Doctoral Dissertation

博士論文

Functional analysis of the RNA helicase, Armitage, in piRNA biogenesis

（piRNA 生合成に寄与する RNA ヘリカーゼ Armitage の機能解析）

A Dissertation Submitted for the Degree of Doctor of Philosophy

December 2019

令和元年１２月博士（理学）申請

Department of Biologics Sciences, Graduate School of Science,

The University of Tokyo

東京大学大学院理学系研究科生物科学専攻

Tatsuki Kinoshita

木下　達貴

Contents

Abstract

PIWI-interacting RNAs (piRNAs) repress transposons to maintain germline genome integrity. Previous study showed that Armitage (Armi) is co-immunoprecipitated with female sterile (1) Yb (Yb) and localized to Yb body, which is considered to be the site for piRNA biogenesis, in a Yb-dependent manner. Other previous studies also showed that artificial tethering of Armi to reporter RNAs induced piRNA biogenesis. However, Yb depletion in *Drosophila* ovarian somatic cells (OSCs) impaired the production of transposon-targeting piRNAs, even in the presence of Armi. Here, I show that the specific interaction of Armi with piRNA precursors, which include both *flamenco* piRNA cluster transcripts and genic piRNA sources, is strictly regulated by Yb. Yb depletion allowed Armi to bind a wider range of RNAs in accordance with their expression levels. The lack of Yb also lead to the production of piRNAs from bound RNAs, which hardly contribute to transposon silencing. Tethering assay supported that Armi has an intrinsic ability to induce piRNA production from bound RNAs Yb-independently. Yb is required for the selection of piRNA precursor, but not required for piRNA processing. *In vitro* assays showed that the ATP hydrolysis-defective mutants of Armi failed to exhibit 5′-to-3′ directional RNA unwinding activity and also failed to detach from bound RNAs. These mutants abolished Yb localization, resulting in the loss of selective association of Armi with piRNA precursors. Therefore, both Yb and ATP hydrolysis activity of Armi is necessary to reduce promiscuous Armi–RNA binding. I also show that Armi is directly associated with Yb. These findings shed light on distinct and collaborative requirements of Yb and Armi in transposon-targeting piRNA biogenesis.

Abbreviations

Abbreviations

Abbreviation	Full name
Ago3	Argonaute 3
Armi	Armitage
ATP	Adenosine triphosphate
Aub	Aubergine
BoYb	Brother of Yb
bp	Base pairs
BSA	Bovine serum albumin
CBB	Coomassie brilliant blue
CDS	Coding sequence
CLIP	Crosslinking and Immunoprecipitation
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
DTT	Dithiothreitol
EDC	1-Ethyl-3-(3-dimethylaminopropyl) carbodiimide
EDTA	Ethylenediaminetetraacetic acid
EGFP	Enhanced green fluorescent protein
FAST-iCLIP	Fully automated and standardized individual-nucleotide resolution CLIP
flam	Flamenco
HEPES	4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid
IP	Immunoprecipitation
kb	Kilobase pairs
KD	Knockdown
kDa	Kilodalton
lincRNA	Long intergenic non-coding RNA
MES	2-(N-morpholino)ethanesulfonic acid
miRNA	Micro RNA
mRNA	Messenger RNA
n.i.	Non-immune immunoglobulin
NP-40	Nonyl phenoxypolyethoxylethanol
nt	Nucleotide
ORF	Open reading frame
OSC	Ovarian somatic cell

PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphor buffered saline
PBS-T	PBS containing 0.1% Tween20
PCR	Polymerase chain reaction
piRISC	piRNA-induced silencing complex
piRNA	PIWI-interacting RNA
Piwi	P-element induced wimpy testis
RISC	RNA-induced silencing complex
RNA	Ribonucleic acid
RNAi	RNA interference
RNase	Ribonuclease
RNP	Ribonucleoprotein
RPKM	Reads per kilobase of exon per million mapped reads
RPM	Reads per million mapped reads
rRNA	Ribosomal RNA
RT-qPCR	Reverse transcription quantitative PCR
S2 cell	Schneider 2 cell
SDS	Sodium dodecyl sulfate
SEM	standard error of the mean
siRNA	small interfering RNA
snoRNA	Small nucleolar RNA
snRNA	Small nuclear RNA
SoYb	Sister of Yb
ssRNA	Single-stranded RNA
TBE	Tris–borate–EDTA
tj	Traffic jam
Tris	Tris(hydroxymethyl)aminomethane
tRNA	Transfer RNA
UTR	Untranslated region
WT	Wildtype
Yb	Female sterile (1) Yb
Zuc	Zucchini

Introduction

Introduction

RNA silencing and transposons

Gene regulation plays an essential role in maintaining proper gene expression. Gene expression is regulated in various steps, from RNA transcription to post-translational protein modification. One of the regulatory mechanisms conserved in almost all eukaryotes is RNA silencing, in which small RNAs, ~20 to 35 nucleotides, recognize target RNAs by their complementarity, leading to transcriptional or post-transcriptional gene silencing (Czech and Hannon, 2011; Ghildiyal and Zamore, 2009; Kim et al., 2009; Meister, 2013; Siomi et al., 2011; Wilson and Doudna, 2013) (Figures 1A and 1B).

Target regulation by small RNAs is mediated by an effector complex called RNA-induced silencing complex (RISC), of which small RNA and an Argonaute-family protein are component. RISCs regulate the expression of target genes by endonuclease activity of Argonaute proteins or recuiting additional proteins, such as RNA degradation machinery or histone methyltransferase. In animals, Argonaute-family proteins are classified into two groups; AGO subfamily and PIWI subfamily. Small RNAs are also classified into three classes based on their functions and biogenesis mechanisms: micro RNAs (miRNAs), small interfering RNAs (siRNAs) and PIWI-interacting RNAs (piRNAs) (Ghildiyal and Zamore, 2009; Kim et al., 2009) (Figure 2).

miRNAs are generated from endogenous hairpin-structured transcripts. Hairpin-RNAs are cleaved by Dicer protein and resulting dsRNAs are loaded onto AGO subfamily protein. miRNAs target mRNA, usually in the 3′ UTR and binding to the target mRNA

induces translational repression by recruiting mRNA decay machinery (Figure 2). siRNAs are typically processed from exogenous or endogenous long double-stranded RNAs or hairpin-RNAs. siRNA sources are also cleaved by Dicer protein and resulting dsRNAs are loaded onto AGO subfamily proteins. siRNAs contribute to antiviral defense and transposon repression (Figure 2). piRNAs, which are ~24 to 31 nt in length, are longer than siRNAs and miRNAs which are ~22 nt in length. Although siRNAs and miRNAs express ubiquitously, piRNAs specifically express in germline tissues and associate with PIWI subfamily proteins. piRNAs are Dicer-independently generated from long single-stranded RNAs and involved in transposon silencing through transcriptional or post-transcriptional gene silencing (Figure 2).

Transposons, or transposable elements, are DNA sequences that can change their positions within a genome. Although transposons could induce genome evolution, transposon insertions also have a harmful effect on their host, such as loss of function and genome destruction. In somatic cells, transposon insertions would not be inherited to the next generation. However, in germline cells, where genome is inherited to the next generation, transposon expressions must be regulated properly. In *Drosophila* somatic cells, transposons are targeted by siRNAs (Chung et al., 2008; Ghildiyal et al., 2008; Kawamura et al., 2008). On the other hand, in *Drosophila* germline, both siRNAs and piRNAs repress transposons (Czech et al., 2008). Considering that loss of piRNA increased transposon expression significantly but loss of siRNA had little effect (Vagin et al., 2006), piRNAs may play an essential role in transposon repression in the germline.

piRNAs in *Drosophila*

piRNA pathway is the most recently discovered pathway among the three classes of small RNA pathways. piRNAs were originally identified in the analysis of small RNAs during *Drosophila* development and termed as repeat-associated small interfering RNAs (rasiRNAs) because most of these RNAs corresponded to repetitive elements, including retrotransposons (Aravin et al., 2004). Piwi and Aub mutations in *Drosophila* ovaries resulted in a lack of rasiRNA, suggesting the correlation between rasiRNAs and PIWI subfamily proteins (Vagin et al., 2006). Immunoprecipitation of PIWI proteins showed that all PIWI subfamily proteins bound to rasiRNAs (Brennecke et al., 2007; Gunawardane et al., 2007; Nishida et al., 2007; Saito et al., 2006). PIWI proteins in mice, rat and zebrafish also bound to rasiRNA-like small RNAs, which are now referred to as piRNAs (Aravin et al., 2006; Girard et al., 2006; Grivna et al., 2006; Lau et al., 2006; Watanabe et al., 2006). Since the discovery of piRNAs, accumulating studies have uncovered the piRNA pathway and its functions, mainly using fly ovaries and mouse testis (Czech et al., 2018; Huang et al., 2017; Iwasaki et al., 2015; Ozata et al., 2019; Yamashiro and Siomi, 2018) (Figure 3).

In *Drosophila* three PIWI subfamily proteins express: Argonaute 3 (Ago3), Aubergine (Aub) and P-element induced wimpy testis (Piwi). A lack of any PIWI protein in *Drosophila* results in sterility (Cox et al., 1998; Li et al., 2009a; Lin and Spradling, 1997; Schmidt et al., 1999; Schüpbach and Wieschaus, 1991), indicating that all members of PIWI subfamily are functionally non-redundant and indispensable for development and transposon silencing of both ovaries and testes.

Aub and Ago3 collaboratively repress transposons in the cytoplasm post-transcriptionally and determine the 5′ ends of "secondary" piRNAs, using their slicer activities. Cleavage products by Aub-piRISC serve as substrates for Ago3-bound secondary piRNAs. In turn, cleavage products by Ago3-bound piRNAs are also utilized to generate Aub-bound piRNAs similarly. Both Aub and Ago3 cleave target RNAs at the position between 10 and 11 nt of the bound piRNAs. Thus, Aub and Ago3-bound piRNAs show 10-nt complementary overlap at their 5′ ends. Notably, as the 5′ end of Aub-bound piRNAs are rich in uridine (1U bias), Ago3-bound piRNAs have 10A bias. In Aub and Ago3-mediated piRNA production, slicer-dependent transposon repression is coupled to secondary piRNA amplification. This amplification pathway is also referred to as the "ping-pong" cycle (Figure 4).

Piwi repress transposons in the nucleus transcriptionally. Piwi-piRISCs are imported into the nucleus by importin α (Yashiro et al., 2018). Although Piwi lacks slicer activity, Piwi repress transposons by inducing heterochromatin formation, with GTSF1/Asterix, Panoramix/Silencio, Maelstrom, Eggless/SetDB1, and linker histone H1 (Dönertas et al., 2013; Iwasaki et al., 2016; Ohtani et al., 2013; Sienski et al., 2012, 2015; Yu et al., 2015). Piwi-bound piRNAs have 1U bias as well as Aub-bound piRNAs.

In *Drosophila* ovaries, all three PIWI proteins express in the germ cells which are differentiated from germline stem cells. On the other hand, in somatic follicle cells surrounding the germ cells, only express Piwi (Figure 3). Thus, transposons are repressed in the nucleus transcriptionally in follicle cells. As the impairment of piRNA in follicle cells induces the transposition of virus-like particles derived from transposons to nurse cells (Song et al., 1997), piRNAs in follicle cells may be indispensable for

repressing the transposition from follicle cells to the germ cells, even if transposon insertions in follicle cells would not be inherited.

piRNA biogenesis in *Drosophila* ovarian somatic cells

Previously, a cultured cell line which is composed of ovarian somatic cells (OSCs) was established (Saito et al., 2009). In OSCs, transposon-repressive piRNAs are generated mainly from *flamenco* (*flam*), which is a soma-specifically expressed, uni-strand piRNA cluster which are rich in transposon remnants (Brennecke et al., 2007). The *flam* transcripts contain fragments of various transposons, most of which are antisense to the transposon sequences. Therefore, piRNAs derived from the *flam* locus are transposon-targeting piRNAs, which contribute to transposon silencing. It is unclear how transposon remnants accumulate in the piRNA clusters. Because transposons can be integrated into various genomic loci other than piRNA clusters (Wang et al., 2018), transposon accumulation in the piRNA clusters might be the final outcome of selective pressure through evolution.

A subset of mRNAs such as *traffic jam* (*tj*) also serve as piRNA sources and these "genic piRNAs" are particularly produced from their 3′ UTRs (Robine et al., 2009; Saito et al., 2009). Genic piRNAs show little complementarity to transposons, so their functions still remain elusive.

The piRNA processing in OSCs requires numerous proteins including female sterile (1) Yb (Yb), Armitage (Armi), Vreteno (Vret), Sister of Yb (SoYb), Shutdown (Shu), Zucchini (Zuc), germ cell protein with ankyrin repeats, sterile alpha motif, and leucine zipper (gasz), and Minotaur (Mino) (Haase et al., 2010; Handler et al., 2013; Hirakata et

al., 2019; Munafò et al., 2019; Olivieri et al., 2010; Preall et al., 2012; Qi et al., 2011; Saito et al., 2010; Szakmary et al., 2009; Vagin et al., 2013; Yamashiro et al., 2019; Zamparini et al., 2011) (Figure 5). Yb, Armi, SoYb and Vret localize to Yb bodies, which are perinuclear non-membranous organelles formed by liquid-liquid phase separation (LLPS) (Hirakata et al., 2019), while Zuc, Gasz, Daed and Mino are located on the surface of mitochondria through their own transmembrane signals. Therefore, both Yb bodies and mitochondria are supposed to be the places where piRNA processing occurs and inter-organelle translocation of piRNA precursors should be required. Supporting this, Yb bodies tend to be surrounded by mitochondria (Murota et al., 2014; Szakmary et al., 2009). This spatial arrangement of the two organelles might increase the efficiency of piRNA biogenesis, by gathering all piRNA factors and piRNA precursors in close proximity.

Armitage and Yb are required for piRNA biogenesis in OSCs

Yb is a Tudor domain-containing protein that possesses an RNA helicase domain (Handler et al., 2011; Szakmary et al., 2009). A single mutation introduced into the RNA helicase domain abolished the RNA-binding activity of Yb and impaired Yb body formation (Murota et al., 2014). As a result, piRNAs were hardly produced and transposons were desilenced. Thus, the RNA-binding activity of Yb is necessary for Yb body assembly, piRNA biogenesis, and transposon silencing. *Drosophila* has two Yb orthologs, SoYb and Brother of Yb (BoYb). OSCs express Yb and SoYb but not BoYb, whereas ovarian germ cells express SoYb and BoYb, but not Yb. Mice have a homolog of SoYb (*i.e.,* Tdrd12), but no counterparts of Yb and BoYb (Pandey et al., 2013).

Previous crosslinking and immunoprecipitation (CLIP) experiments in OSCs revealed that Yb specifically associated with piRNA precursors (Ishizu et al., 2015; Murota et al., 2014). Analysis of piRNA precursors based on CLIP signals revealed that Yb induce piRNA processing from the downstream regions and also specify the regions from which piRNAs are produced. The production of artificial piRNAs from reporter transcripts is triggered by Yb–RNA (*i.e., trans-cis*) association, and reporter-derived artificial piRNAs repressed complementary genes (Ishizu et al., 2015). The first exon of the *flam* transcripts exhibited the ability to trigger artificial piRNA production (Homolka et al., 2015). Within the first exon, approximately 200 nt fragment was identified as *flam-cis*-element (Ishizu et al., 2015), which is necessary and sufficient for piRNA production from the downstream region. *tj-cis*-element was also identified within the 3′ UTR of the *tj* transcripts (Ishizu et al., 2015).

Armi contains a Upf1-like superfamily1 (SF1) RNA helicase domain and the domain conserves sequential motifs characteristic among the members of the SF1 helicase family (Cook et al., 2004). Armi co-purified with Yb and Piwi, and the complex contained piRNA precursors (Saito et al., 2010). As Piwi in the complex was devoid of piRNAs, the RNA–protein (RNP) complex is formed during piRNA biogenesis (Saito et al., 2010). Although Armi was dispersed in the cytosol upon Yb depletion, Yb bodies were still assembled in Armi-depleted OSCs. These results suggested that Yb control Armi's behavior and function in piRNA processing.

Armi exhibits 5′-to-3′ directional RNA unwinding activity *in vitro* (Pandey et al., 2017). When Armi is artificially tethered to reporter transcripts in cultured OSCs through the affinity of λN and BoxB, phased piRNAs production was induced (Pandey et al., 2017). These findings support the idea that unwinding of piRNA precursors by Armi

contribute to Zuc-dependent phased piRNA biogenesis. However, the requirement of Yb in this processing is unknown because the tethering assays were performed where endogenous Yb expresses. MOV10L1 in mice, a homolog of Armi, also unwinds RNA duplexes from 5′ to 3′ *in vitro* and a mutant defective of this activity failed to generate piRNAs in the testes (Vourekas et al., 2015). MOV10L1 also specifically associate with piRNA precursors that often have secondary structures such as G quadruplexes. A model was proposed that MOV10L1 translocates and remodels the secondary structures of piRNA precursors and facilitate cleavage by nuclease(s), promoting piRNA biogenesis (Vourekas et al., 2015). However, how piRNA precursors are discriminated from other transcripts in the mouse testes remains unknown.

piRNAs are produced in a phased manner

Genome-wide sequencing of Piwi-bound piRNAs in cultured OSCs and *Drosophila* ovaries led to the proposal that piRNA production occurs in a phased manner, where endonuclease Zuc (Haase et al., 2010; Ipsaro et al., 2012; Nishimasu et al., 2012; Olivieri et al., 2010; Saito et al., 2010) excises mature piRNAs successively from the 5′ end of piRNA precursors (Han et al., 2015; Homolka et al., 2015; Mohn et al., 2015) (Figures 4 and 5). During the processing, Piwi binds to a piRNA precursor through its 5′ end through the binding pocket embedded in the MID domain of Piwi. The precursor is cleaved by Zuc and a piRISC is released into the cytosol. Following this reaction, the cleavage product exposes a new 5′ end to which another piRNA-free Piwi binds. Repeating these reactions, piRISCs are successively released from single piRNA precursor. Therefore, Zuc determines both 5′ and 3′ ends of phased piRNAs. However,

even the loss of Zuc accumulates fragmented piRNA precursors (Murota et al., 2014), implying that the very first cleavage of the primary precursors is carried out by nuclease(s) other than Zuc, bearing a monophosphate at 5′ ends for Piwi loading, although such factor(s) have not been identified. The precise mechanism of phased piRNA biogenesis also remains elusive.

RNA helicases in piRNA biogenesis

RNA helicases are highly conserved enzymes that use energy from cycles of ATP binding and hydrolysis to bind or remodel RNAs or RNA–protein complexes (Pyle, 2008; Singleton et al., 2007; Tanner and Linder, 2001). RNA helicases are found in all species (Anantharaman et al., 2002) and involved in nearly all aspects of RNA metabolism (Jankowsky and Fairman, 2007; Tanner and Linder, 2001).

RNA helicases from both eykaryotes and prokaryotes are classified into six superfamilies (SF1 to SF6), according to the motifs in their amino acid sequences (Gorbalenya and Koonin, 1993; Singleton et al., 2007). SF3 to SF6 helicases, which form toroidal, predominantly hexameric structures, are found in bacteria and viruses (Kainov et al., 2006). All eukaryotes RNA helicases belong to SF1 and SF2, which are non-ring forming (Singleton et al., 2007) (Figure 6A). SF1 and SF2 helicases contain structurally conserved helicase core, formed by two tandem RecA-like folds, RecA1 and RecA2. Both SF1 and SF2 helicases contain sequence motifs characteristic of RNA helicases reside within these domains (Figure 6B). Furthermore, C- and/or N-terminal additional domains are typically connected to helicase core, which specify its cellular functions by recruitment of other proteins or recognition of specific RNAs.

RNA helicase activity is classified into two types: "canonical duplex unwinding" and "local strand separation" (Jankowsky, 2011). In canonical duplex unwinding, helicases translocate along the bound strand with defined directionality, either 5′ to 3′ or 3′ to 5′, thereby displacing the complementary strand (Jankowsky, 2011) (Figure 6C). This type of unwinding usually requires a single-stranded overhang region in a defined orientation (polarity) with the duplex. RNA helicases of the Upf1-like families exhibit from 5′ to 3′ unwinding polarities (Czaplinski et al., 1995; De et al., 2015). In contrast, RNA helicases of the DEAD-box family unwind duplexes by local strand separation (Jankowsky, 2011). Although the unwinding mechanism of DEAD-box proteins is not based on translocation, DEAD-box proteins directly load onto the duplex region, destabilize the duplex and accelerate strand separation (Figure 6D) (Bizebard et al., 2004; Tijerina et al., 2006; Yang and Jankowsky, 2006; Yang et al., 2007). Therefore, unwinding occurs without defined polarity (Tijerina et al., 2006; Yang and Jankowsky, 2006).

Many helicases are involved in piRNA pathway. Yb, SoYb, BoYb, Spindle-E (Spn-E), Vasa and Armi have been identified as piRNA-related helicases and all of them except for Armi belong to DEAD-box family, in SF2 superfamily (Iwasaki et al., 2015). Accumulating genetic and biochemical studies have shown that point mutations to conserved motifs in piRNA-related helicases impaired piRNA pathway (Murota et al., 2014; Nishida et al., 2015; Pandey et al., 2013; Xiol et al., 2014). For instance, Vasa is involved in ping-pong cycle by displacing the cleavage products from Siwi-piRISC ATP-dependently (Nishida et al., 2015; Xiol et al., 2014) (Figure 4). However, several mutations introduced to conserved motifs in Vasa abolished the RNA-release activity, resulting in defect in Ago3-piRNA biogenesis.

Summary of this study

Previous studies suggested that Armi is involved in the piRNA biogenesis Yb-dependently. However, the functions of Armi and the relationship with Yb are still unknown. Particularly, although tethering assay suggested the involvement of the helicase activity of Armi in piRNA biogenesis, its functions about helicase activity of Armi have not been sufficiently investigated.

To gain insights into the roles of Armi in somatic piRNA biogenesis and the relationship with Yb, fully automated and standardized individual-nucleotide-resolution CLIP (FAST-iCLIP) (Flynn et al., 2015) were performed for Armi in OSCs in the presense and absence of Yb. CLIP analysis revealed that Armi in OSCs specifically bound with *flam* transcripts and genic piRNA precursors similarly to Yb (Ishizu et al., 2015). Yb depletion greatly reduced the specificity in Armi–piRNA precursor interaction. Armi without Yb bound promiscuously with RNAs largely in accordance with the expression levels of the cellular RNAs. Strikingly, Armi in Yb-depleted OSCs could produce piRNAs from its bound RNAs regardless of its dispersal at cytosol. Those piRNAs were mostly not complementary to transposon sequences, so they failed to repress transposons. The ATP hydrolysis-defective mutants of Armi failed to unwind RNAs *in vitro* and were retained on bound RNAs at cytosol in OSCs, abolishing piRNA production. These results suggest that both Yb and ATP-dependent, RNA unwinding/dissociating activities of Armi are necessary for generating transposon-targeting piRNAs.

Materials and Methods

Materials and Methods

Cell culture

OSCs were cultured at 25°C in Shields and Sang M3 Insect Medium (Sigma) supplemented with 10% fly extract, 10% fetal bovine serum (Atlas Biologicals), 0.6 mg/ml glutathione, and 10 mU/ml insulin, as described previously (Niki et al., 2006; Saito et al., 2009). S2 cells were cultured at 25°C in Schneider's *Drosophila* Medium (Gibco) supplemented with 10% fetal bovine serum (Equitech-Bio) and 1× penicillin–streptomycin–glutamine (Gibco).

