, C.K., Dong, X., Bolcun-Filas, E., Wang, J., Han, B.W., Xu, J., Moore, M.J., Schimenti, J.C., Weng, Z., and Zamore, P.D. (2013). An ancient transcription factor initiates the burst of piRNA production during early meiosis in mouse testes. Mol. Cell *50*, 67–81.

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.

Luteijn, M.J., and Ketting, R.F. (2013). PIWI-interacting RNAs: from generation to transgenerational epigenetics. Nat. Rev. Genet. *14*, 523–534.

Malone, C.D., Brennecke, J., Dus, M., Stark, A., McCombie, W.R., Sachidanandam, R., and Hannon, G.J. (2009). Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary. Cell *137*, 522–535.

Nishida, K.M., Iwasaki, Y.W., Murota, Y., Nagao, A., Mannen, T., Kato, Y., Siomi, H., and Siomi, M.C. (2015). Respective functions of two distinct Siwi complexes assembled during PIWI-interacting RNA biogenesis in Bombyx germ cells. Cell Rep. *10*, 193–203.

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.

Robine, N., Lau, N.C., Balla, S., Jin, Z., Okamura, K., Kuramochi-Miyagawa, S., Blower, M.D., and Lai, E.C. (2009). A broadly conserved pathway generates 3'UTR-directed primary piRNAs. Curr. Biol. *19*, 2066–2076.

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.

Vourekas, A., Zheng, K., Fu, Q., Maragkakis, M., Alexiou, P., Ma, J., Pillai, R.S., Mourelatos, Z., and Wang, P.J. (2015). The RNA helicase MOV10L1 binds piRNA precursors to initiate piRNA processing. Genes Dev. *29*, 617–629.

Xiol, J., Cora, E., Koglgruber, R., Chuma, S., Subramanian, S., Hosokawa, M., Reuter, M., Yang, Z., Berninger, P., Palencia, A., et al. (2012). A role for Fkbp6 and the chaperone machinery in piRNA amplification and transposon silencing. Mol. Cell *47*, 970–979.

Xiol, J., Spinelli, P., Laussmann, M.A., Homolka, D., Yang, Z., Cora, E., Couté, Y., Conn, S., Kadlec, J., Sachidanandam, R., et al. (2014). RNA clamping by Vasa assembles a piRNA amplifier complex on transposon transcripts. Cell *157*, 1698–1711.

Zheng, K., and Wang, P.J. (2012). Blockade of pachytene piRNA biogenesis reveals a novel requirement for maintaining post-meiotic germline genome integrity. PLoS Genet. *8*, e1003038.