FAST-iCLIP

FAST-iCLIP was performed as previously described (Flynn et al., 2015) with some modifications. For each iCLIP experiment, a total of 1×10^8 OSCs were washed once with ice-cold PBS and UV-crosslinked by irradiating with 200 mJ/cm^2 of 254 nm UV. Cells were harvested and cell pellets were lysed with lysis buffer [20 mM HEPES-KOH (pH7.3), 150 mM NaCl, 1 mM EDTA, 1 mM DTT, 0.5% NP-40, 2 μg/ml pepstatin, 2 μg/ml leupeptin, 0.5% aprotinin]. For immunoprecipitation of RNA–protein complexes, 10 μg of mouse monoclonal antibodies [anti-Armi and anti-c-Myc (Sigma)] were immobilized on 100 μl of SureBeads Protein G magnetic beads (Bio-Rad). The cleared lysates are incubated with beads for 2h at 4°C, after which beads were washed three times with wash buffer [20 mM HEPES-KOH, (pH7.3), 300 mM NaCl, 1 mM DTT, 0.05% NP-40, 2 μg/ml pepstatin, 2 μg/ml leupeptin, 0.5% aprotinin], following 1 U/μl RNase T1 (Thermo Fisher) treatment for 15 min at room temperature. Beads were washed three times with high-salt wash buffer [20 mM HEPES-KOH (pH7.3), 500 mM NaCl, 1 mM DTT, 0.05% NP-40, 2 μg/ml

21

pepstatin, 2 μg/ml leupeptin, 0.5% aprotinin]. After 3′ end RNA dephosphorylation with

T4 PNK (NEB), 3′ ends of the RNAs were ligated to a preadenylated DNA adapter with

T4 RNA ligase2, truncated KQ (NEB), and then radioactively labeled. The samples were

resolved by NuPAGE gel electrophoresis (Life Technologies) and RNA–protein

complexes were transferred to nitrocellulose membrane. RNA–protein complexes were

detected by phosphorimaging and cutting out from the membrane. For RNA isolation,

the membrane was treated with Proteinase K (Roche). RNAs were reverse transcribed

and resulting cDNAs were purified, following circularization of cDNAs as previously

described (Flynn et al., 2015). Libraries were then amplified with Q5 DNA polymerase

(NEB) and size-selected via polyacrylamide gel electrophoresis. iCLIP libraries were

sequenced using the HiSeq 2000 or 2500 platform as single-end 50-bp reads.

Immunoprecipitation of Piwi

For piRNA-seq library preparation, small RNA northern blotting and visuallizaiton of Piwi-

associated RNAs, immunoprecipitation of Piwi was performed using $1×10^8$ OSCs, which

were lysed in lysis buffer [20 mM HEPES-KOH (pH7.3), 150 mM NaCl, 1 mM EDTA, 1

mM DTT, 0.1% NP-40, 40 U/ml RNasin (Promega), 2 μg/ml pepstatin, 2 μg/ml leupeptin,

0.5% aprotinin]. The lysate was cleared by centrifugation at 20,000*g* for 15 min at 4°C.

The supernatant was collected and cooled on ice. For immunoprecipitation, Anti-Piwi

antibody (Saito et al., 2006), anti-Myc antibody (DSHB), or anti-Flag (MBL) antibody was

immobilized on Dynabeads protein G (Invitrogen). Beads were incubated with lysates for

2 h at 4°C, and then washed five times with high-salt wash buffer [20 mM HEPES-KOH

(pH7.3), 500 mM NaCl, 1 mM EDTA, 1 mM DTT, 0.1% NP-40, 2 μg/ml pepstatin, 2 μg/ml

leupeptin, 0.5% aprotinin]. For sample preparations, RNAs were isolated from the beads

by acid phenol–chloroform (Ambion) treatment and ethanol precipitated. To visualize the RNAs, 1% of them were dephosphorylated with Antarctic Phosphatase (NEB) and 5′-end-labeled with [γ-^{32}P]ATP using T4 PNK. The labeled RNAs were denatured with Gel Loading Buffer II (Ambion) for 5 min at 95°C, and resolved by electrophoresis on a 12% urea-PAGE gel. The RNAs were visualized by phosphorimaging.

Small RNA libraries

Small RNA cloning was performed as previously described (Iwasaki et al., 2017). In summary, co-immunoprecipitated piRNAs were gel-purified and cloned using the NEBNext Small RNA Library Sample Prep Set (NEB). The libraries were analyzed on a Miseq platform (Illumina).

OSC transfection

Transfection of siRNAs was performed as previously described (Iwasaki et al., 2016). In brief, 200 pmol siRNA duplex and $5×10^6$ cells were suspended in 20 µl of SF Cell Line 96-well Nucleofector Solution (Lonza Bioscience) and then nucleofected using the program DG-150 of the 96-well Shuttle Device (Lonza), following the manufacturer's protocol. After 48 h of incubation, a second transfection was performed as before and cells were harvested after 48 h of incubation (total of 96 h). Cotransfection of siRNA and plasmid DNA was performed as previously described (Hirakata et al., 2019). In brief, 200 pmol siRNA duplex and 4 µg of plasmid DNA and $1.5×10^7$ cells were suspended in 100 µl of the transfection buffer [180 mM Sodium Dihydrogen Phosphate (pH 7.2), 5 mM KCl, 15 mM MgCl$_2$ and 50 mM D-mannitol] and nucleofected using the program N-020 of the Nucleofector II Device (Lonza), following the manufacturer's protocol. Cells were

harvested after 48 h of incubation.

RT-qPCR

Total RNAs were isolated from OSCs using Isogen II (Nippon Gene) following the manufacturer's instructions. A total of 1 μg RNA was reverse-transcribed using ReverTra Ace qPCR RT Master Mix (Toyobo) following the manufacturer's instructions. One-tenth of the resulting cDNAs were amplified with PowerUp SYBR Green Master Mix (Thermo Fisher) and primers (0.5 μM final concentration) using StepOnePlus (Thermo Fisher). The primers used are shown in Table 3.

Plasmid construction

An expression vector for Myc-Armi WT was generated by inserting the wild-type Armi coding region into the pAcM vector (Saito et al., 2009). First-strand cDNA was synthesized from total RNA isolated from OSCs using Transcriptor First Strand cDNA Synthesis Kit (Roche). Full-length Armi was amplified by KOD plus DNA polymerase (Toyobo) and cloned into the EcoRI-NotI-digested pAcM vector. The wild-type Armi expression plasmid was mutagenized by inverse PCR using PrimeSTAR Max DNA Polymerase (Takara) following the manufacturer's protocol to generate an siRNA-resistant Armi WT plasmid that contained seven silent mutations introduced into the 21-bp Armi-siRNA target sequence. The point mutation was introduced into the siRNA-resistant Armi WT plasmid using PrimeSTAR Max DNA Polymerase. Expression vectors for Flag-Armi WT and mutants were generated from pAcM-Armi vectors by substituting the Myc-tag coding region with the 3×Flag coding region using NEBuilder HiFi DNA assembly (NEB), in accordance with the manufacturer's instructions. 3×Flag coding

24

region was amplified by PrimeSTAR Max DNA polymerase with the primers Flag F/R using pIB-3×FLAG-Siwi (Nishida et al., 2015) as a template and then cloned into the PCR-amplified linearized vector generated by PrimeSTAR Max DNA polymerase with the primers Flag vector F and pAc vector R. Expression vectors for λN fused protein were generated by inserting the λN coding region into the pAcM-Armi or -Yb vector using NEBuilder HiFi DNA assembly, in accordance with the manufacturer's instructions. λN coding region was amplified by PrimeSTAR Max DNA polymerase with the primers lambda F/R using pAC5.1-λN-HA-LacZ (Fukaya and Tomari, 2011) as a template and then cloned into the PCR-amplified linearized pAcM-Armi or Yb with the primers lambdaN vector F and pAc vector R. pAcM-Yb was as described previously (Murota et al., 2014). pAcM-LacZ was as described previously (Hirakata et al., 2019). For pAc-EGFP-4BoxB-10repeat plasmid production, BoxB cloning oligonucleotides were designed to form double strands with SalI for 5′-end and XhoI for 3′-end overhangs. These pairs of oligonucleotides were annealed, 5′ phosphorylated, and ligated to each other to form self-assembled double-stranded DNA, followed by digestion with XhoI to remove undesired products. Self-assembled DNAs were separated in 6% polyacrylamide gel, and cut out with a razor blade, followed by elution in 0.4 M NaCl for 12 h at 4°C. The products were purified by ethanol precipitation, and ligated into the XhoI site of the pAcE-mcs-10repeat plasmid (Ishizu et al., 2015). Plasmids were sequenced with the pAcE sequencing primer using an ABI Prism BigDye Terminator v3.1 cycle sequencer, in accordance with the manufacturer's protocol. The primers used are listed in Table 3.

Small RNA northern blotting

For total RNA sample preparation, total RNA was isolated from OSCs using Isogen II following the manufacturer's instructions. To remove DNA contamination, the RNA pellets were treated for 10 min at 37°C with 1 μl (2U) of TURBO DNase (Thermo Fisher) in 50 μl of 1×TURBO DNase buffer. RNA was then purified by phenol–chloroform extraction and ethanol precipitation. For small RNA detection, northern blotting analysis was carried out as described previously (Pall and Hamilton, 2008). In brief, 5 μg of total RNAs or Piwi immunoprecipitates (described above) were denatured with Gel Loading Buffer II for 5 min at 95°C, resolved by electrophoresis on a 12% urea-PAGE gel in TBE buffer for 1–1.5 h at 200 V, and transferred onto a Hybond N membrane (GE Healthcare) using a semi-dry electroblotter at 20 V for 60 min. RNAs were crosslinked with EDC crosslinking solution [0.16 M 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide in 0.13 M 1-methylimidazole at pH 8] at 60°C for 2 h. The hybridizations were performed at 42°C overnight in 0.2 M sodium phosphate (pH 7.2), 7% SDS, and 1 mM EDTA with end-labeled antisense DNA oligonucleotide, and washed at 42°C in 2× saline sodium citrate containing 0.1% SDS. The DNA oligonucleotides used are summarized in Table 3.

mRNA northern blotting

Total RNA was isolated from OSCs using Isogen II following the manufacturer's instructions. mRNAs were extracted from 50 μg of total RNA using Oligotex-dT30 (Takara) following the manufacturer's instructions. Purified mRNAs were eluted and denatured with NorthernMax Formaldehyde Load Dye (Invitrogen) for 15 min at 65°C, resolved by electrophoresis on 1.2% agarose gel containing 2.2 M formaldehyde, and passively transferred onto a Hybond N$^+$ membrane (GE Healthcare) with 50 mM NaOH.

RNAs were crosslinked to the membrane by irradiating 240 mJ/cm^2 of 254 nm UV. The hybridizations were performed at 65°C overnight in 0.5 M Sodium Dihydrogen Phosphate (pH 7.2), 7% SDS, and 1 mM EDTA with body-labeled DNA probe, and washed at 65°C in 2× saline sodium citrate containing 0.1% SDS. DNA probes were synthesized using a Random Primer DNA Labeling Kit Ver. 2 (Takara) following the manufacturer's instructions. Templates of randomly primed probes were generated as previously described (Ishizu et al., 2015). Briefly, for the EGFP probe, the EGFP ORF was amplified using the primers EGFP F/R with pAc-EGFP-4BoxB-10repeat as a template. For the GAPDH probe, the GAPDH ORF was amplified using the primers GAPDH F/R with OSC cDNA as a template. The primers and DNA oligonucleotides used are listed in Table 3.

Co-immunoprecipitation

For co-IP of Yb complex, OSCs transfected with pAcM-Armi plasmids were used. A total of 1×10^8 cells were washed in PBS and lysed in lysis buffer [20 mM HEPES-KOH, pH7.3, 150 mM NaCl, 2 mM MgCl$_2$, 1 mM DTT, 0.5% NP-40, 40 U/ml RNasin (Promega), 2 μg/ml pepstatin, 2 μg/ml leupeptin, 0.5% aprotinin]. The lysate was cleared by centrifugation at 20,000g for 15 min at 4°C. The supernatant was collected and cooled on ice. Anti-Yb antibody (Murota et al., 2014) and mouse non-immune IgG (IBL) as a negative control were immobilized on Dynabeads protein G. Beads were incubated with lysates for 2 h at 4°C, and then beads were washed five times with lysis buffer. Washed beads were resuspended in 10 μl of sample buffer containing 100 mM DTT and incubated for 10 min at room temperature. Supernatant containing immunoprecipitates was transferred to a new tube and incubated at 95°C for 5 min. Co-immunoprecipitates were analyzed by western blotting.

Western blotting

For input sample preparation, cells were harvested and pelleted by centrifugation at 440g for 5 min. They were then lysed in 2× loading dye [100 mM Tris-HCl (pH 6.8), 4% (w/v) SDS, 0.2% (w/v) bromophenol blue, 20% (v/v) glycerol, and 200 mM DTT] and heated to 95°C for 5 min. The samples were resolved in SDS-PAGE gels and transferred to nitrocellulose membrane (GE Healthcare) using a semi-dry electroblotter by applying 120 mA for 60 min in transfer buffer [25 mM Tris-HCl (pH 7.5), 192 mM glycine, 20% methanol]. After blocking for 30 min with 5% skim milk in PBS containing 0.1% Tween 20 (PBS-T), the membrane was incubated with primary antibodies at room temperature for 1 h. Membranes were washed four times for 5 min each in PBS-T, and then incubated with HRP-conjugated secondary antibody (Cappel) at 1:5,000 in PBS-T for 30 min at room temperature. The membranes were washed four times for 5 min each with PBS-T, and ECL substrate (Bio-Rad) was then applied to the membrane for imaging. The ECL signals were visualized using ImageQuant LAS 4000 (GE Healthcare). Anti-Armi (Saito et al., 2010) (1:1,000 dilution), anti-Piwi (Saito et al., 2006) (1:1,000 dilution), anti-Yb (Murota et al., 2014) (1:1,000 dilution), anti-Myc (9E10, DSHB) (1:1,000 dilution), anti-Flag (MBL, PM020) (1:1,000 dilution) and anti-tubulin (DSHB) (1:2,000 dilution) antibodies were used as primary antibodies.

Recombinant Armi purification and characterization

Armi purification

A total of 1×10^8 S2 cells were transfected with 50 µg of DNA and 100 µl of Lipofectamine 2000 (Invitrogen) and incubated at 25°C. 48 h after transfection, transfected cells were harvested and cell pellets were lysed in 5 ml of lysis buffer [50 mM Tris-HCl (pH8.0), 150

mM NaCl, 1 mM DTT, 0.1% NP-40, 1× EDTA-free protease inhibitor cocktail (Roche)].

To remove RNA contamination, RNase A (Roche) was added to a final concentration of

25 µg/ml and lysed cells were incubated for 15 min at room temperature. Lysates were

cleared by centrifugation at 16,500g for 15 min at 4°C. 5 ml lysates were incubated with

a 500-µl bed volume of Anti-Flag M2 Affinity Gel (Sigma) for 2 h at 4°C. Beads were

washed three times with five volumes of wash buffer [50 mM Tris-HCl (pH8.0), 1 M NaCl,

1 mM DTT, 0.1% NP-40] and then washed three times with five volumes of elution buffer

[50 mM Tris-HCl (pH8.0), 150 mM NaCl, 1 mM DTT, 0.1% NP-40, 10% glycerol]. Proteins

were eluted by overnight incubation in 750 µl of elution buffer containing 0.5 µg/µl 3×Flag

peptide (Sigma) at 4°C. Only supernatant was collected by centrifugation in Ultrafree-

MC GV 0.22 µm Centrifugal Filter Units (Merck), at 12,000g for 5 min. The supernatant

was concentrated approximately fivefold with Amicon Ultra with a membrane NMWL of

30 kDa (Merck). The concentration of the purified proteins was determined using Pierce

660 nm Protein Assay Reagent (Thermo Fisher) based on a BSA standard.

Helicase assay

Helicase assays were carried out as previously described with some modifications

(Jankowsky and Putnam, 2009). Custom RNA oligonucleotides were purchased from

Sigma. All strands were purified by urea-PAGE. Short strands were radioactively 5′-end-

labeled with [γ-^{32}P]ATP using T4 PNK. For annealing, 2 µM of radiolabeled short strands

and 5 µM unlabeled long strands were mixed in 20 mM Tris-HCl (pH 7.5), 50 mM NaCl

and 1 mM EDTA and then heated to 95°C for 5 min and cooled down slowly to 26°C over

1 h. Resulting mixtures were resolved by nondenaturing PAGE and RNA duplexes were

detected by phosphorimaging. Gel slices containing RNA duplexes were excised and

eluted by overnight incubation in 50 mM MES-KOH (pH 6.5), 0.3 M sodium acetate, and

1 mM EDTA at 4°C. RNA duplexes were ethanol-precipitated, dissolved in 50 mM MES-KOH (pH 6.5) and 50 mM KCl and stored at –80°C. The concentration of the duplexes was determined by scintillation counting. Helicase assays were performed in 30-μl reaction mixtures containing 50 mM Tris-HCl (pH 7.5), 50 mM KOAc, 1 mM DTT, 0.01% NP-40, 1 U/μL RNasin, 50 nM Flag-purified Armi and 1 nM RNA duplexes. Reaction mixtures were pre-incubated for 5 min at 25°C. Reactions were initiated by adding a mixture of $MgCl_2$/ATP (0.5 mM final concentration) and unlabeled short strands (500 nM final concentration) to prevent re-annealing of the released strands. Aliquots of 3 μl were taken at the time points indicated and mixed with 3 μl of quench buffer (50 mM EDTA, 0.5% SDS, 0.04% bromophenol blue, 0.04% xylene cyanol, 20% glycerol) The reaction products were separated by 15% nondenaturing PAGE. The labeled strands were visualized by phosphorimaging. The RNA substrates used are summarized in Table 3.

ATPase assay

ATPase assays were performed in 20-μl reaction mixtures containing 50 mM MES-KOH (pH 6.5), 50 mM KOAc, 1 mM $Mg(OAc)_2$, 1 mM DTT, 0.5 mg/ml BSA, 1 U/μl RNasin, 500 nM 16-nt ssRNA (short RNA also used for helicase assay), and 50 nM Flag-purified Armi. Reactions were initiated by adding a mixture of cold ATP and traces of [γ-^{32}P]ATP at a final concentration of 100 μM. Aliquots of 3 μl were taken at the time points of 10, 20, and 30 min. Reactions were stopped by dilution in 750 μl of ice-cold 5% (w/v) charcoal (Nacalai Tesque) in 50 mM NaH_2PO_4 with a vortex and incubation on ice. After centrifugation at 21,500g for 10 min at 4°C, 500 μl of the supernatant was collected and the amount of released PO_4 was determined by counting the radioactivity of the supernatant in a scintillation counter.

Gel shift assay

For the gel shift assay, 16-nt ssRNA (short RNA also used for helicase assay) was purified by urea-PAGE and radioactively 5′-end-labeled with [γ-^{32}P]ATP using T4 PNK. 10-µl reaction mixtures containing 50 mM Tris-HCl (pH 7.5), 50 mM KOAc, 2 mM Mg(OAc)$_2$, 1 mM DTT, 1 U/µl RNasin, 1 nM ssRNA, and 1 µM Flag-purified Armi were incubated for 15 min at 25°C. After incubation, ATP was added to a final concentration of 100 µM. 2 µl of nondenaturing PAGE sample buffer (0.1% bromophenol blue, 0.1% xylene cyanol, and 50% glycerol) was then added. The RNA–protein complexes were separated in 5% nondenaturing Tris-glycine gel. The gels were dried and the RNA–protein complexes were visualized by phosphorimaging.

ATP binding assay

Flag-tagged Armi was expressed in S2 cells as described above. S2 cells were harvested and lysed in binding buffer [30 mM HEPES-KOH (pH 7.3), 150 mM NaCl, 5 mM Mg(OAc)$_2$, 1 mM DTT, 0.1 % NP-40, 2 µg/ml leupeptin, 2 µg/ml pepstatin and 0.5 % aprotinin]. The lysate was cleared by centrifugation at 16,500g for 15 min at 4°C. 14 µg of mouse monoclonal anti-Flag antibody (MBL) was immobilized on 70 µl (2.1 mg) of Dynabeads protein G. The beads were incubated with lysates for 2 h at 4°C, after which they were washed three times with wash buffer [30 mM HEPES-KOH (pH 7.3), 500 mM NaCl, 5 mM Mg(OAc)$_2$, 1 mM DTT, 0.1% NP-40, 2 µg/ml leupeptin, 2 µg/ml pepstatin and 0.5% aprotinin]. A total of 0.3 mg of the beads were used in order to determine the immunoprecipitation efficiency. The rest of the beads were separated into 0.3 mg, 0.6 mg, and 0.9 mg. 40 µl of binding buffer containing 1 mM cold ATP and traces of [γ-^{32}P]ATP was added to the beads. The mixture was then incubated for 1 h at 4°C. After incubation, the beads were washed four times with binding buffer. The amount of bound

ATP was determined by counting the radioactivity of the beads in a scintillation counter.

In vitro binding assay

For *in vitro* binding assay, total of 5×10^7 OSCs transfected with pAcM-Lacz or pAcM-Yb plasmids were used. Cells were harvested and lysed in lysis buffer [20 mM HEPES-KOH, pH7.3, 150 mM KOAc, 5 mM Mg(OAc)$_2$, 5 mM DTT, 0.1% NP-40, 2 µg/ml pepstatin, 2 µg/ml leupeptin, 0.5% aprotinin]. To minimize RNA contamination, RNase A (Roche) was added to a final concentration of 100 µg/ml and lysed cells were incubated for 5 min at room temperature. Lysates were cleared by centrifugation at 16,500g for 15 min at 4°C. The supernatant was collected and cooled on ice. For immunoprecipitation, 4 µl of Anti-Myc antibody (9E10, DSHB) was immobilized on 20 µl of Dynabeads protein G. Beads and lysates were incubated for 2 h at 4°C. Beads were washed five times with high salt wash buffer [20 mM HEPES-KOH, pH7.3, 500 mM NaCl, 1 mM DTT, 0.1% NP-40, 2 µg/ml pepstatin, 2 µg/ml leupeptin, 0.5% aprotinin] and then washed two times with binding buffer [20 mM HEPES-KOH, pH7.3, 150 mM NaCl, 1 mM DTT, 0.1% NP-40, 2 µg/ml pepstatin, 2 µg/ml leupeptin, 0.5% aprotinin]. 100 µl of binding buffer containing 10 nM of Flag-Armi and 300 nM of BSA was added to beads and incubated for 2 h at 4°C. Beads were washed eight times with binding buffer and beads were resuspended in 10 µl of sample buffer containing 100 mM DTT and incubated for 10 min at room temperature. Supernatant containing immunoprecipitates was transferred to a new tube and incubated at 95°C for 5 min. Co-immunoprecipitates were analyzed by western blotting.

Immunofluorescence

For immunofluorescence staining, OSCs were fixed in 4% formaldehyde in PBS. Fixed cells were then rinsed twice with PBS and permeabilized with incubation in PBS containing 0.1% Triton X-100 at room temperature for 15 min. Cells were then incubated with primary antibodies in a blocking solution (3% BSA in PBS) at room temperature for 30 min. Samples were washed three times for 10 min each with PBS, and then incubated with fluorescence-conjugated secondary antibodies in a blocking solution at room temperature for 30 min. Samples were washed three times for 10 min each with PBS, and then mounted using VECTASHIELD Mounting Medium with DAPI (Vector Laboratories). Images were captured using a ZEISS LSM 710. For primary antibodies, anti-Yb IgG2b (Murota et al., 2014) (1:300 dilution), and anti-Myc IgG1 (9E10, DSHB) (1:500 dilution) were used. For secondary antibodies, Alexa Fluor 488-conjugated anti-mouse IgG2b (Molecular Probes) and Alexa Fluor 594-conjugated anti-mouse IgG1 (Molecular Probes) were used (1:1,000 dilution).

Data analysis

Read mapping of iCLIP data

First, 3′ adapter sequences were trimmed using Cutadapt v1.14 and reads less than 30 bp were discarded. Next, PCR-derived duplicates were removed by collapsing all identical reads containing the same random barcode at the 5′ end with fastx-collapser from the FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/index.html). Finally, the 5′-end random barcode located at the positions 1–9 of the reads were trimmed. Sequence tags were then mapped against the Release 6 assembly of the *Drosophila melanogaster* genome (dm6) with STAR (2.5.3a). Alignment output BAM files were

sorted and indexed by SAMtools (Li et al., 2009b). The first nucleotide in the genome upstream of a mapped read sequences was defined as a crosslinked nucleotide. The annotation of reads was determined as previously described using custom scripts (Saito et al., 2009). Reads were assigned to genome regions in the following order: transposon, repeat, rRNA, tRNA, snRNA, snoRNA, pseudogene, lincRNA, pre-miRNA, 3′ UTR, CDS, 5′ UTR, and intergenic. Only uniquely mapped reads were visualized on the UCSC Genome Browser; all tracks were normalized to RPM (reads per million) against the total number of reads in that dataset.

Read mapping of piRNA data

Adapter sequences were removed from the obtained reads and mapped to the Release 6 assembly of the *Drosophila* genome using Bowtie (Langmead et al., 2009), allowing two mismatches and extracting the reads mapped uniquely to the genome. The reads were normalized to RPM by the number of *Drosophila* genome-mapped reads. For mapping to transposable elements, the reads were mapped to RepBase 17.7 using Bowtie, allowing two mismatches.

Read count calculation

The mapped reads from FAST-iCLIP and piRNA-seq data for transcripts annotated in Flybase release 6.13 were counted. The counts were normalized using RPM. Expression data of transcripts (RPKM) was obtained from a previous publication (Ohtani et al., 2013).

Results

Results

Armi selectively binds piRNA precursors similarly to Yb in OSCs

I with help from Dr. Hirotsugu Ishizu first performed Armi FAST-iCLIP in normal OSCs (Figures 7 and 8). Two biological replicate libraries, Rep1 and Rep2, were generated, and the sequencing library statistics were summarized in Table 1. A high correlation between these two libraries was observed in the normalized number of reads [as reads per million (RPM) value] (Spearman's rho = 0.94) (Figure 9). Therefore, these sequence tags were combined for further analysis.

The genomic distribution of Armi-FAST-iCLIP nucleotides was highly similar to that of Piwi-loaded piRNAs in OSCs (Figure 10A). Both Armi-bound RNAs and Piwi-loaded piRNAs are mainly derived from antisense transposons and 3′ UTR of mRNAs. For this analysis, Piwi-loaded piRNAs were newly isolated by immunoprecipitation from OSCs and sequenced as in our previous study (Ohtani et al., 2013). Armi-FAST-iCLIP tags were uniquely mapped on the *flam* piRNA cluster and the *tj* gene, of which transcripts were major piRNA sources in OSCs, and significantly overlapped with unique mapping of Piwi-bound piRNAs (Figures 10B and 10C). The Armi-FAST-iCLIP tags were hardly mapped on *tj* CDS and 5′ UTR, which produced few piRNAs (Figure 10C). These results suggest that Armi specifically binds to piRNA precursors and might play an important role in piRNA precursor selection.

For further analysis, the genic piRNA coding genes (Saito et al., 2009) were categorized into three groups: "High," "Intermediate," and "Low," based on their piRNA expression. A total of 258 mRNAs whose RPM was 100 or higher (RPM≥100) were

classified as "High." In addition, 857 and 6,249 mRNAs categorized as "Intermediate" $(20\leq RPM<100)$ and "Low" $(RPM<20)$, respectively.

I with help from Dr. Hirotsugu Ishizu then examined the correlation between the abundances of Piwi-bound genic piRNAs and mRNAs in OSCs. The value of Spearman's rho for the correlation was relatively low (0.69; left panel in Figure 11). This suggests that genic piRNA is not produced in accordance with their expression levels, indicating that genic piRNA precursors is specifically selected. In contrast, the correlation between the abundances of Piwi-loaded piRNAs and Armi-FAST-iCLIP tags was higher (Spearman's rho = 0.86; right panel in Figure 11). These results support the notion that Armi contribute to discriminating genic piRNA precursors among all cellular mRNAs, irrespective of their expression levels *in vivo*.

Yb controls the specificity in Armi–piRNA precursor association

The mapping patterns of Armi-FAST-iCLIP tags on *flam* and *tj* (Figures 10B and 10C) were similar to those of Yb-CLIP tags (Ishizu et al., 2015). Armi co-immunopurified with Yb and localized to Yb bodies in a Yb-dependent manner. Furthermore, Armi co-immunoprecipitates in Yb-depleted OSCs no longer contained *tj*-derived piRNA precursors (Saito et al., 2010), indicating that the selective Armi–piRNA precursor association could be regulated by Yb. I therefore examined how the selective Armi–piRNA precursor association is influenced by the loss of Yb. Armi FAST-iCLIP was conducted once again, but in the presence or absence of Yb in OSCs (Table 1). The loss of Yb hardly affected the level of Armi protein (Figure 12A). The CLIP signal of Armi was also not markedly reduced by Yb depletion (Figures 12B and 12C). Supporting this, our

previous study showed that *tj* piRNA precursors contained in the Armi co-immunoprecipitates were almost completely disappeared when Yb was depleted (Saito et al., 2010). These results suggest that RNA binding activity of Armi is still maintained even in the absence of Yb, although the CLIP signal slightly mapped on *tj* might be the outcomes of UV irradiations which leads RNAs to covalently and irreversibly attach to proteins.

Two biological replicate libraries were generated in the presence and absence of Yb. The reproducibility of the libraries was significantly high (Figure 13); thus, the sequence tags were individually combined for further analysis. Genomic distribution analysis showed that the rate of Armi-FAST-iCLIP tags mapped to transposons was drastically reduced from 48.59% to 19.80% by the loss of Yb (Figure 14A). In contrast, Yb depletion increased Armi binding to other RNAs, except those in the "intergenic" group (Figure 14A). The "intergenic" group includes sequences of *flam,* but not transposon sequences. Armi-FAST-iCLIP tags uniquely mapped to the *flam* locus disappeared almost completely by the loss of Yb (Figure 14B), although the expression level of *flam* transcripts was little changed upon Yb loss (Figure 15). The changes in Armi binding to RNAs in "others" (see the legend for details) and "CDS" were significant (~3.4-fold). The changes in "5′ UTR" were also notable (~2.7-fold), but those in "3′ UTR" were rather small (~1.3-fold). Overall, the lack of Yb drastically change the RNA-binding preference of Armi, resulting in the decrease in Armi binding to piRNA precursors to some extent.

Armi-FAST-iCLIP tags mapped on genes in the "High" group, such as *tj*, apparently decreased by Yb loss (Figures 14C and 16), although the expression level of *tj* RNA transcripts was hardly changed (Figure 15). However, Armi-FAST-iCLIP tags

mapped on genes in the "Intermediate" and "Low" groups increased upon Yb depletion (Figures 14C and 16) The comprehensive analysis also agreed with these trends; the median log2 fold changes were -0.55 (High), 0.90 (Intermediate), and 1.31 (Low) (Figure 17). It appears that Armi without Yb tends to bind mRNAs in a more randomized fashion.

To examine whether Armi without Yb bind to mRNAs in accordance with their expression, the correlations between Armi FAST-iCLIP tags and mRNAs were assessed. The value of Spearman's rho for the correlation in normal OSCs was 0.78 (left panel in Figure 18), similar to that in Figure 11 (left panel). However, Yb depletion increased this value to 0.83 (right panel in Figure 18). All of these findings suggest that Armi without Yb bind to cellular RNAs promiscuously in accordance with their expression levels in OSCs.

Yb is necessary for selective Armi–piRNA precursor binding but not for piRNA processing

In Yb-depleted OSCs, fewer piRNAs were loaded onto Piwi, but residual piRNAs were detected, as reported previously (Saito et al., 2010) (Figure 19). Armi-FAST-iCLIP analysis revealed that Armi exhibited promiscuous RNA binding activity upon Yb depletion. This raised a possibility that piRNAs could be produced from Armi-bound RNAs even after Yb depletion. Therefore, piRNA-seq from control and Yb-depleted OSCs was performed and their sequences were analysed (Table 1). A high correlation was observed between Armi-FAST-iCLIP tags and Piwi-bound piRNAs in control OSCs (left panel in Figure 20A). However, in Yb-depleted OSCs, major changes were observed in "CDS" and "others" (right panel in Figure 20A). The detailed analyses clarified that Armi binding to rRNAs in the group of "others" significantly increased by the loss of Yb

(3.77% to 20.18%) (Table 2), although rRNA-derived piRNA production was only slightly increased from 2.74% to 3.40% upon Yb depletion (Table 2). The substantial changes observed in "CDS" and "others" may be attributed to this cryptic phenomenon. One potential explanation for this is that Armi fails to process rRNAs into mature piRNAs due to their highly structures.

Genic piRNA mapping onto the genome showed that the levels of genic piRNAs derived from "High" group genes were reduced by the loss of Yb, while those from "Intermediate" and "Low" group genes increased moderately (Figure 20B). These tendencies agreed with the comprehensive analysis of the changes in genic piRNA expression levels among three groups (Figure 21); the median log2 fold changes were 0.19 (High), 1.50 (Intermediate), and 1.74 (Low). It should be noted that these changes were basically similar to those observed in Armi-FAST-iCLIP tags, indicating that the proportion of piRNAs is strongly influenced by the specificity of Armi–RNA binding and piRNA is produced in accordance with the RNA binding activity of Armi in OSCs.

The value of Spearman's rho for the correlation between the abundances of Piwi-bound genic piRNAs and mRNAs in normal OSCs was relatively low (Spearman's rho = 0.72; left panel in Figure 22), as in Figure 11 (left panel). However, the value increased upon Yb depletion (Spearman's rho = 0.80; right panel in Figure 22). These results also suggested that piRNAs were produced specifically from Armi-bound RNAs even in Yb-depleted OSCs. Supporting this, the strong correlation between Piwi-bound genic piRNAs and Armi-FAST-iCLIP tags was still observed even in Yb-depleted OSCs (Figure 23). These results suggest that the loss of Yb significantly decreased the specificity of Armi–RNA association, and piRNA production is dependent on the RNA recognition by Armi. These findings indicate that Yb is required to control Armi and

ensure that it binds to the "correct" piRNA precursors but that it is not necessarily required for the downstream processing steps. The findings also indicate that Armi is capable of producing piRNAs without Yb, although the resulting piRNA were mostly non-transposon-targeting. Thus, Yb is not absolutely necessary for piRNA production, but is necessary for producing transposon-targeting piRNAs.

Previous studies demonstrated that tethering of Armi to reporter transcripts, induces piRNA processing in OSCs and also in the germ cells of the ovaries (Pandey et al., 2017; Rogers et al., 2017). These results suggested that recruiting of RNA transcripts to piRNA processing machinery was sufficient to trigger piRNA production. However, the experiments were conducted in the presence of Yb and the requirement for piRNA processing event remains unknown. Therefore, I performed similar assays in both control and Yb-depleted OSCs, which revealed that reporter-derived artificial piRNA production were induced upon Armi tethering even in Yb-depleted OSCs (Figures 24A, 24B and 24C). This result indicated that Yb was not necessarily involved in piRNA processing event. On the other hand, Yb tethering in Armi-depleted OSCs failed to induce piRNA production from the reporter transcripts (Figures 25A and 25B). These results confirmed that Armi has an intrinsic ability to induce piRNA generation in a Yb-independent manner.

ATP hydrolysis by Armi is necessary for piRNA biogenesis and transposon silencing

In vitro assays have shown that Armi in *Drosophila,* and also in mice (*a.k.a.* MOV10L1), exhibited 5′-to-3′ directional RNA unwinding activity (Pandey et al., 2017; Vourekas et al., 2015). In contrast to wild-type (WT) Armi, Armi mutants lacking this

activity failed to produce piRNAs from the reporter genes (Pandey et al., 2017), indicating that the RNA unwinding activity of Armi plays an essential role in piRNA biogenesis. In this study, I mutated two residues, Asp862 and Arg949 in motifs II and IIIa of the RecA1/RecA2 domain, respectively, to alanine and conducted *in vitro* assays to further investigate the characteristic of Armi, using the recombinant proteins (Figures 26 and 27). Although the D862A mutant slightly bound ATP, it failed to hydrolyze ATP (Figures 28 and 29). The Armi R949A mutant exhibited neither ATP binding nor hydrolysis activity (Figures 28 and 29). WT Armi exhibited the directional unwinding activity in the presence of, but not in the absence of, ATP (Figures 30A and 30B). However, neither of the mutants unwound double-stranded (ds) RNA in the presence of ATP (Figure 30A), indicating that the RNA unwinding activity was coupled to the ATP hydrolysis activity. Gel shift assay suggested that the ATP hydrolysis activity was also required for Armi to be released from bound RNAs (Figure 31). Indeed, the rate of the ATP hydrolysis activity increases approximately twice when RNA was added (Figure 32), supporting the notion that RNA release activity of Armi is coupled to the ATP hydrolysis activity. To examine whether ATP hydrolysis activity of Armi is required for piRNA biogenesis in OSCs, rescue assays were performed in Armi-depleted OSCs. Both D862A and R949A mutants failed to restore piRNA production impaired by the depletion of endogenous Armi and the *mdg1* transposon remained desilenced (Figures 33A and 33B). These results suggest that the ATP hydrolysis-dependent, RNA unwinding and/or dissociation activities of Armi are essential for piRNA production and piRNA-mediated transposon silencing in OSCs.

The ATP hydrolysis activity of Armi is necessary for minimizing nonspecific Armi–RNA binding

To reveal how piRNA production is inhibited in OSCs expressing Armi mutants, I with help from Dr. Hirotsugu Ishizu performed FAST-iCLIP using Armi D862A and R949A mutants expressed in OSCs (Figure 34 and Table 1). The genomic distribution of Armi-FAST-iCLIP tags of each mutant was greatly changed from those of WT Armi, but in similar manners (Figure 35A and Table 2): the rate of WT Armi-FAST-iCLIP tags mapped to transposons was 36.10%, while those of Armi D862A and R949A mutants were decreased to 11.11% and 10.91%, respectively. In contrast, the distribution of Armi-FAST-iCLIP tags to 3′ UTR was greatly increased from 33.52% to 62.47% and 66.77%, respectively. The genomic distribution of WT Armi-FAST-iCLIP shown in Figure 35A is to some extent different from that shown in Figure 10A because different antibodies were used in each FAST-iCLIP experiment; anti-Armi and anti-Myc, respectively. Altogether, the ATP hydrolysis activity of Armi appeared to have a great impact on the selective association of Armi with piRNA precursors *in vivo.*

Armi-bound nucleotides of both Armi mutants hardly mapped on *flam* locus (Figure 35B). This matched our observations that *idefix*-piRNAs originating solely from the *flam* locus were no longer detected in OSCs expressing the Armi mutants (Figure 33A). Armi FAST-iCLIP tags of both mutants were mapped on *tj* 3′ UTR, but to a much lesser extent to WT control (Figure 35C). On the other hand, expression of WT Armi in Armi-lacking OSCs restored the the production of *tj*-piRNAs, but not the D862A mutant (Figure 36). These results raised a possibility that Armi mutants bound to cellular RNAs promiscuously. The correlation analysis revealed that both mutants bound a wider range

of mRNAs in accordance with their expression levels (Figure 37). It should be noted that all of these phenotypes were greatly similar to those of endogenous Armi in Yb-depleted OSCs (Figure 22). Therefore, not only Yb but also ATP hydrolysis activity of Armi is necessary for accumulating transposon-targeting piRNAs and consequently transposon silencing. Although piRNAs were still produced upon Yb depletion (Figure19), when the Armi mutants were expressed in OSCs lacking endogenous Armi, piRNAs bound with Piwi were barely detected (Figure 38). In this context, the defect in ATP hydrolysis by Armi more severely impaired piRNA production than Yb loss. The role of Armi may slightly be more significant than that of Yb in piRNA production.

Armi-FAST-iCLIP analysis suggested that Armi mutants bound to cellular RNAs promiscuously in the cytosol. Immunofluorescence analysis showed that Armi mutants partially accumulated at Yb bodies, and mainly dispersed in the cytosol (Figure 39). WT Armi was localized to Yb bodies, as was endogenous Armi. Because both WT Armi and Armi mutants co-immunoprecipitated with Yb (Figure 40), the cytoplasmic dispersal of Armi mutants was not due to the loss of Yb–Armi interaction. I speculated that the cytoplasmic dispersal of Armi mutants would be induced by their lack of ATP hydrolysis activity. Rather, Armi mutants would be retained on cellular RNAs, thereby failing to accumulate at Yb bodies (Figure 41) This was very likely considering the fact that both Armi mutants bound a wider range of RNAs in accordance with their expression levels, irrespective of the participation of Yb in determining substrate binding. *In vitro* assays also supported this idea (Figure 31).

ATP hydrolysis activity of Armi is required for both the directional unwinding activity and RNA releasing activity (Figures 30 and 31). To examine whether the unwinding activity of Armi is also required for piRNA production, I tethered Armi mutants

in Armi-depleted OSCs. Although the production of artificial piRNAs was induced upon WT Armi tethering in Armi-depleted OSCs, the production of artificial piRNA from the downstream of the BoxB sites was impaired by the tethering of Armi mutants in Armi-depleted OSCs (Figure 42). Previous study demonstrated that the tethering of Armi mutants induced but reduced the generation of piRNAs to some extent under conditions where endogenous Armi still expressed (Pandey et al., 2017). Consistent with this, slight piRNA production from the downstream of the BoxB sites was observed when Armi mutants were tethered in control OSCs (Figure 42). This slight piRNA production might be due to the presence of endogenous Armi. Immunofluorescence analysis also agreed with this notion; Armi mutants partially localized to Yb bodies. These results suggested that the directional unwinding activity of Armi was also required for piRNA processing.

How does Armi WT dissociate from abruptly bound RNAs? One possibility is that Armi naturally get rid of the bound RNAs after completing the RNA unwinding reaction. In cells, however, RNAs should be mostly covered with numerous proteins such as poly(A) binding proteins and/or ribosomes, suggesting that RNA dissociation and RNA unwinding are uncoupled events. Otherwise, Armi could bind cellular RNAs within the exposed regions. Considering that genic piRNAs are produced from various mRNAs, bound RNAs could be recruited to piRNA processing machinery without dissociation and remodeled by the unwinding activity of Armi, resulting in piRNA production, of which the efficiency might be dependent on bound proteins or RNA structures. I propose that ATP hydrolysis activity of Armi is required for two consecutive steps in piRNA biogenesis, namely, prior to piRNA processing (for minimizing promiscuous Armi–RNA binding) and during piRNA processing (for relaxing and translocating piRNA precursors for Zuc cleavage).

Armi is recruited to Yb body through protein–protein interaction with Yb

Recent studies suggested that Armi is directly translocated from Yb body to mitochondria (Munafò et al., 2019; Yamashiro et al., 2019). Furthermore, it was also shown that unlike Armi ATPase mutants, some Armi mutants were localized to Yb bodies although they are defective in piRNA biogenesis; Armi N756A and ΔN34 mutants (Yamashiro et al., 2019). These also indicate that the cytoplasmic dispersal of Armi D862A and R949A mutants was not from Yb bodies. Armi ΔN34 mutant, which lacked 34 N-terminal amino acids, localized to Yb bodies (Yamashiro et al., 2019) (Figure 43). Armi ΔN34 mutant exhibited RNA dependent ATPase activity and directional RNA unwinding activity (Figures 44 and 45). These results also support the notion that ATPase activity is required for Armi to accumulation at Yb bodies. Armi N756A mutant, which lacked RNA binding activity, also localized to Yb bodies (Yamashiro et al., 2019) (Figure 43). This result indicates that the localization of Armi at Yb bodies is RNA-independent and Armi is recruited to Yb body through proein–protein interaction with Yb. *In vitro* binding assay demonstrated that recombinant Armi bound specifically with myc-Yb immunoprecipitated from transfected OSCs (Figure 46). Supporting this result, recent study also showed that Yb associates with Armi through the extended Tudor domain of Yb (Hirakata et al., 2019).

Discussion

Discussion

In OSCs, both Yb and Armi are indispensable for piRNA biogenesis and they associate with piRNA precursors during piRNA processing. Both Yb and Armi localize to Yb bodies, where piRNA biogenesis occurs in OSCs. A clear difference between Yb and Armi emerged when their interdependence was examined in terms of Yb localization and RNA binding activity of Armi: depletion of Armi had little change in Yb localization to Yb body. However, in the absence of Yb, Armi not only dispersed in the cytosol but also exhibited a decreased ability to associate with piRNA precursors (Saito et al., 2010). These findings implied that Yb regulates the behavior of Armi in terms of both cellular localization and RNA binding. Therefore, Yb should be located at a higher level than Armi in the hierarchy in piRNA biogenesis in OSCs.

This study provided further evidence to support this. FAST-iCLIP showed that Armi bound piRNA precursors similarly to Yb, but the loss of Yb allowed Armi to bind cellular RNAs promiscuously in accordance with their expression levels. In this context, Armi bound with *flam* RNAs only weakly. As a result, the abundance of transposon-targeting piRNAs was severely reduced, resulting in transposon desilencing. These results indicate that the specificity of Armi–RNA binding is controlled by Yb and further support the original concept that Yb is located at a higher level than Armi in the hierarchy in the piRNA pathway.

I also provided support for a new concept that Yb is not absolutely required for piRNA processing in OSCs and that Armi has an intrinsic ability to produce piRNAs from its bound RNAs. I presented two lines of evidence supporting this concept. The first one was that, although the production efficiency was rather low, piRNAs were still produced

from various Armi-bound RNAs in Yb-depleted OSCs. The second was provided by the tethering assays, in which Armi produced artificial piRNAs regardless of the presence or absence of Yb, from reporter RNAs to which Armi is forced to bind. In the tethering assays, Armi was physically attached to reporter RNAs through the affinity of λN and BoxB, which I assumed was sufficient to localize Armi in quite close proximity to the RNA substrates, thereby increasing the efficiency of piRNA production.

From the perspective of piRNA precursors, the piRNA processing is an RNA degradation process. Therefore, the degradation machinery should strictly define the substrates which should be degraded. I found that OSCs possess two means of controlling it. The first one is governed by Yb, as has already been noted above (Figure 47A). Yb ensures piRNA precursor selection by increasing the specificity of Armi–RNA (*i.e.,* piRNA precursor) binding. The second one is governed by Armi itself (Figure 47B). *In vitro* assays demonstrated that WT form of Armi is released from the bound RNAs using its ATP hydrolysis activity. Based on this, it is implied that Armi might repeat the desorption until it encounters the "correct RNAs" (*i.e.,* piRNA precursors) that have primarily been bound with Yb. When the ATP hydrolysis activity of Armi was somewhat impaired, it is no longer released from bound RNAs. As a result, Armi with non-piRNA-precursors failed to localize at Yb bodies. Unlike in the case of Yb depletion, they were not led to the processing step because the tethering assays suggested that ATP hydrolysis activity was also required for Armi to translocate and unwind the bound RNAs to facilitate Zuc-mediated piRNA production. Thus, it is claimed that piRNA biogenesis in OSCs is controlled by multiple layers, in which Yb and Armi play distinct but collaborative roles to ensure the production of transposon-targeting piRNAs and transposon repression.

The important issue to be solved here is how Armi is retained onto correct RNAs to which Yb has already bound in normal OSCs. Earlier studies showed that Upf1, the central player in nonsense-mediated mRNA decay (NMD) (Leeds et al., 1991; Sun et al., 1998), is selectively released from bound non-target mRNAs using its ATP hydrolysis activity (Lee et al., 2015). This situation is similar to what I found for Armi in piRNA biogenesis. Both Armi and Upf1 belong to the same SF1 helicase family. In *S. cerevisiae*, the ATP hydrolysis activity of Upf1 is inhibited by the translation release factors eRF1 and eRF3 to keep the protein on the RNAs which should be led to NMD (Czaplinski et al., 1998). I postulate that Yb and/or other factors analogously inhibit the ATP hydrolysis activity of Armi, preventing the protein from detaching from the target RNAs, on which processing machineries might be formed stably. I am currently engaged in identifying such factors by conducting Yb body proteome and RNAi-based gene screening by assuming that, in the absence of such factors, Armi should fail to localize to Yb bodies. In *in vitro* assays, the presence of Yb hardly affected the efficiency of the ATP hydrolysis activity of Armi (data not shown), suggesting that Yb itself might not be the factor I am looking for.

Recent study showed that ATP hydrolysis-defective mutants of Armi in germline cells, where Yb does not express, lost substrates selectivity and failed to localize to Nuage or mitochondria and dispersed in the cytoplasm (Ge et al., 2019). The ATP hydrolysis activity of Armi in germline cells should be also required for releasing from bound non-specific RNAs, as well as in follicle cells. Thus, piRNA factors expressed in both germline cells and follicle cells might repress the ATP hydrolysis activity of Armi. One of the possibilities is Piwi; considering the requirement of Piwi for shuttling of Armi–piRNA precursors from Yb bodies to mitochondria (Yamashiro et al., 2019), Piwi (and

Aub in germline cells) might regulate the ATP hydrolysis activity of Armi.

Tethering assay suggested the possibility that Armi is involved in the phased piRNA biogenesis, using 5′-to-3′ directional RNA unwinding activity. However, the direct evidence for the involvement has not been obtained. Recent studies showed that Armi bound to piRNA precursors both before and after Zuc-mediated cleavage (Ge et al., 2019; Ishizu et al., 2019) (Figure 48). It is also shown that some 5′ ends of Armi-bound RNAs are generated Zuc-independently, suggesting that other unknown nuclease(s) determine the "first" 5′ ends of piRNA precursors. Considering that tethering of Armi without Yb induced piRNA biogenesis, Armi might recruit such nuclease to piRNA precursors. However, such nuclease(s) have not been identified by a comprehensive mass spectrometry analysis of Armi co-immunoprecipitates (data not shown).

In human cells, Upf2 binds Upf1 and enhances ATP hydrolysis activity, leading to mRNP remodeling and recruitment of degradation machinery (Chamieh et al., 2008). Considering that the directional RNA unwinding activity of Armi is involved in phased piRNA biogenesis at mitochondria, a functional counterpart of Upf2 may play an important role at mitochondria. As Armi is interacted with Daed and Gasz at mitochondria (Munafò et al., 2019; Yamashiro et al., 2019), such factors might enhance the ATP hydrolysis activity of Armi.

Conclusion

Conclusion

In this study, I confirmed that Yb and Armi have distinct roles in piRNA biogenesis. Although Armi has an intrinsic ability to produce piRNAs from its bound RNAs, Yb defines the piRNA precursors and funnels them to Armi to ensure the production of transposon-targeting piRNAs. In addition, I revealed ATP hydrolysis activity of Armi is required for the detachment from non-target RNAs in the cytoplasm. This detachment facilitates the localization of Armi to Yb body. Armi is recruited to Yb body by direct interaction with Yb. ATP hydrolysis activity is also required for phased piRNA biogenesis, translocating on piRNA precursors (Figures 41 and 48). In this context, the function of Armi is similar to that of MOV10L1. MOV10L1 selectively binds to piRNA precursors and facilitates piRNA biogenesis using its directional unwinding activity. However, the model for piRNA biogenesis in mouse male germ cells is slightly different from the model for piRNA biogenesis in fly ovarian somatic cells. At present, the functional counterpart of Yb in mice and other species is still unknown. The mechanism of how ATP hydrolysis activity of Armi is regulated by other piRNA factors is also still elusive. Future analyses to solve these issues are thus anticipated.

Figures
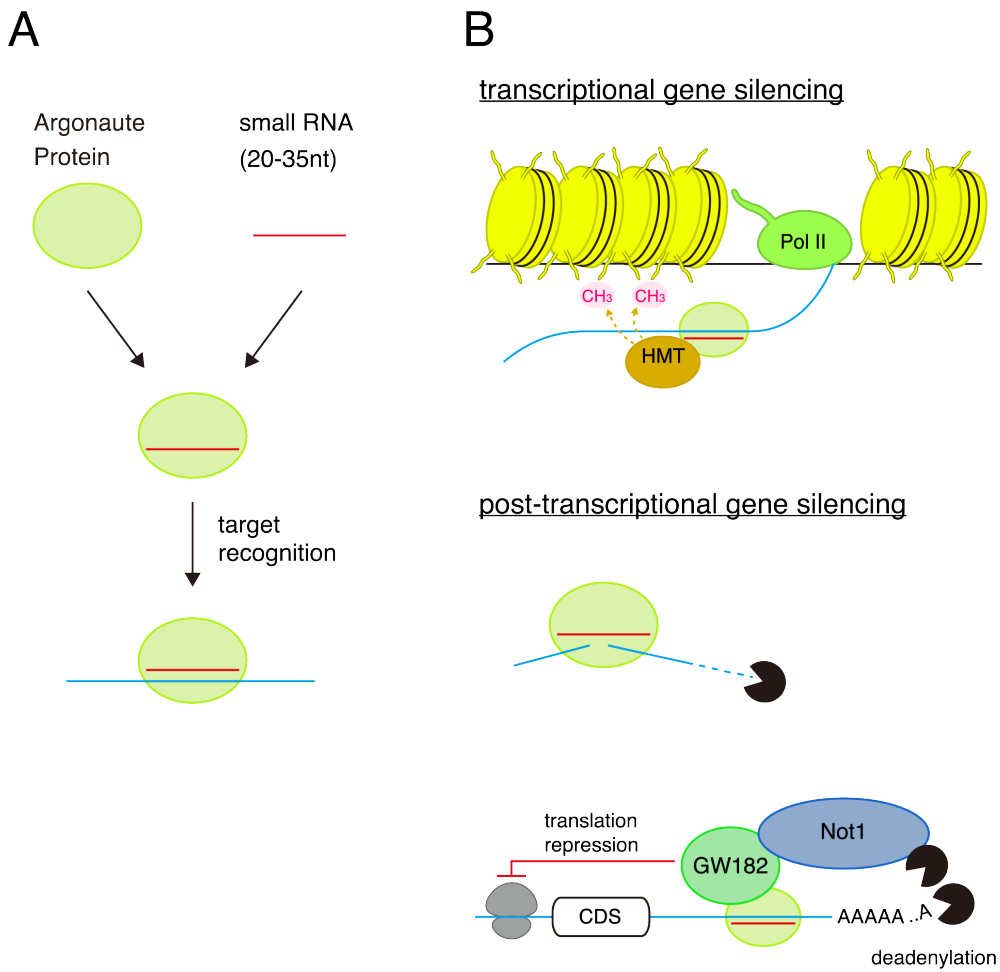
Figure 1. An overview of RNA silencing

(A) RNA-induced silencing complex (RISC) comprised of Argonaute protein and small RNA recognizes a target RNA by Watson–Crick base paring.

(B) The mechanism of RNA silencing. In transcriptional gene silencing (upper panel), RISC recruits histone methyl transferase (HMT), resulting in repressive histone modifications; i.e. trimethylation of histone H3 lysine 9. In post-transcriptional gene silencing (lower panel), RISC cleaves target RNAs or recruit translation repressive complexes, inducing RNA degradation.
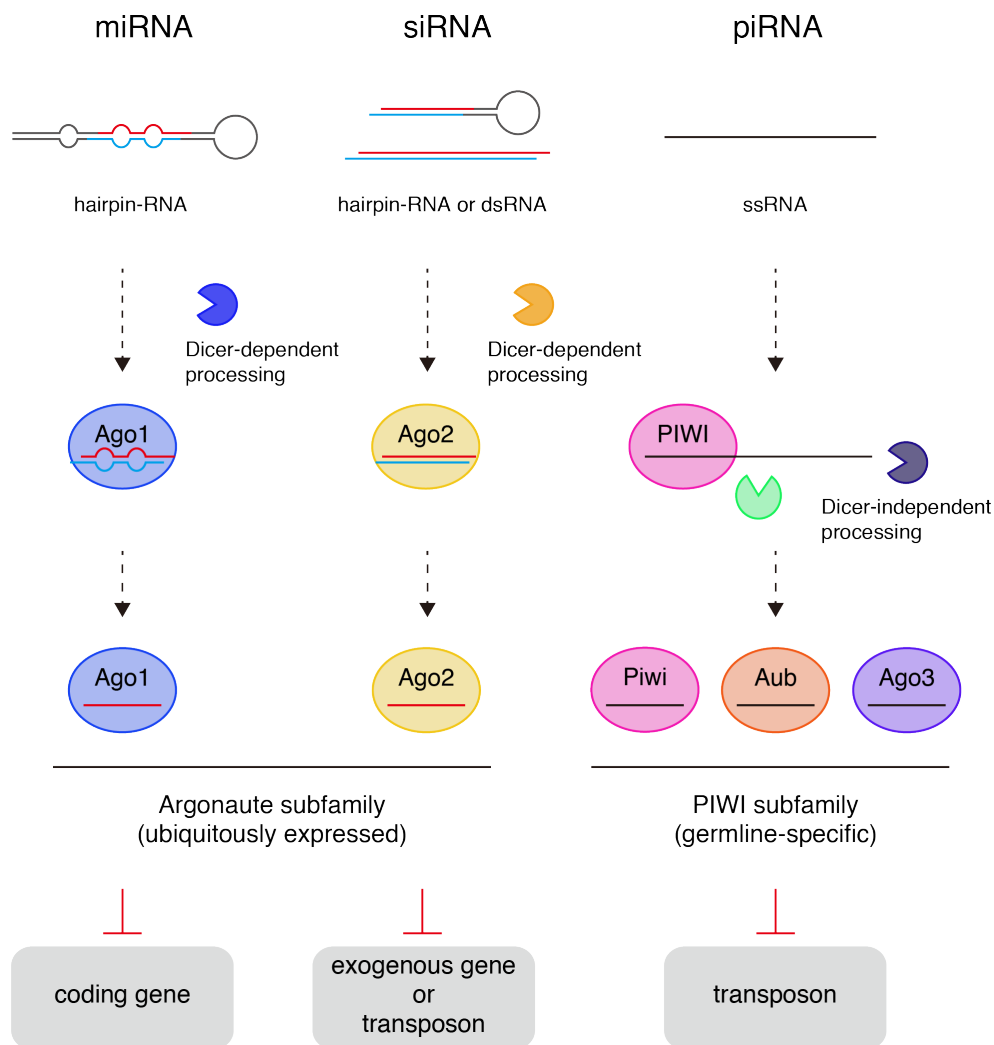
Figure 2. Classification of Argonaute proteins and small RNAs in *Drosophila*

An overview of three RNA silencing pathways in *Drosophila*. These pathways differ in their substrates, biogenesis, effector proteins and target RNAs.
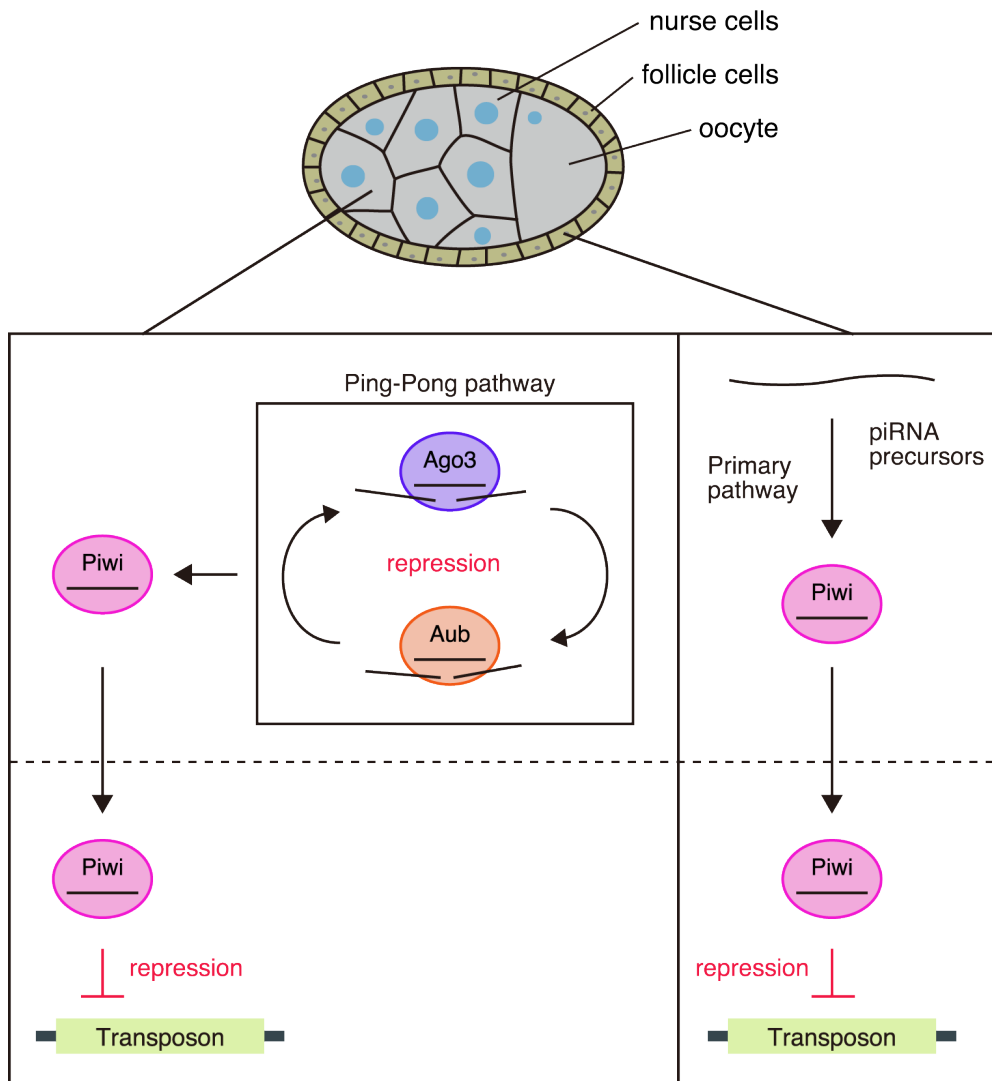
Figure 3. An overview of piRNA pathway in *Drosophila* ovary

Cartoon of *Drosophila* egg chamber and overview of piRNA pathway in *Drosophila* ovary.

In germ cells, three PIWI proteins express. In contrast, somatic cells express only Piwi.
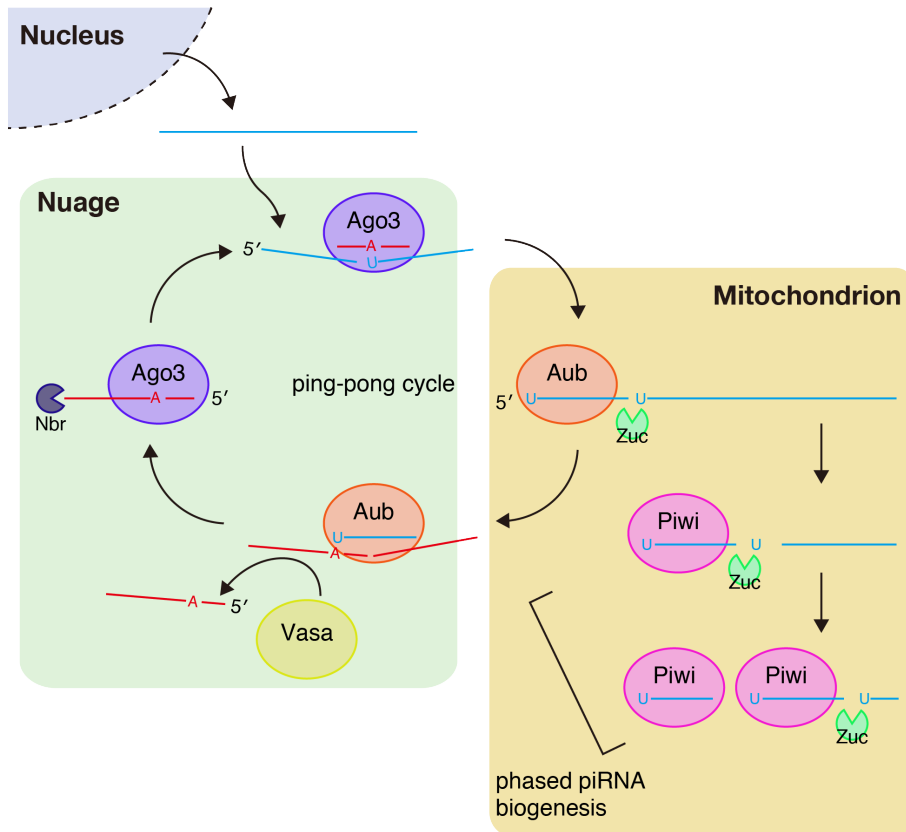
Figure 4. Current model of piRNA biogenesis in *Drosophila* germ cells

Current model of piRNA biogenesis in *Drosophila* germ cells.
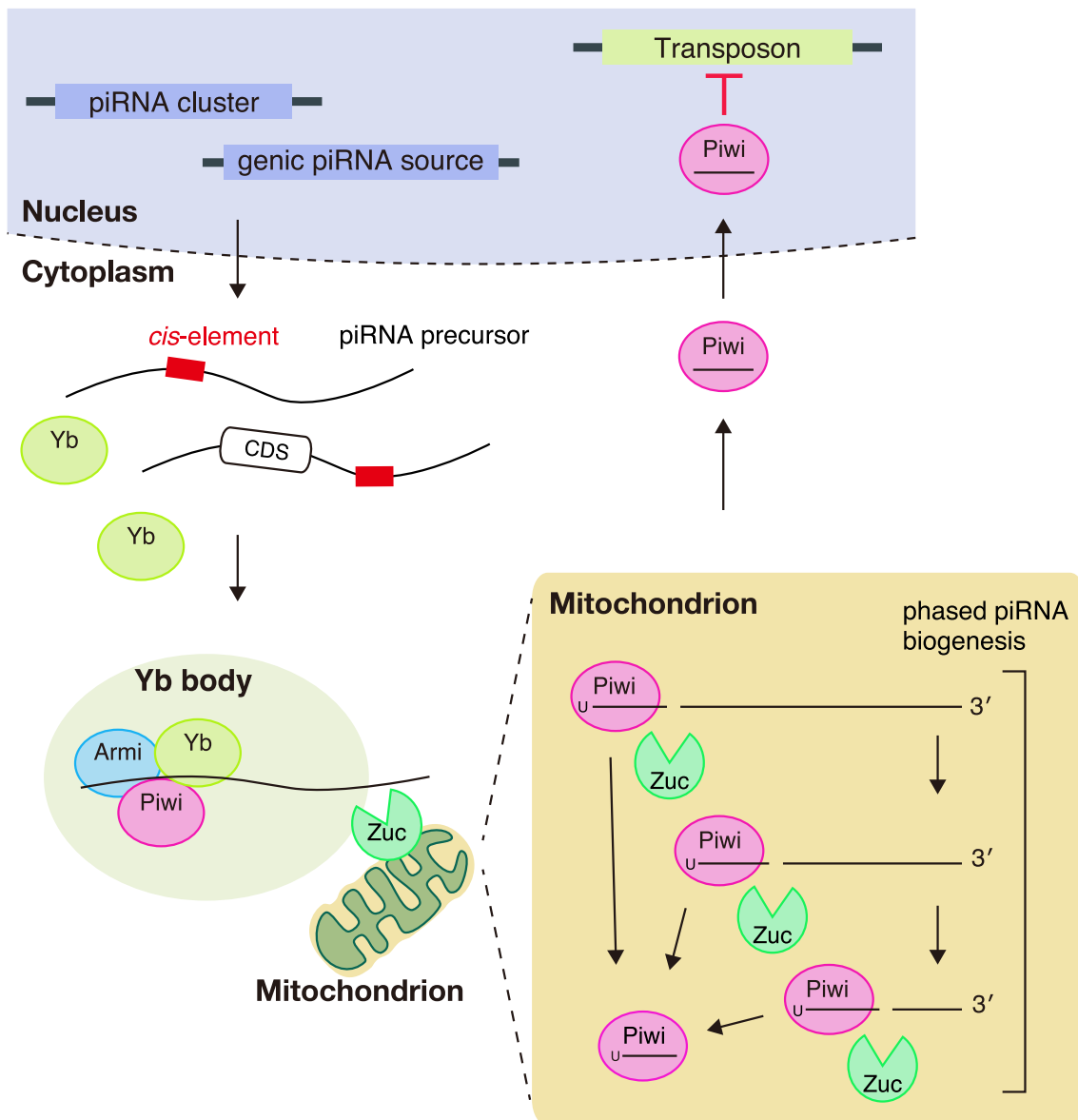
Figure 5. Current model of piRNA biogenesis in *Drosophila* ovarian somatic cells

Current model of piRNA biogenesis in *Drosophila* ovarian somatic cells.

Figure 6
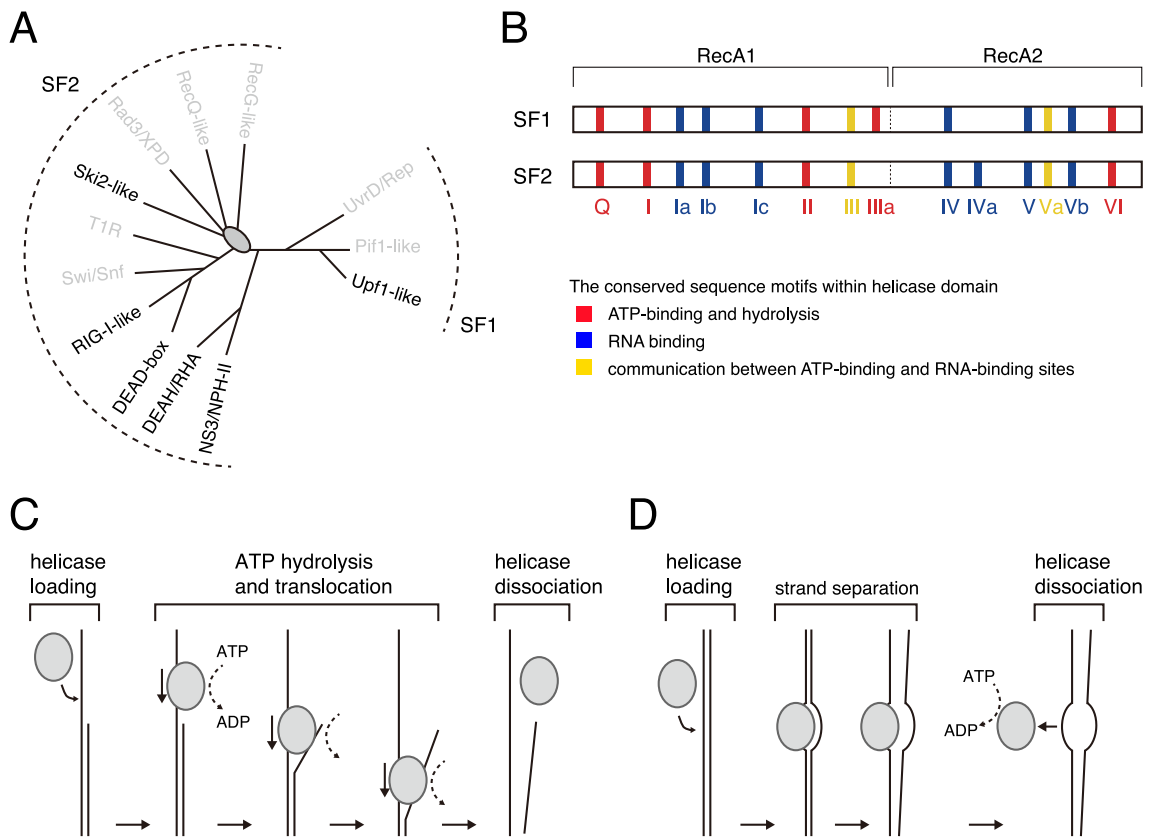
Figure 6. Classification and function of RNA helicases in eukaryotes

(A) Unrooted cladogram of the SF1 (right) and SF2 (left) helicase families. This figure is modified from Fairman-Williams et al., 2010. RNA helicases are shown in black.

(B) Schematic representation of the helicase core of SF1 and SF2 helicase families. The sequence motifs are colored according to their predominant biochemical function.

(C) Schematic representation of canonical duplex unwinding.

(D) Schematic representation of local strand separation.
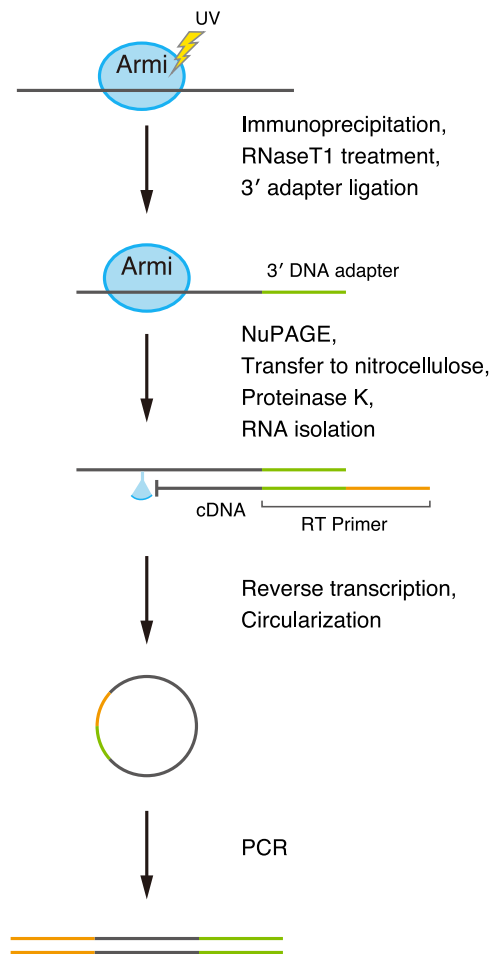
Figure 7. Schematic representation of FAST-iCLIP protocol

After UV irradiation, the covalently linked RNA is co-immunoprecipitated with Armi, treated with RNase T1 and ligated to an adapter at the 3′ end. After protein gel electrophoresis and nitrocellulose membrane transfer, the region corresponding to the expected size of the Armi–RNA complex is excised and subjected to proteinase K treatment to isolate RNA. Isolated RNA is converted into a cDNA library, circularized and deep sequenced.
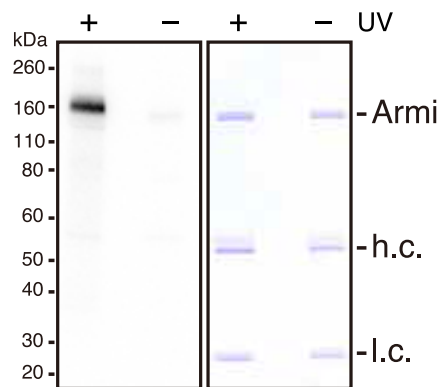
Figure 8. Analysis of UV-crosslinked Armi–RNA complexes

Crosslinked RNA was ^{32}P-labeled for detection (left panel). Armi–RNA complexes were separated by SDS-PAGE, exposed to a Phosphorimager, and subjected to CBB staining (right panel). Cells that were not irradiated with UV were used as controls for the specificity of the Armi–RNA signal.
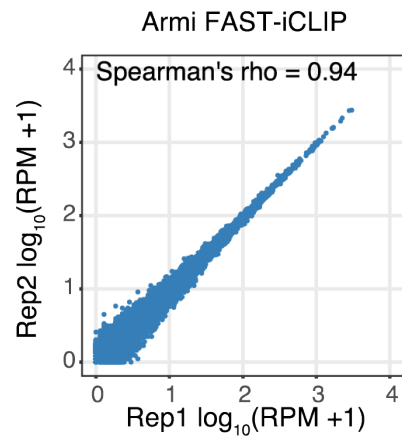
Figure 9. Reproducibility of Armi FAST-iCLIP

Scatter plots of normalized mRNA abundance in two Armi FAST-iCLIP replicate experiments. Each dot represents a protein-coding gene.

Figure 10. Genomic distribution and mapping of Armi FAST-iCLIP tags and piRNA

(A) Annotation of *Drosophila* genome regions where Armi FAST-iCLIP tags (red) and piRNAs (blue) were mapped. Eight other categories of genome region (repeat, rRNA, tRNA, snRNA, snoRNA, pseudogene, lincRNA, pre-miRNA) are summarized as "others." Replicates were merged into one dataset.

(B and C) Browser view of Armi crosslinking nucleotides and piRNAs found in the genomic *flam* locus corresponding to 21,631,891 to 21,790,731 on the X chromosome (B) and *tj* (C). Replicates were merged into one dataset.

Figure 11. Assessment of specific binding of Armi to piRNA precursors

Left panel: Scatter plots of normalized mRNA abundance for piRNA-seq (x-axis) and RNA-seq (y-axis). Right panel: Scatter plots of normalized mRNA abundance for piRNA-seq (x-axis) and Armi FAST-iCLIP (y-axis). Each dot represents a protein-coding gene. Genic piRNA coding genes are classified into three groups based on their piRNA expression: corresponding to low (purple), intermediate (ocher), and high (green) levels of expression. Magenta dots represent mRNAs in which no reads were mapped in the piRNA-seq data.
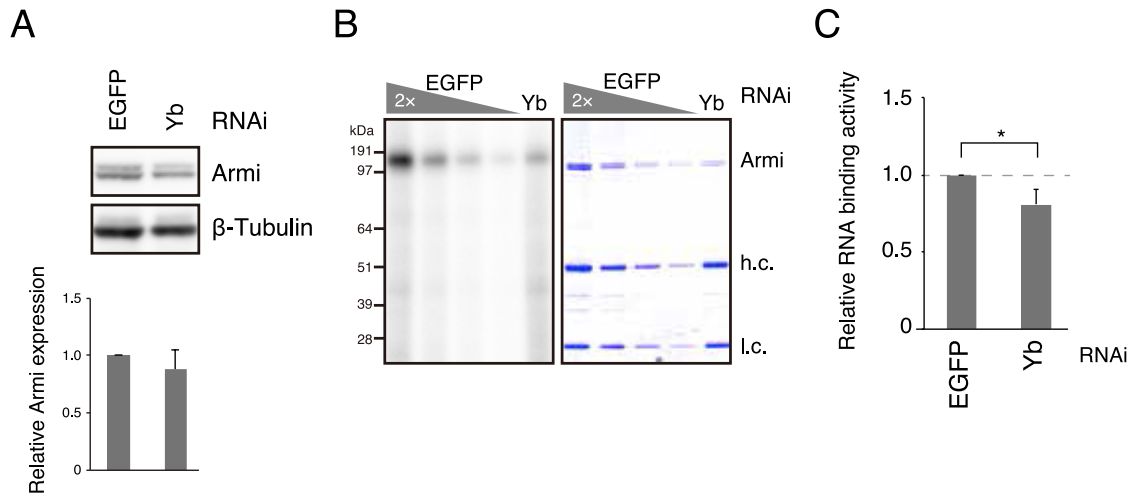
Figure 12. Analysis of UV-crosslinked Armi–RNA complexes in EGFP- or Yb-depleted OSCs

(A) The expression levels of Armi in EGFP knockdown or Yb knockdown cells were examined by western blotting. Graph shows the relative expression level of Armi normalized to tubulin. Data are presented as mean ± SEM for biological triplicates.

(B) Left: Autoradiography of an SDS-PAGE gel for detecting the Armi–RNA complexes isolated from EGFP knockdown or Yb knockdown cells. Right: Coomassie-stained SDS-PAGE gel.

(C) The graph on the right represents the ratio of RNA binding activity of Armi in Yb-depleted cells normalized by RNA binding activity of Armi in EGFP-depleted cells. Data are presented as mean ± SEM for biological triplicates. *$p < 0.05$ (paired Student's *t*-test, two-tailed).
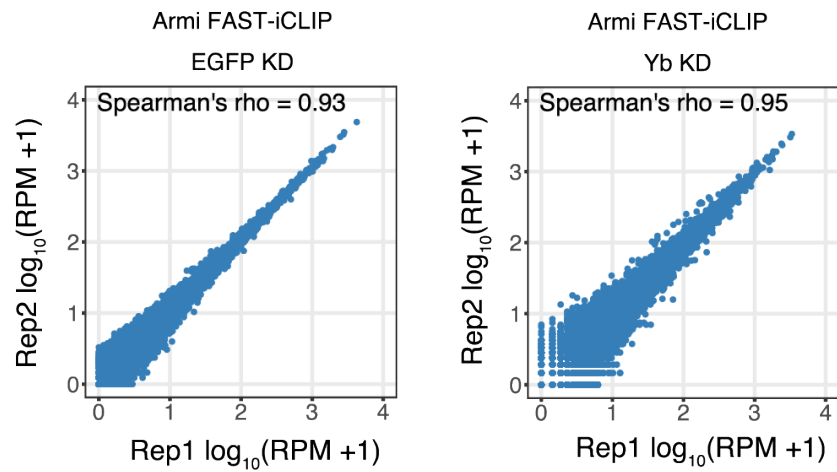
Figure 13. Reproducibility of Armi FAST-iCLIP from EGFP- or Yb-depleted OSCs

Left panel: Scatter plots of normalized mRNA abundance in two Armi FAST-iCLIP replicate experiments of EGFP knockdown samples (Spearman's rho = 0.93). Right panel: Scatter plots of normalized mRNA abundance in two Armi FAST-iCLIP replicate experiments of Yb knockdown samples (Spearman's rho = 0.95). Each dot represents a protein-coding gene.
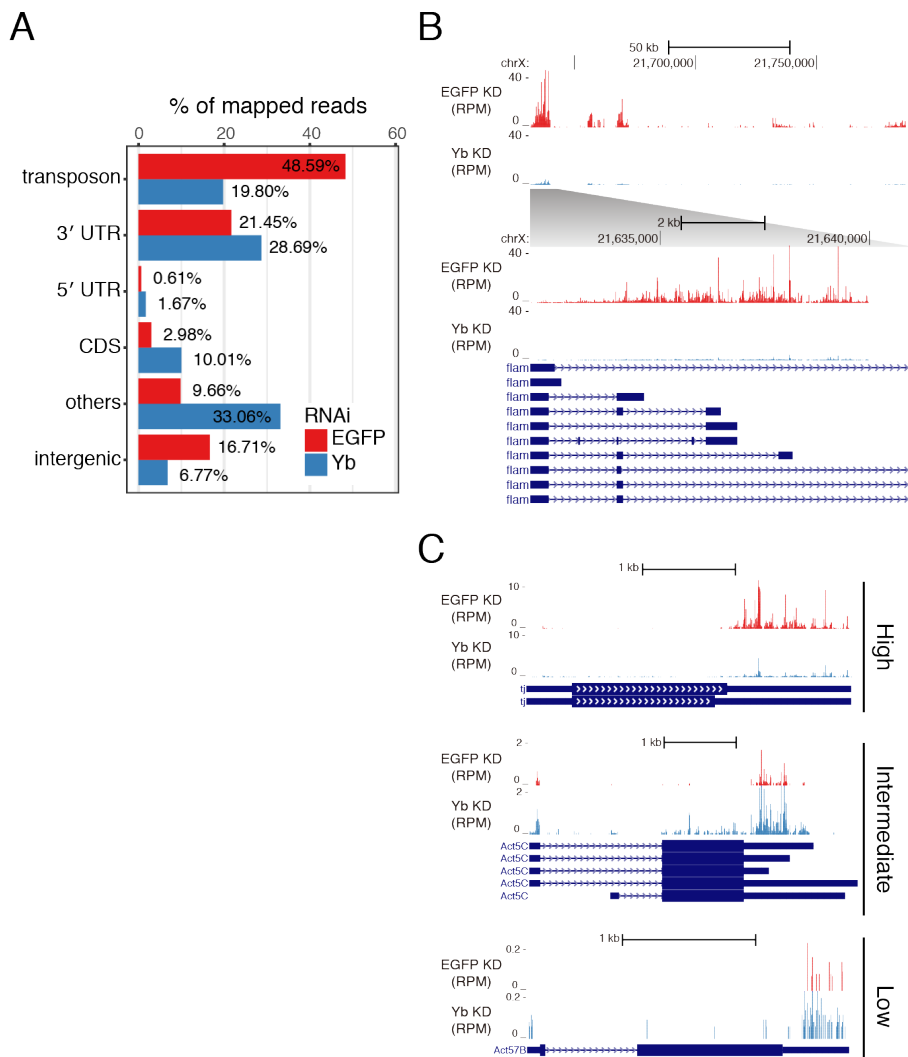
Figure 14. Genomic distribution and mapping of Armi FAST-iCLIP tags from EGFP-or Yb-depleted OSCs

(A) Annotation of *Drosophila* genome regions where Armi FAST-iCLIP tags from EGFP knockdown (red) and Yb knockdown (blue) in OSCs were mapped. Eight other categories of genome region (repeat, rRNA, tRNA, snRNA, snoRNA, pseudogene, lincRNA, pre-miRNA) are summarized as "others." Replicates were merged into one dataset.

(B and C) Browser view of Armi crosslinking nucleotides found in the genomic *flam* locus corresponding to 21,631,891 to 21,790,731 on the X chromosome (B) and *tj*, *Act5C*, and *Act57B* (C). *tj*, *Act5C*, and *Act57B* are classified into "high", "intermediate", and "low" group, respectively. Replicates were merged into one dataset.
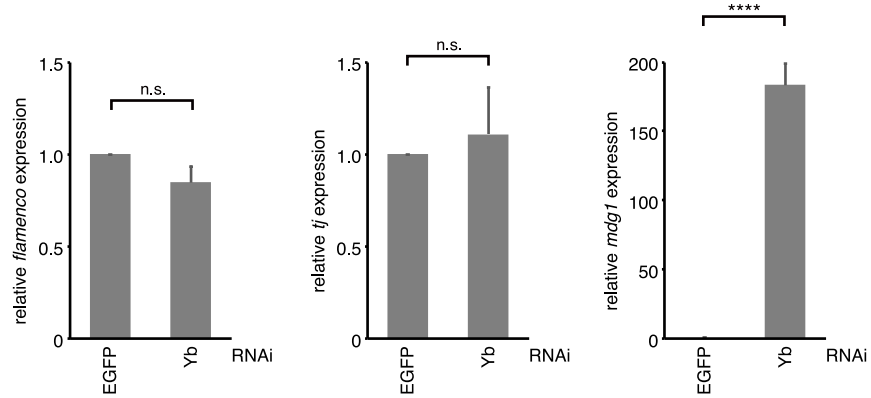
Figure 15. Comparison of RNA expression in EGFP- or Yb-depleted OSCs

The expression levels of the *flam*, *tj* and *mdg1* were monitored by RT-qPCR. Data are presented as mean ± SEM for biological triplicates. ****$p < 0.0001$ (paired Student's *t*-test, two-tailed).
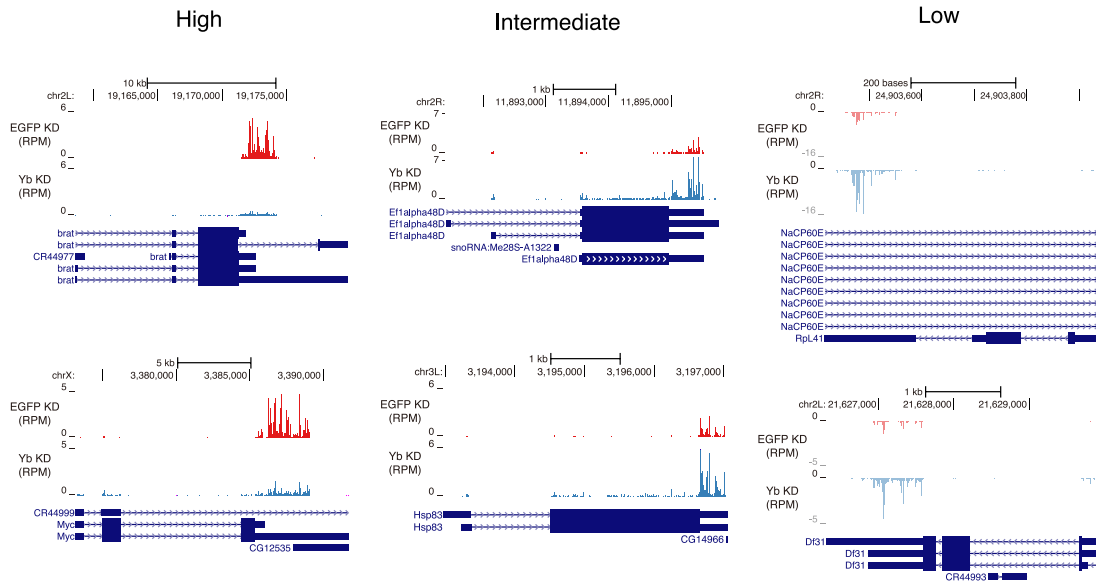
Figure 16. Mapping of Armi FAST-iCLIP tags from EGFP- or Yb-depleted OSCs

Browser view of Armi crosslinking nucleotides found in *brat*, *Myc*, *Ef1alpha48D*, *Hsp83*, *RpL41*, and *Df31*. *brat* and *Myc* are representative examples of genes in the "High" group. *Ef1alpha48D* and *Hsp83* are representative examples of genes in the "Intermediate" group. *RpL41* and *Df31* are representative examples of genes in the "Low" group. Replicates were merged into one dataset.
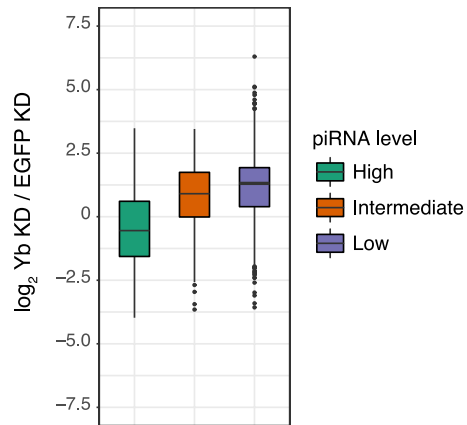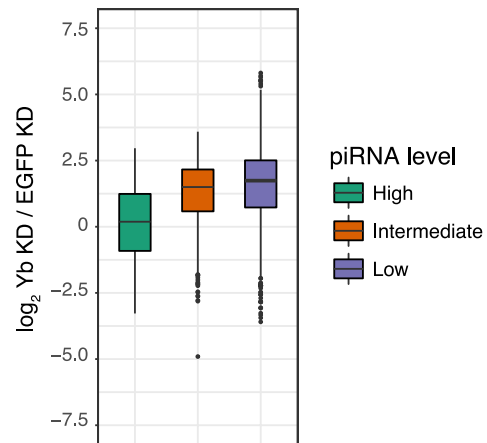
Figure 17. Fold changes in Armi FAST-iCLIP tags in Yb- vs. EGFP-depleted OSCs
Boxplot showing log2 fold change of RPM values of genes in "High," "Intermediate," and "Low" groups in Armi-FAST-iCLIP upon Yb depletion.
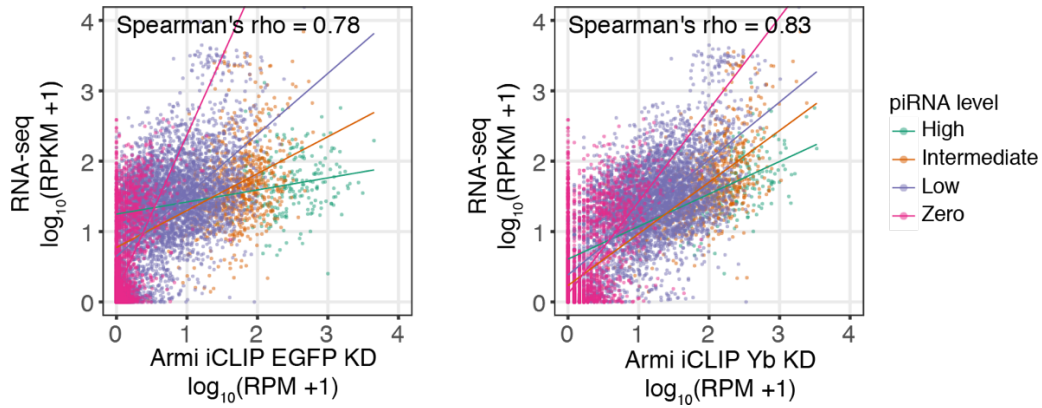
Figure 18. Assessment of promiscuous binding of Armi to cellular RNAs in EGFP-or Yb-depleted OSCs

Left panel: Scatter plots of normalized mRNA abundance for Armi FAST-iCLIP in EGFP knockdown OSCs (x-axis) and RNA-seq (y-axis). Right panel: Scatter plots of normalized mRNA abundance for Armi FAST-iCLIP in Yb knockdown OSCs (x-axis) and RNA-seq (y-axis). Each dot represents a protein-coding gene. Colors correspond to the genic piRNA expression level. The regression lines are indicated by solid lines.
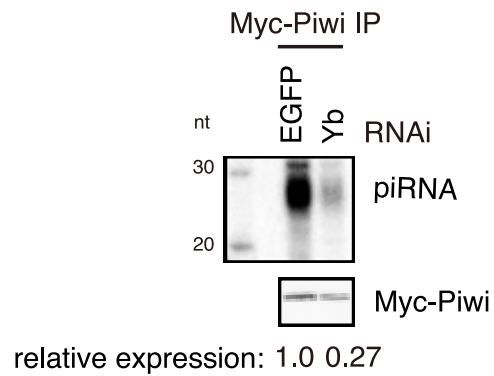
Figure 19. Detection of piRNAs associated with Myc-Piwi from EGFP- or Yb-depleted OSCs

Analysis of piRNAs associated with Myc-Piwi in EGFP knockdown OSCs and Yb knockdown OSCs. RNAs isolated from Myc-Piwi immunoprecipitates were ^{32}P-labeled for detection. Immunoprecipitation efficiency was determined by western blotting. Relative expression level of piRNAs is indicated.
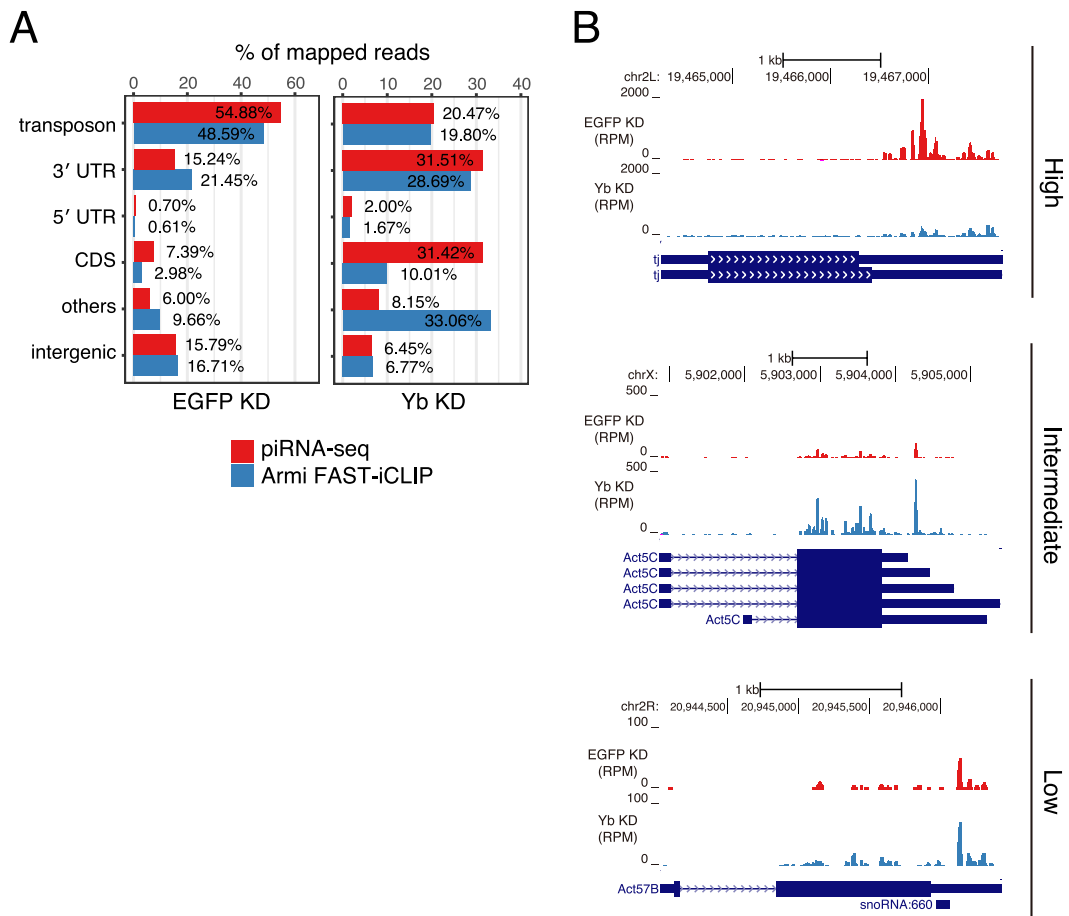
Figure 20. Genomic distribution and mapping of piRNA-seq from EGFP- or Yb-depleted OSCs

(A) Annotation of *Drosophila* genome regions where Armi FAST-iCLIP tags (blue) and piRNAs (red) from EGFP knockdown and Yb knockdown in OSCs were mapped. Replicates were merged into one dataset.

(B) Browser view of piRNA reads found in *tj*, *Act5C*, and *Act57B*. *tj*, *Act5C*, and *Act57B* are classified into "high", "intermediate", and "low" group, respectively. Replicates were merged into one dataset.

Figure 21. Fold changes in piRNA-seq in Yb- vs. EGFP-depleted OSCs

Boxplot showing log2 fold change of RPM values of genes in "High," "Intermediate," and "Low" groups in piRNA-seq upon Yb depletion.
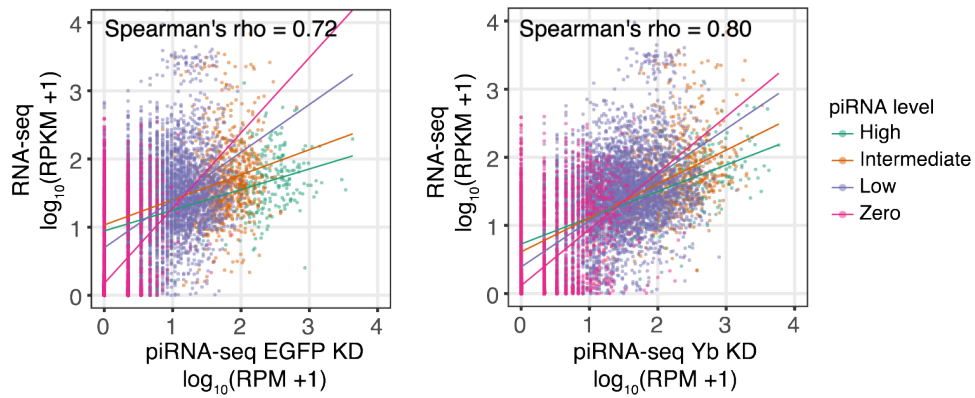
Figure 22. Assessment of promiscuous piRNA production in EGFP- or Yb-depleted OSCs

Left panel: Scatter plots of normalized mRNA abundance for piRNA-seq from EGFP knockdown OSCs (x-axis) and RNA-seq (y-axis). Right panel: Scatter plots of normalized mRNA abundance for piRNA-seq from Yb knockdown OSCs (x-axis) and RNA-seq (y-axis). Each dot represents a protein-coding gene. Colors correspond to genic piRNA expression level. The regression lines are indicated by solid lines.
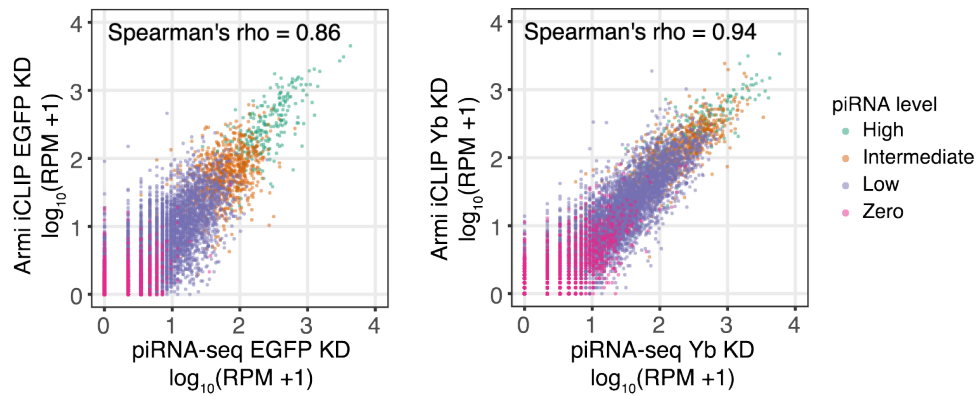
Figure 23. Assessment of correlation between Armi FAST-iCLIP tags and piRNA-seq from EGFP- or Yb-depleted OSCs

Left panel: Scatter plots of normalized mRNA abundance for piRNA-seq from EGFP knockdown OSCs (x-axis) and Armi FAST-iCLIP (y-axis). Right panel: Scatter plots of normalized mRNA abundance for piRNA-seq from Yb knockdown OSCs (x-axis) and Armi FAST-iCLIP (y-axis). Each dot represents a protein-coding gene. Colors correspond to genic piRNA expression level.
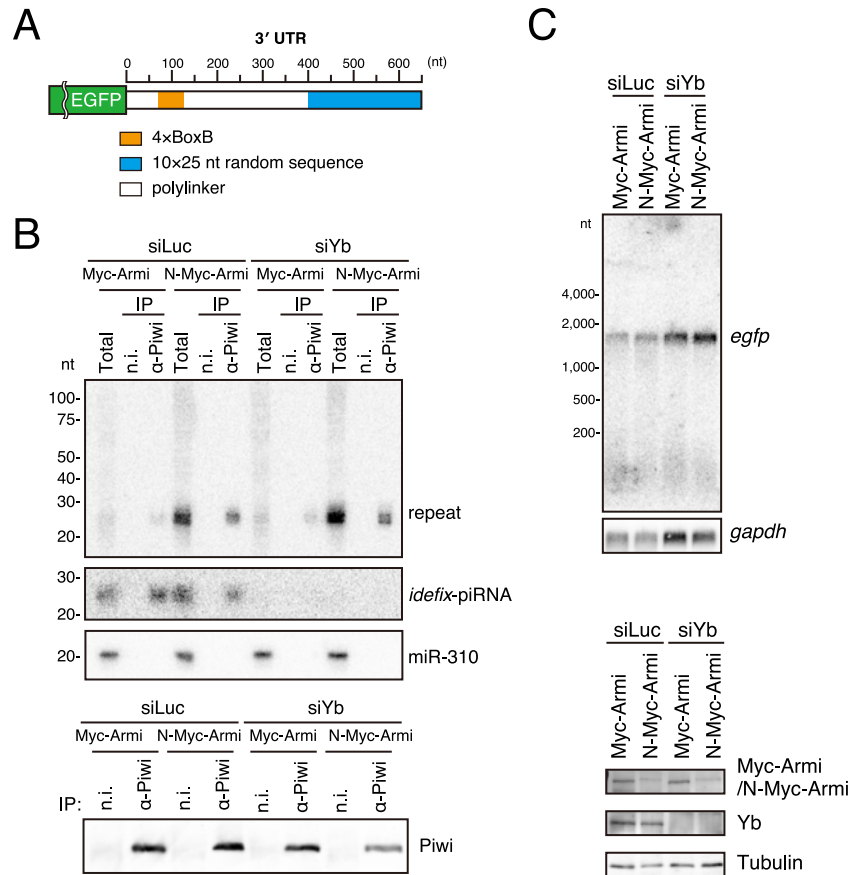
Figure 24. Tethering of Armi in Yb-depleted OSCs

(A) Schematic of the construct for expressing artificial piRNAs. Four BoxB sequences are shown in orange. Blue box indicates 25-nt random sequences, which are repeated ten times.

(B) Top: Artificial piRNAs derived from the random sequences were detected by northern blotting using a 50-nt probe complementary to two tandem repeats. As a control, endogenous *idefix*-piRNAs and miR-310 were detected. Bottom: Immunoprecipitation efficiency was determined by western blotting.

(C) Top: Northern blotting using an *egfp* probe shows the expression levels of full-length transcripts. The *gapdh* transcript was visualized as a loading control. Bottom: Western blotting shows the expression levels of Myc-tagged or λN-Myc-tagged protein and endogenous Yb in OSCs after transfection. Tubulin was detected as a loading control.
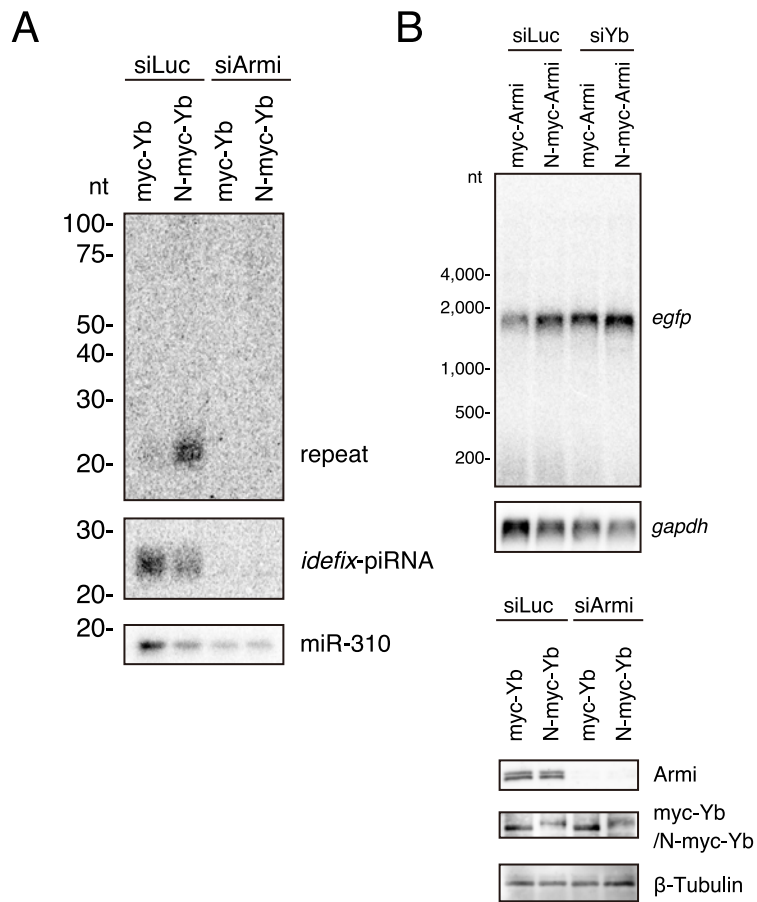
Figure 25. Tethering of Yb in Armi-depleted OSCs

(A) Artificial piRNAs derived from the random sequences were detected by Northern blotting using a 50-nt probe complementary to two tandem repeats. As a control, endogenous *idefix*-piRNAs and miR-310 were detected.

(B) Top: Northern blotting using an *egfp* probe shows the expression levels of full-length transcripts. The *gapdh* transcript was visualized as a loading control. Bottom: Western blotting shows the expression levels of Myc-tagged or λN-Myc-tagged protein and endogenous Armi in OSCs after transfection. Tubulin was detected as a loading control.
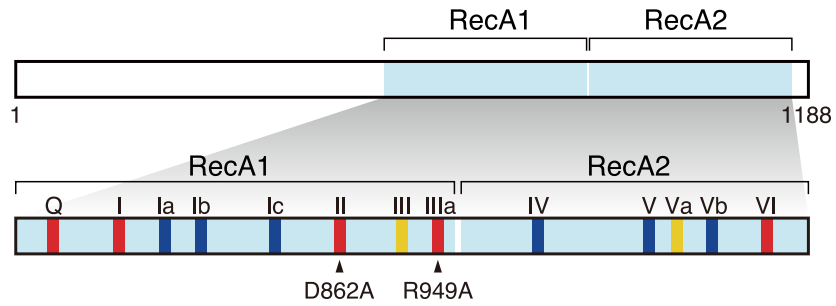
Figure 26. Schematic representation of the Armi protein domain structure

Armi contains a C-terminal helicase core consisting of two RecA-like domains RecA1 and RecA2. Amino acids in motif II or IIIa that were mutated to create Armi mutants are indicated.
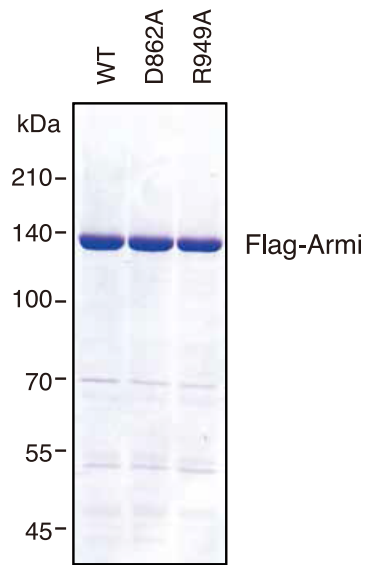
Figure 27. Quality of purified Flag-tagged Armi WT or point mutants

CBB-stained SDS-PAGE of purified Flag-tagged Armi WT, D862A, and R949A.
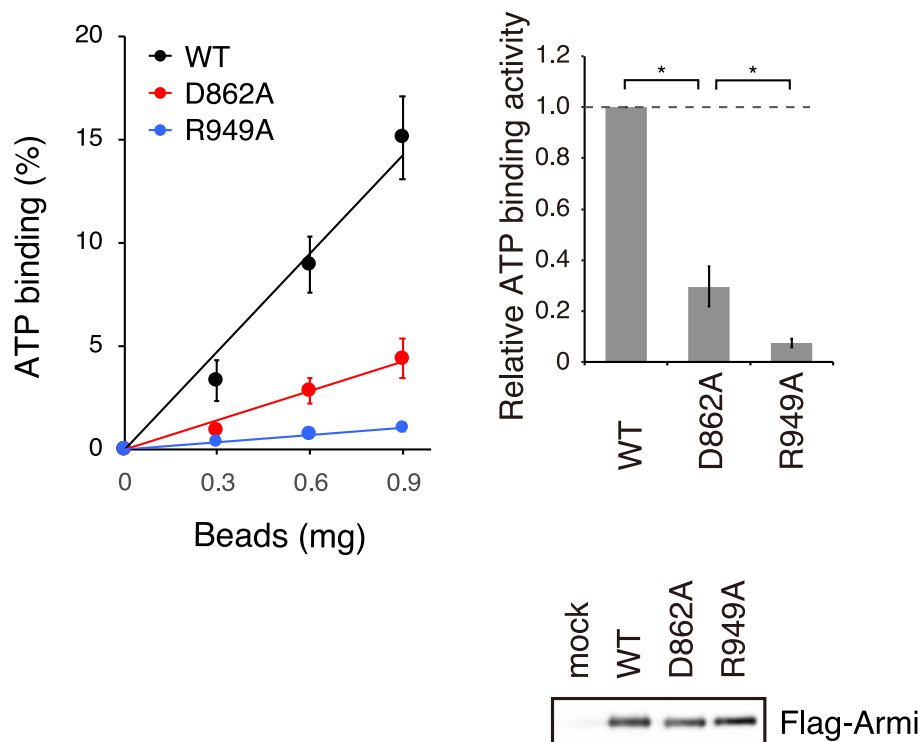
Figure 28. Quantitative analysis of ATP bound to Flag-Armi WT or point mutants

Quantitative analysis of ATP bound to Flag-Armi WT, D862A, and R949A purified from *Drosophila* S2 cells. Data are presented as mean ± SEM for three independent experiments. The graph on the right represents the ratio of ATP binding activity of mutants to that of WT normalized to the corresponding protein level. The lower graph of a western blot shows the levels of Flag-Armi used in the assay. *$p < 0.05$ (paired Student's *t*-test, two-tailed).
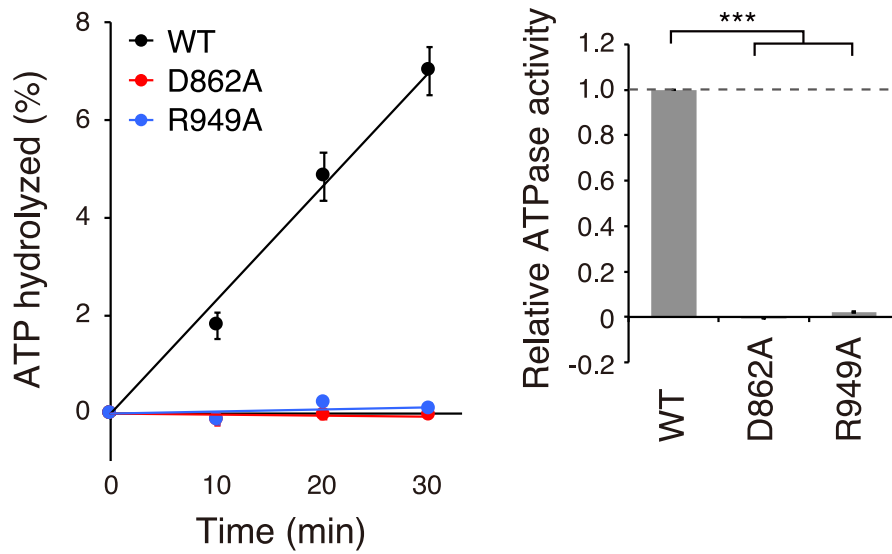
Figure 29. Quantitative analysis of ATPase activity of Flag-Armi WT and point mutants

Quantitative analysis of ATPase activity of Flag-Armi WT, D862A, and R949A. Data are presented as mean ± SEM for three independent experiments. The graph on the right presents the ratio of ATPase activity of mutants to that of WT. ***$p < 0.001$ (paired Student's *t*-test, two-tailed).
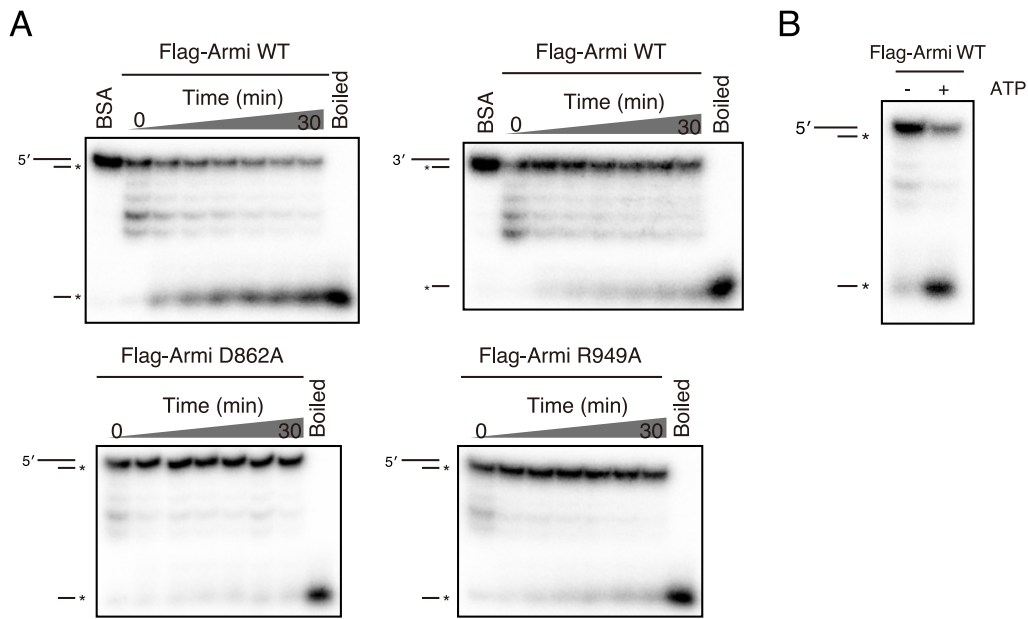
Figure 30. Analysis of unwinding activity of Flag-Armi WT and point mutants

(A) Top: Unwinding activity of Flag-Armi WT. RNA duplexes containing 5′ or 3′ overhangs were incubated with Flag-Armi WT over an increasing amount of time and separated by nondenaturing PAGE. Bottom: RNA duplexes containing 5′ overhangs were incubated with Flag-Armi D862A and R949A. Incubation with BSA for 30 min served as a negative control, while boiled RNA duplexes served as a positive control. The asterisk (*) denotes the position of the ^{32}P label.

(B) Unwinding activity of Flag-Armi WT in the presence or absence of ATP. RNA duplexes with 5′ overhang were incubated with Flag-Armi WT for 30 min and separated by nondenaturing PAGE. The asterisk (*) denotes the position of the ^{32}P label.
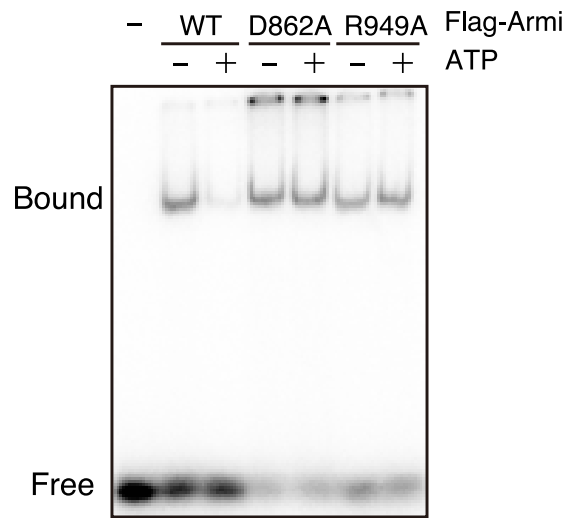
Figure 31. Analysis of RNA binding activity of Flag-Armi WT or point mutants in the presence or absence of ATP

RNA binding activity of Flag-Armi WT, D862A, and R949A. 16-nt ssRNA was 5′-end-labeled and incubated with Flag-Armi in the presence and absence of ATP.
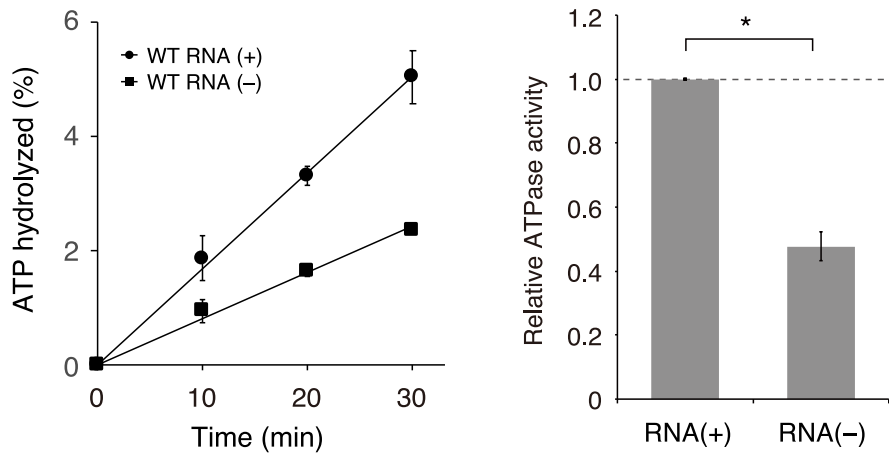
Figure 32. Quantitative analysis of ATPase activity of Flag-Armi WT in the presence or absence of RNA

Quantitative analysis of Flag-Armi WT in the presence or absence of ssRNA. Data are presented as mean ± SEM for three independent experiments. The graph on the right presents the ratio of ATPase activity of mutants to that of WT. ***$p < 0.001$ (paired Student's t-test, two-tailed).
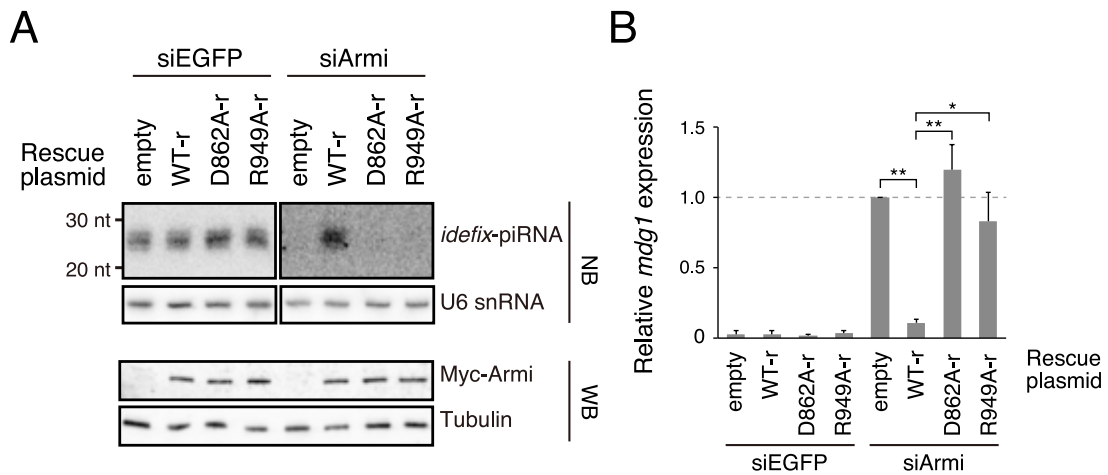
Figure 33. piRNA accumulation and transposon expression in OSCs expressing Myc-Armi WT and point mutants

(A) Top: Northern blotting with *idefix*-piRNA probe showed that siRNA-resistant WT Myc-Armi, but not D862A and R949A mutants, rescued the defects in piRNA production in endogenous Armi-depleted OSCs. As a control, OSCs were transfected with an empty vector. U6 snRNA was used as a loading control. Bottom: The expression levels of Myc-Armi WT, D862A, and R949A were examined by western blotting using an anti-Myc antibody. Tubulin was used as a loading control.

(B) The expression levels of the *mdg1* transposon were monitored by RT-qPCR. Data are presented as mean ± SEM for biological triplicates. *p < 0.05, **p < 0.01 (paired Student's *t*-test, two-tailed).
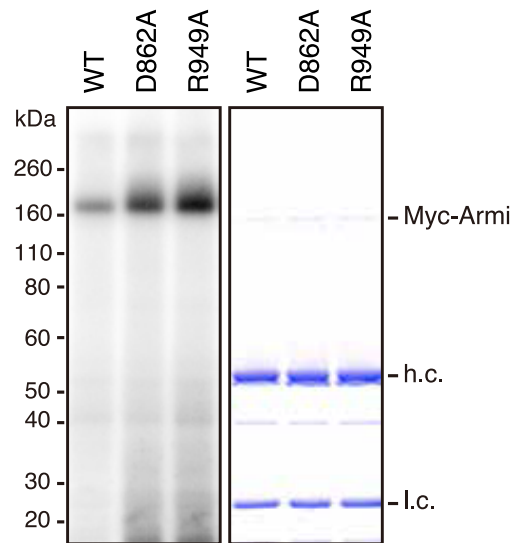
Figure 34. Analysis of UV-crosslinked Myc-Armi–RNA complexes in OSCs expressing Myc-Armi WT or point mutants

Analysis of UV-crosslinked Myc-Armi–RNA complexes using autoradiography (left panel) and CBB staining (right panel). Crosslinked RNA was radiolabeled for detection. Myc-Armi–RNA complexes were separated by SDS-PAGE, exposed to a Phosphorimager, and subjected to CBB staining.
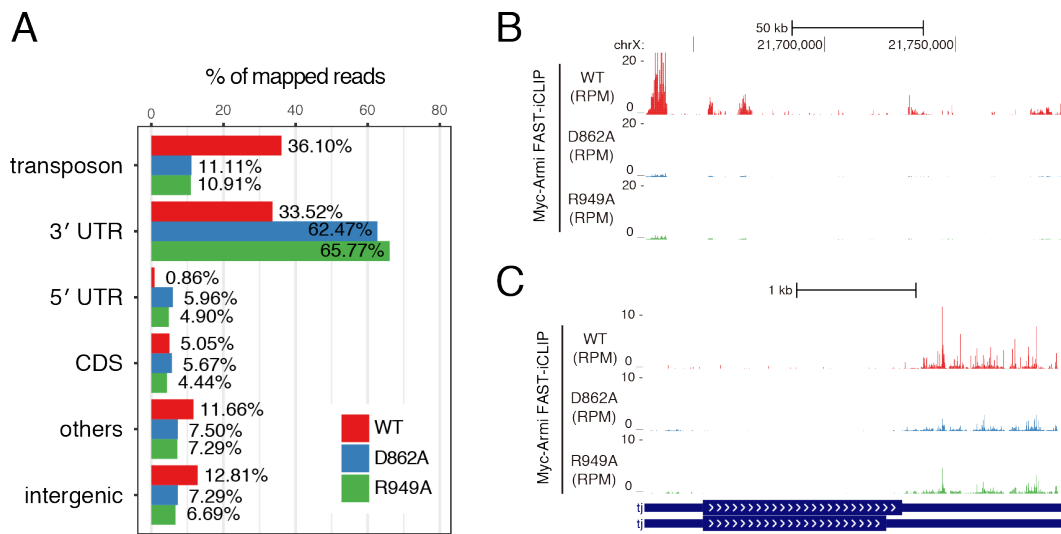
Figure 35. Genomic distribution and mapping of Myc-Armi FAST-iCLIP tags from OSCs expressing Myc-Armi WT or point mutants

(A) Annotation of *Drosophila* genome regions where FAST-iCLIP tags of Myc-Armi WT (red), D862A (blue), and R949A (green) were mapped. Eight other categories of genome region (repeat, rRNA, tRNA, snRNA, snoRNA, pseudogene, lincRNA, pre-miRNA) are summarized as "others." Replicates were merged into one dataset.

(B and C) Browser view of Myc-Armi crosslinking nucleotides found in the genomic *flam* locus corresponding to 21,631,891 to 21,790,731 on the X chromosome (B) and *tj* (C). Replicates were merged into one dataset.
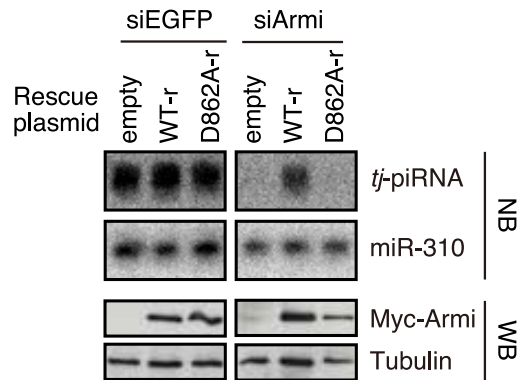
Figure 36. Detection of genic piRNAs from OSCs expressing Myc-Armi WT or D862A

Top: Northern blotting with *tj*-piRNA probe showed that siRNA-resistant WT Myc-Armi, but not D862A mutant, rescued the defects in piRNA production in endogenous Armi-depleted OSCs. As a control, OSCs were transfected with an empty vector. miR-310 was used as a loading control. Bottom: The expression levels of Myc-Armi WT and D862A were examined by western blotting using an anti-Myc antibody. Tubulin was used as a loading control.
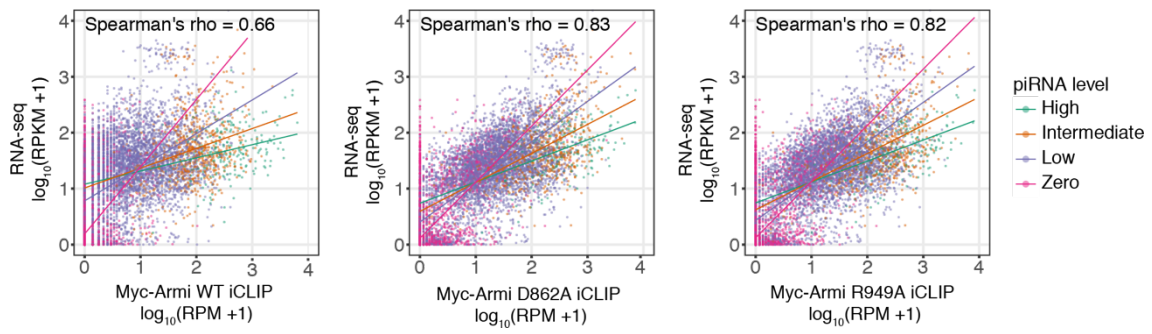
Figure 37. Assessment of promiscuous binding of Myc-Armi to cellular RNAs in OSCs expressing Myc-Armi WT or point mutants

Left panel: Scatter plots of normalized mRNA abundance for RNA-seq (y-axis) versus FAST-iCLIP (x axis) of Myc-Armi WT (left), Myc-Armi D862A (middle), and Myc-Armi R949A (right). Each dot represents a protein-coding gene. Colors correspond to genic piRNA expression level. The regression lines are indicated by solid lines.
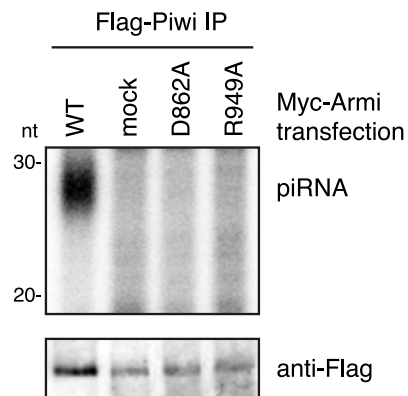
Figure 38. Detection of piRNAs associated with Flag-Piwi from OSCs expressing Myc-Armi WT or point mutants

Analysis of piRNAs associated with Flag-Piwi in OSCs expressing Myc-Armi WT, D862A, or R949A. Isolated RNA from Flag-Piwi immunoprecipitates was ^{32}P-labeled for detection. Immunoprecipitation efficiency was determined by western blotting.
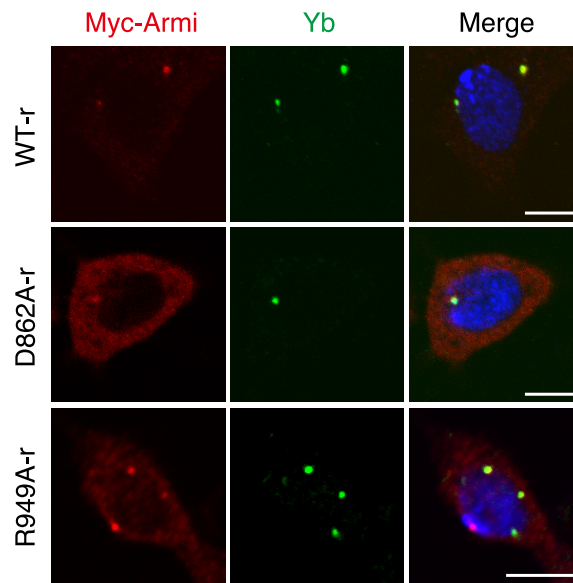
Figure 39. Subcellular localizations of Myc-Armi WT or point mutants in OSCs
Immunofluorescence using anti-Myc and anti-Yb antibodies was performed on OSCs expressing Myc-Armi WT, D862A, or R949A. Scale bars represent 10 µm.
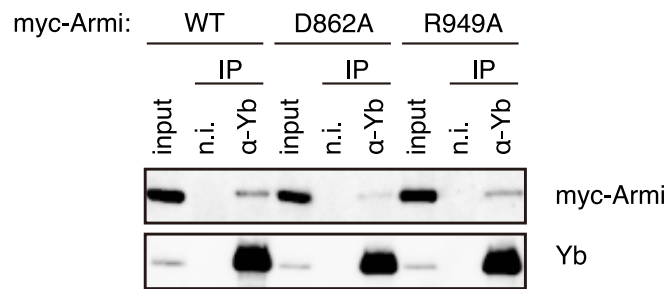
Figure 40. Interaction analysis of Yb and Myc-Armi WT or point mutants in OSCs

The Myc-tagged Armi complexes were immunoprecipitated using anti-Yb antibody. "n.i." indicates non-immune IgG used as a negative control.
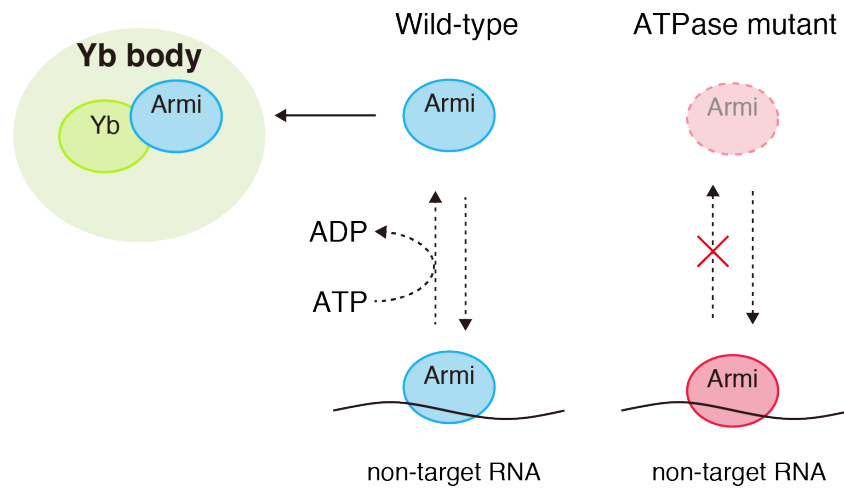
Figure 41. Proposed model for ATP-dependent RNA discrimination step by Armi preceding Yb body localization

Armi in the cytoplasm bind to non-piRNA precursors. WT Armi dissociate from bound RNAs using its ATP-dependent RNA unwinding reaction. Armi mutants are retained on bound RNAs owing to their lack of ATP hydrolysis activity.
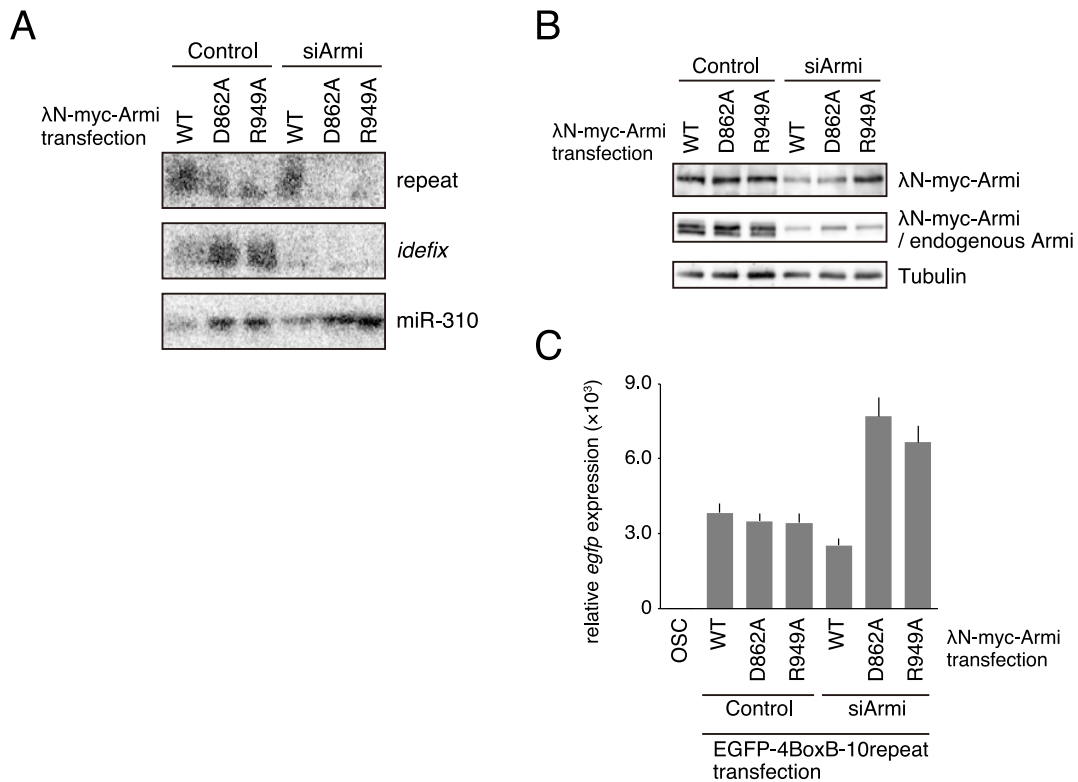
Figure 42. Tethering of Armi WT or point mutants in OSCs

(A) Artificial piRNAs derived from the random sequences were detected by Northern blotting using a 50-nt probe complementary to two tandem repeats. As a control, endogenous *idefix*-piRNAs and miR-310 were detected.

(B) Western blotting shows the expression levels of λN-Myc-tagged protein and endogenous Armi in OSCs after transfection. Tubulin was detected as a loading control.

(C) The expression levels of the *egfp* derived from the construct were monitored by RT-qPCR. Data are presented as mean ± SEM for technical triplicates.
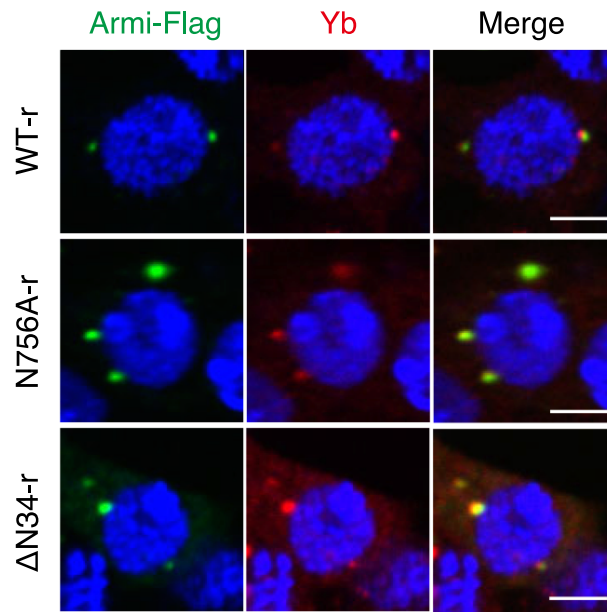
Figure 43. Subcellular localization of Armi-Flag WT or mutants in OSCs

This figure is modified from Yamashiro et al., 2019. Immunofluorescence using anti-Flag and anti-Yb antibodies was performed on OSCs expressing Armi-Flag WT, N756A, or ΔN34. Scale bars represent 5 μm.
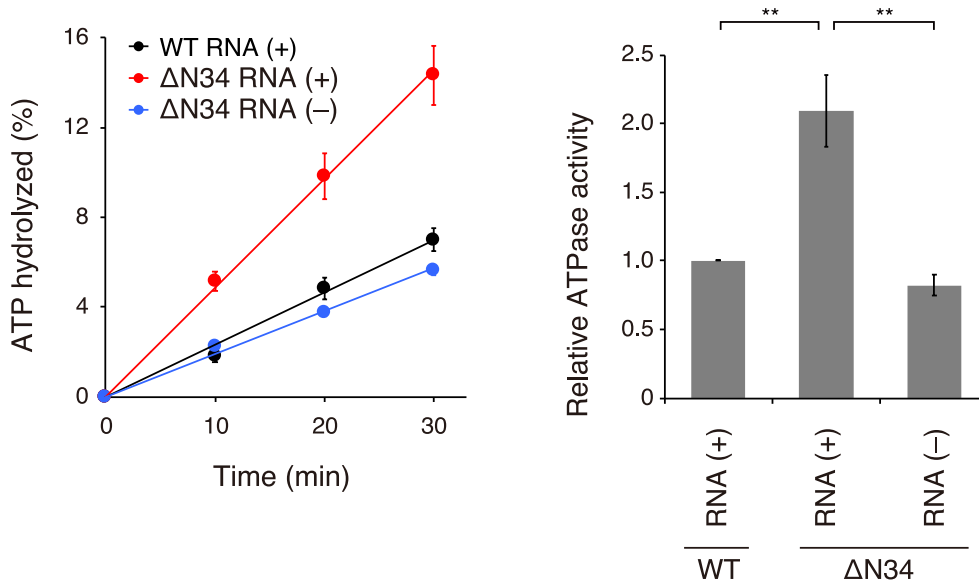
Figure 44. Quantitative analysis of ATPase activity of Flag-Armi ΔN34

Quantitative analysis of ATPase activity of Flag-Armi WT and ΔN34 in the presence or absence of ssRNA. Data are presented as mean ± SEM for three independent experiments. The graph on the right presents the ratio of ATPase activity of mutants to that of WT. ***p < 0.001 (paired Student's *t*-test, two-tailed).
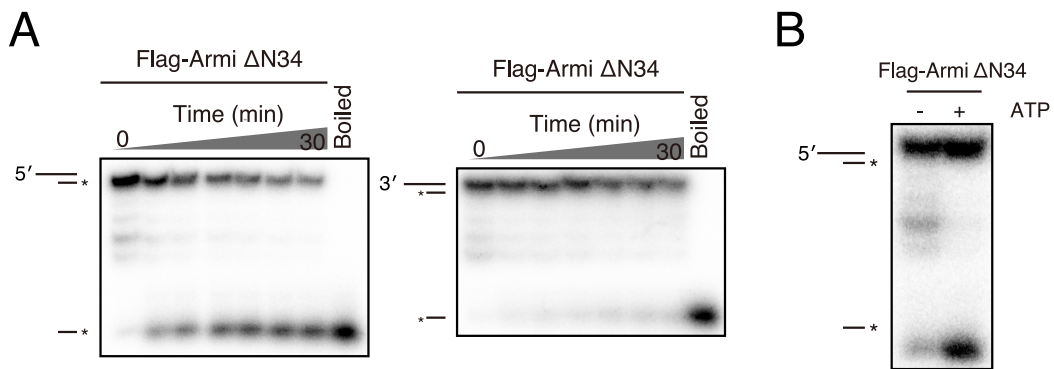
Figure 45. Analysis of unwinding activity of Flag-Armi ΔN34

(A) Unwinding activity of Flag-Armi ΔN34. RNA duplexes containing 5′ or 3′ overhangs were incubated with Flag-Armi ΔN34 over an increasing amount of time and separated by nondenaturing PAGE. Boiled RNA duplexes served as a positive control. The asterisk (*) denotes the position of the ^{32}P label.

(B) Unwinding activity of Flag-Armi ΔN34 in the presence or absence of ATP. RNA duplexes with 5′ overhang were incubated with Flag-Armi WT for 30 min and separated by nondenaturing PAGE. The asterisk (*) denotes the position of the ^{32}P label.
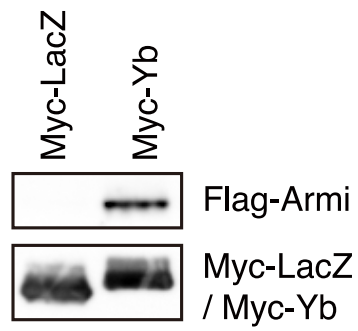
Figure 46. Protein–protein interaction analysis of Myc-Yb and Flag-Armi

In vitro binding assay of Armi and Yb. Flag-Armi was incubated with the Myc tagged proteins immunoprecipitated from transfected OSCs. Protein binding affinity and immunoprecipitation efficiency and was determined by western blotting.
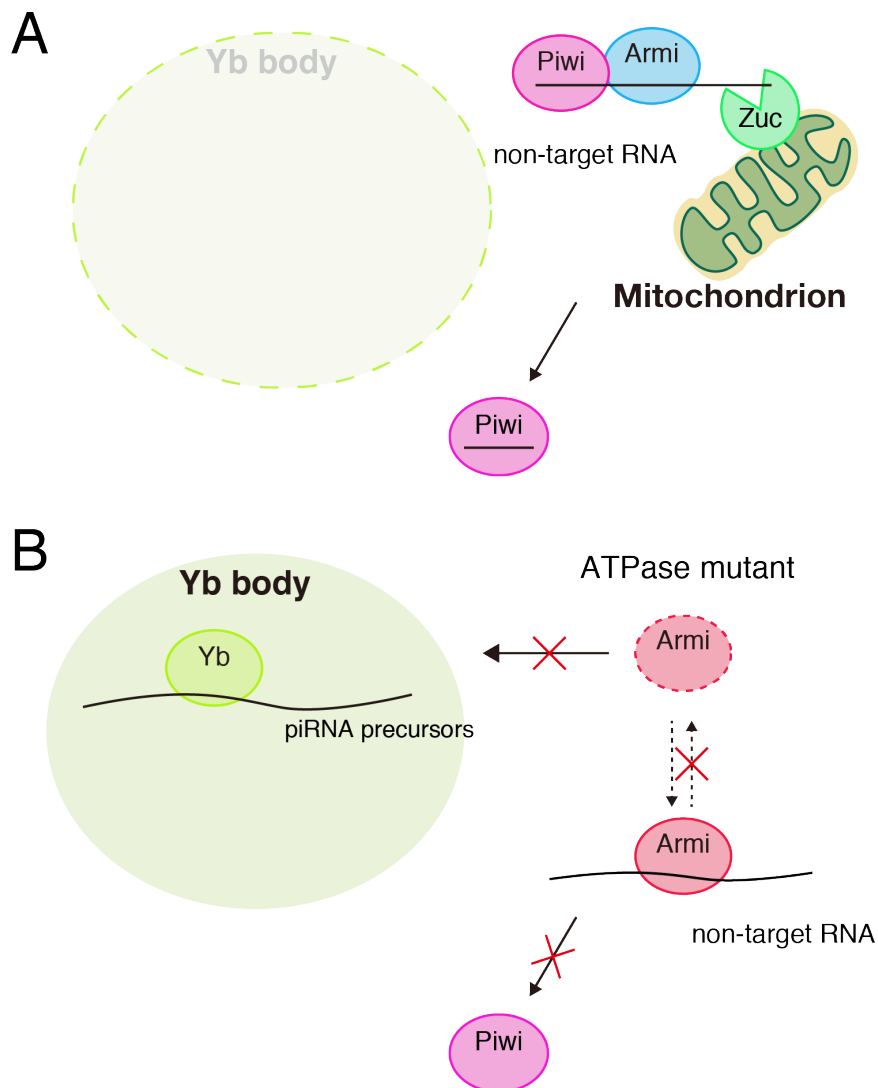
Figure 47. The behaviors of Armi in OSCs upon knockdown or mutant expressions

(A) In Yb-depleted OSCs, Armi disperses at cytosol and binds to non-piRNA precursors. Although the efficiency is rather low, Armi-bound RNAs are processed into piRNAs, which are non-transposon-targeting.

(B) In OSCs expressing ATP hydrolysis-defective mutants of Armi, Armi fails to detach from bound RNAs, resulting in cytoplasmic dispersal. As ATP hydrolysis activity of Armi is necessary for Zuc-mediated piRNA processing step, Armi-bound RNAs are not processed into mature piRNAs.
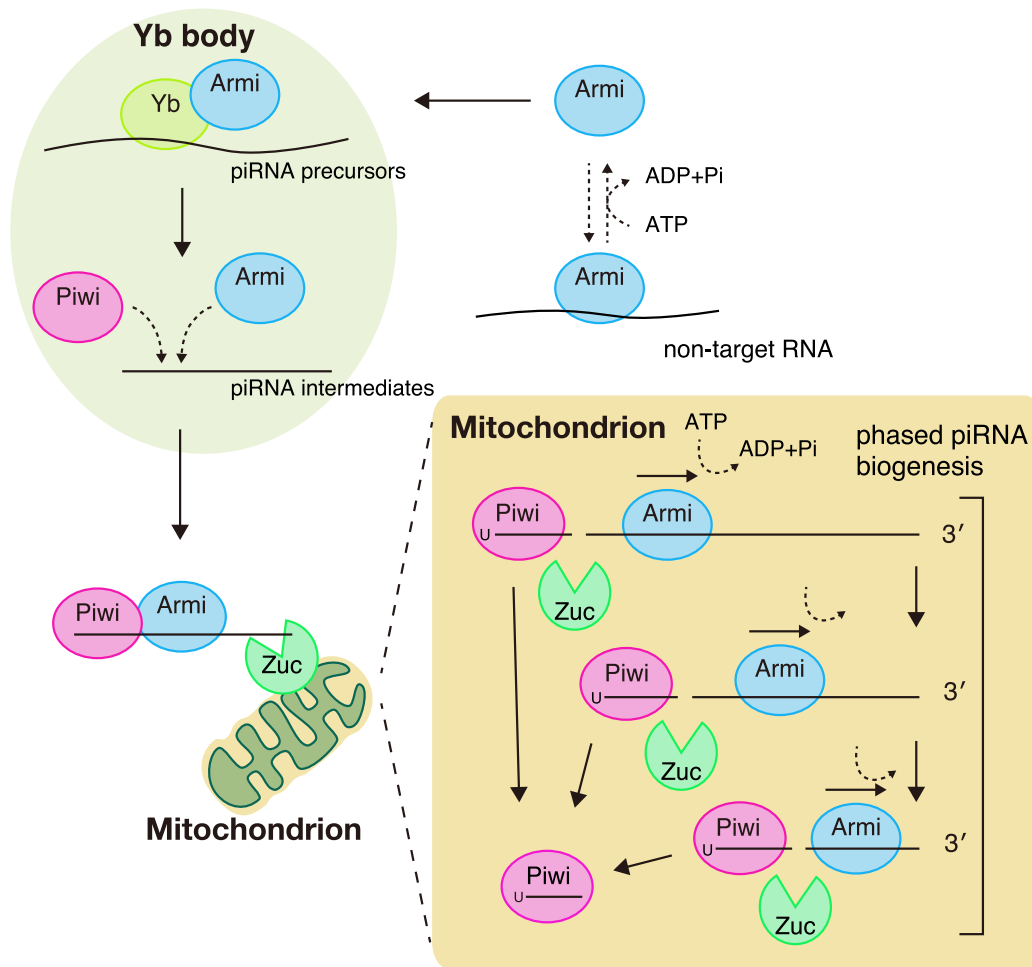
Figure 48. Proposed model for the role of Armi in piRNA biogenesis

Armi has an intrinsic ability to induce piRNA production Yb-independently. Yb is required for the selection of piRNA precursor, but not required for piRNA processing. Armi detaches from the bound RNAs by hydrolyzing ATP and is recruited to Yb body, thereby Armi binds to piRNA precursors. In piRNA processing, Armi translocates piRNA precursors from 5′ to 3′, relaxing RNA precursors for Zuc cleavage.

Tables

Table 1. Deep sequencing library statistics

Armi FAST-iCLIP relate to Figures 8-11	Replicate 1	Replicate 2
Total sequencing reads	13,431,626	29,125,864
Unique sequencing reads (Removal of PCR duplicates)	9,626,860 (100%)	12,782,772 (100%)
Uniquely mapped reads	2,920.164 (30.33%)	3,638,453 (28.46%)
Multi-mapped reads	4,052,751 (42.10%)	5,278,261 (41.29%)

piRNA-seq relate to Figures 10 and 11	
Total sequencing reads	1,890,344
Adapter removed reads	1,877,050 (100%)
Uniquely mapped reads	700,795 (37.33%)
Multi-mapped reads	1,089,329 (58.03%)

Armi FAST-iCLIP	EGFP KD		Yb KD	
relate to Figures 12-18 and 23	Replicate 1	Replicate 2	Replicate 1	Replicate 2
Total sequencing reads	16,897,392	8,684,413	9,955,067	8,268,587
Unique sequencing reads (Removal of PCR duplicates)	8,549,354 (100%)	6,547,087 (100%)	4,096,102 (100%)	3,445,622 (100%)
Uniquely mapped reads	2,865,331 (33.52%)	2,149,276 (32.83%)	1,241,545 (30.31%)	1,149,123 (33.35%)
Multi-mapped reads	3,197,186 (37.40%)	2,215,835 (33.84%)	1,040,281 (25.40%)	1,014,529 (29.44%)

piRNA-seq relate to Figures 19-23	EGFP KD	Yb KD
Total sequencing reads	942,220	1,007,430
Adapter removed reads	917,067 (100%)	992,733 (100%)
Uniquely mapped reads	347,906 (37.94%)	628,699 (63.33%)
Multi-mapped reads	464,728 (50.68%)	224,641 (22.66%)

Myc-Armi FAST-iCLIP	WT		D862A		R949A	
relate to Figures 34-37	Replicate 1	Replicate 2	Replicate 1	Replicate 2	Replicate 1	Replicate 2
Total sequencing reads	6,934,555	17,727,062	18,758,899	17,983,946	18,243,204	14,280,917
Unique sequencing reads (Removal of PCR duplicates)	1,932,190 (100%)	1,916,821 (100%)	4,937,850 (100%)	3,296,597 (100%)	4,471,764 (100%)	3,019,768 (100%)
Uniquely mapped reads	779,306 (40.33%)	736,962 (38.45%)	3,111,292 (63.01%)	2,240,368 (67.96%)	2,883,691 (64.49%)	2,033,916 (67.35%)
Multi-mapped reads	534,682 (27.67%	493,885 (25.77%)	562,123 (11.38%)	360,705 (10.94%)	467,983 (10.47%)	314,412 (10.41%)

Table 2. Annotation of mapped reads

Armi FAST-iCLIP, related to Figures 8-11

Type	
transposon	8742989 (55.02%)
repeat	365652 (2.30%)
rRNA	597041 (3.76%)
tRNA	340649 (2.14%)
snRNA	11455 (0.07%)
snoRNA	1852 (0.01%)
pseudogene	17888 (0.11%)
lincRNA	182143 (1.15%)
pre_miRNA	173 (0.00%)
3′UTR	2576773 (16.22%)
5′UTR	90636 (0.57%)
CDS	480813 (3.03%)
intergenic	241565 (15.62%)

piRNA-seq, related to Figures 10 and 11

Type	
transposon	1119978 (62.56%)
repeat	15155 (0.85%)
rRNA	6548 (0.37%)
tRNA	1720 (0.10%)
snRNA	452 (0.03%)
snoRNA	231 (0.01%)
pseudogene	1770 (0.10%)
lincRNA	62620 (3.50%)
pre_miRNA	15196 (0.85%)
3′UTR	215503 (12.04%)
5′UTR	7787 (0.43%)
CDS	78806 (4.40%)
intergenic	264358 (14.77%)

Armi FAST-iCLIP, related to Figures 12-18 and 23

Type	EGFP KD	Yb KD
transposon	5067043 (48.59%)	880004 (19.80%)
repeat	290243 (2.78%)	128122 (2.88%)
rRNA	393616 (3.77%)	897014 (20.18%)
tRNA	166602 (1.60%)	373775 (8.41%)
snRNA	7842 (0.08%)	16996 (0.38%)
snoRNA	1249 (0.01%)	1952 (0.04%)
pseudogene	15101 (0.14%)	15704 (0.35%)
lincRNA	132487 (1.27%)	36490 (0.82%)
pre_miRNA	153 (0.00%)	206 (0.00%)
3′UTR	2237113 (21.45%)	1275554 (28.69%)
5′UTR	63731 (0.61%)	74051 (1.67%)
CDS	310249 (2.98%)	444810 (10.01%)
intergenic	1742189 (16.71%)	300800 (6.77%)

piRNA-seq, related to Figures 19-23

Type	EGFP KD	Yb KD
transposon	445939 (54.88%)	174707 (20.47%)
repeat	10886 (1.34%)	13573 (1.59%)
rRNA	22291 (2.74%)	29043 (3.40%)
tRNA	652 (0.08%)	7328 (0.86%)
snRNA	201 (0.02%)	895 (0.10%)
snoRNA	143 (0.02%)	2798 (0.33%)
pseudogene	1324 (0.16%)	3576 (0.42%)
lincRNA	13513 (1.66%)	12194 (1.43%)
pre_miRNA	45 (0.01%)	202 (0.02%)
3′UTR	123844 (15.24%)	268895 (31.51%)
5′UTR	5698 (0.70%)	17032 (2.00%)
CDS	59781 (7.39%)	268088 (31.42%)
intergenic	128317 (15.79%)	55009 (6.45%)

Myc-Armi FAST-iCLIP, related to Figures 34-37

Type	WT	D862A	R949A
transposon	918572 (36.10%)	697099 (11.11%)	622121 (10.91%)
repeat	86331 (3.39%)	226714 (3.61%)	235561 (4.13%)
rRNA	105166 (4.13%)	88293 (1.41%)	47422 (0.83%)
tRNA	67501 (2.65%)	19633 (0.31%)	16031 (0.28%)
snRNA	2447 (0.10%)	39465 (0.63%)	32423 (0.57%)
snoRNA	435 (0.02%)	2919 (0.05%)	2307 (0.04%)
pseudogene	3637 (0.14%)	12606 (0.20%)	9853 (0.17%)
lincRNA	31287 (1.23%)	80610 (1.28%)	71343 (1.25%)
pre_miRNA	16 (0.00%)	170 (0.00%)	154 (0.00%)
3′UTR	853141 (33.52%)	3919934 (62.47%)	3749062 (65.77%)
5′UTR	21841 (0.86%)	373917 (5.96%)	279188 (4.90%)
CDS	128437 (5.05%)	355968 (5.67%)	253115 (4.44%)
intergenic	326024 (12.81%)	457160 (7.29%)	381422 (6.69%)

Table 3. List of oligonucleotides related to Material and Methods

Experiment	Name	Sequence	Description
vector construction	Armi D862A F	GCCGAAGCTGGTCAATGCACCGA	site-directed mutation
	Armi D862A R	GAACAGCACGTGAGTGAAAT	site-directed mutation
	Armi R949A F	GCCGCATTACCGTCGATTATGAG	site-directed mutation
	Armi R949A R	GTAATTGTACAACAATTTGG	site-directed mutation
	Flag F	ATCGGGCTAGCCACCATGGACTACAAAGACCATGA	HiFi DNA assembly
	Flag R	GGTACCAAGCTTGTTCTTGTCATCGTCATCCTTGT	HiFi DNA assembly
	Flag vector F	AACAAGCTTGGTACCGAATT	inverse PCR
	lambdaN vector F	ATGGGAGAGCAGAAACTGAT	inverse PCR
	pAc vector R	GGTGGCTAGCCCGATCCGGG	inverse PCR
	lambdaN F	ATCGGGCTAGCCACCATGGACGCACAAACACGACG	HiFi DNA assembly
	lambdaN R	TTTCTGCTCTCCCATAGCCATCAAGCTTGGAGGAG	HiFi DNA assembly
	BoxB oligo F	TCGACTAAGTCCAACTACTAAACTGGGCCCTGAAGAAGGGCCCATAT AGGGCCCTGAAGAAGGGCCCTATCGAGGATATTATC	BoxB cloning oligo
	BoxB oligo R	TCGAGATAATATCCTCGATAGGGCCCTTCTTCAGGGCCCTATATGGG CCCTTCTTCAGGGCCCAGTTTAGTAGTTGGACTTAG	BoxB cloning oligo
	pAcE sequencing primer	CGACAACCACTACCTGAGCA	sequencing primer
Nortern Blotting	*idefix*-piRNA	AAACTACTGGCAATCGTTTGGGAA	probe, *idefix*-piRNA
	tj-piRNA	GGTAATGGGAATGCACTTCTCTTGAA	probe, *tj*-piRNA
	repeat probe	GACCGAGTCAAGTTGAAAACCTATGGACCGAGTCAAGTTGAAAACCTA	probe, artificial piRNA
	miR-310	AAAGGCCGGGAAGTGTGCAATA	probe, miR-310
	U6 snRNA	GGGCCATGCTAATCTTCTCTGTA	probe, U6 snRNA
	EGFP F	ATGGTGAGCAAGGGCGAGGA	PCR F
	EGFP R	GTACAGCTCGTCCATGCCGA	PCR R
	GAPDH F	ATCGTCGAGGGTCTGATGAC	PCR F
	GAPDH R	ACCGAACTCGTTGTCGTACC	PCR R
RNAi	EGFP-si sense	GGCAAGCUGACCCUGAAGUTT	RNAi control
	EGFP-si antisense	ACUUCAGGGUCAGCUUGCCTT	RNAi control
	Luc-si sense	CGUACGCGGAAUACUUCGATT	RNAi control
	Luc-si antisense	UCGAAGUAUUCCGCGUACGTT	RNAi control
	Zuc-si sense	GCAUUGCCGUCAGCACUGUTT	siRNA for Zuc
	Zuc-si antisense	ACAGUGCUGACGGCAAUGCTT	siRNA for Zuc
	Armi-si sense	CAGCUUAAGGCGAAUCCAATT	siRNA for Armi
	Armi-si antisense	UUGGAUUCGCCUUAAGCUGTT	siRNA for Armi
	Yb-si sense	GGGGUACAUCAACCAAGCUTT	siRNA for Yb
	Yb-si antisense	AGCUUGGUUGAUGUACCCCTT	siRNA for Yb
RT-qPCR	rp49 F	CCGCTTCAAGGGACAGTATCTG	qPCR F
	rp49 R	ATCTCGCCGCAGTAAACGC	qPCR R
	mdg1 F	AACAGAAACGCCAGCAACAGC	qPCR F
	mdg1 R	CGTTCCCATGTCCGTTGTGAT	qPCR R
RNA	5′ overhang long strand	AAACAAAACAAAACAAAACAAAAUAGCACCGUAAAGACGC	
	5′ overhang short strand	GCGUCUUUACGGUGCU	
	3′ overhang long strand	GCGUCUUUACGGUGCUUAAAACAAAACAAAACAAAACAAA	
	3′ overhang short strand	AGCACCGUAAAGACGC	also used for ATPase assay and Gel shift assay

References

Anantharaman, V., Koonin, E. V, and Aravind, L. (2002). Comparative genomics and evolution of proteins involved in RNA metabolism. Nucleic Acids Res. *30*, 1427–1464.

Aravin, A., Gaidatzis, D., Pfeffer, S., Lagos-Quintana, M., Landgraf, P., Iovino, N., Morris, P., Brownstein, M.J., Kuramochi-Miyagawa, S., Nakano, T., et al. (2006). A novel class of small RNAs bind to MILI protein in mouse testes. Nature *442*, 203–207.

Aravin, A.A., Klenov, M.S., Vagin, V. V, Bantignies, F., Cavalli, G., and Gvozdev, V.A. (2004). Dissection of a natural RNA silencing process in the *Drosophila melanogaster* germ line. Mol. Cell. Biol. *24*, 6742–6750.

Bizebard, T., Ferlenghi, I., Iost, I., and Dreyfus, M. (2004). Studies on three *E. coli* DEAD-box helicases point to an unwinding mechanism different from that of model DNA helicases. Biochemistry *43*, 7857–7866.

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.

Chamieh, H., Ballut, L., Bonneau, F., and Le Hir, H. (2008). NMD factors UPF2 and UPF3 bridge UPF1 to the exon junction complex and stimulate its RNA helicase activity. Nat. Struct. Mol. Biol. *15*, 85–93.

Chung, W.J., Okamura, K., Martin, R., and Lai, E.C. (2008). Endogenous RNA interference provides a somatic defense against *Drosophila* transposons. Curr. Biol. *18*, 795–802.

Cook, H.A., Koppetsch, B.S., Wu, J., and Theurkauf, W.E. (2004). The *Drosophila* SDE3 homolog Armitage is required for *oskar* mRNA silencing and embryonic axis specification. Cell *116*, 817–829.

Cox, D.N., Chao, A., Baker, J., Chang, L., Qiao, D., and Lin, H. (1998). A novel class of evolutionarily conserved genes defined by *piwi* are essential for stem cell self-renewal. Genes Dev. *12*, 3715–3727.

Czaplinski, K., Weng, Y., Hagan, K.W., and Peltz, S.W. (1995). Purification and characterization of the Upf1 protein: a factor involved in translation and mRNA degradation. RNA *1*, 610–623.

Czaplinski, K., Ruiz-Echevarria, M.J., Paushkin, S. V, Han, X., Weng, Y., Perlick, H.A., Dietz, H.C., Ter-Avanesyan, M.D., and Peltz, S.W. (1998). The surveillance complex interacts with the translation release factors to enhance termination and degrade aberrant mRNAs. Genes Dev. *12*, 1665–1677.

Czech, B., and Hannon, G.J. (2011). Small RNA sorting: matchmaking for Argonautes. Nat. Rev. Genet. *12*, 19–31.

Czech, B., Malone, C.D., Zhou, R., Stark, A., Schlingeheyde, C., Dus, M., Perrimon, N., Kellis, M., Wohlschlegel, J.A., Sachidanandam, R., et al. (2008). An endogenous small interfering RNA pathway in *Drosophila*. Nature *453*, 798–802.

Czech, B., Munafò, M., Ciabrelli, F., Eastwood, E.L., Fabry, M.H., Kneuss, E., and Hannon, G.J. (2018). piRNA-guided genome defense: from biogenesis to silencing. Annu. Rev. Genet. *52*, 131–157.

De, I., Bessonov, S., Hofele, R., dos Santos, K., Will, C.L., Urlaub, H., Lührmann, R., and Pena, V. (2015). The RNA helicase Aquarius exhibits structural adaptations mediating its recruitment to spliceosomes. Nat. Struct. Mol. Biol. *22*, 138–144.

Dönertas, D., Sienski, G., and Brennecke, J. (2013). *Drosophila* Gtsf1 is an essential component of the Piwi-mediated transcriptional silencing complex. Genes Dev. *27*,

1693–1705.

Fairman-Williams, M.E., Guenther, U.P., and Jankowsky, E. (2010). SF1 and SF2 helicases: family matters. Curr. Opin. Struct. Biol. *20*, 313–324.

Flynn, R.A., Martin, L., Spitale, R.C., Do, B.T., Sagan, S.M., Zarnegar, B., Qu, K., Khavari, P.A., Quake, S.R., Sarnow, P., et al. (2015). Dissecting noncoding and pathogen RNA-protein interactomes. RNA *21*, 135–143.

Fukaya, T., and Tomari, Y. (2011). PABP is not essential for microRNA-mediated translational repression and deadenylation *in vitro*. EMBO J. *30*, 4998–5009.

Ge, D.T., Wang, W., Tipping, C., Gainetdinov, I., Weng, Z., and Zamore, P.D. (2019). The RNA-binding ATPase, Armitage, couples piRNA amplification in Nuage to phased piRNA production on mitochondria. Mol. Cell *74*, 982–995.e6.

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

Ghildiyal, M., Seitz, H., Horwich, M.D., Li, C., Du, T., Lee, S., Xu, J., Kittler, E.L.W., Zapp, M.L., Weng, Z., et al. (2008). Endogenous siRNAs derived from transposons and mRNAs in *Drosophila* somatic cells. Science *320*, 1077–1081.

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.

Gorbalenya, A.E., and Koonin, A.E. (1993). Helicases: amino acid sequence comparisons and structure-function relationships. Curr. Opin. Struct. Biol. *3*, 419–429.

Grivna, S.T., Beyret, E., Wang, Z., and Lin, H. (2006). A novel class of small RNAs in

mouse spermatogenic cells. Genes Dev. *20*, 1709–1714.

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.

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.

Han, B.W., Wang, W., Li, C., Weng, Z., and Zamore, P.D. (2015). piRNA-guided transposon cleavage initiates Zucchini-dependent, phased piRNA production. Science *348*, 817–821.

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.

Handler, D., Meixner, K., Pizka, M., Lauss, K., Schmied, C., Gruber, F., and Brennecke, J. (2013). The genetic makeup of the *Drosophila* piRNA pathway. Mol. Cell *50*, 762–777.

Hirakata, S., Ishizu, H., Fujita, A., Tomoe, Y., and Siomi, M.C. (2019). Requirements for multivalent Yb body assembly in transposon silencing in *Drosophila*. EMBO Rep. *20*, e47708.

Homolka, D., Pandey, R.R., Goriaux, C., Brasset, E., Vaury, C., Sachidanandam, R., Fauvarque, M.O., and Pillai, R.S. (2015). PIWI slicing and RNA elements in precursors instruct directional primary piRNA biogenesis. Cell Rep. *12*, 418–428.

Huang, H., Li, Y., Szulwach, K.E., Zhang, G., Jin, P., and Chen, D. (2014). AGO3 Slicer activity regulates mitochondria-nuage localization of Armitage and piRNA amplification. J. Cell Biol. *206*, 217–230.

Huang, X., Fejes Toth, K., and Aravin, A.A. (2017). piRNA Biogenesis in *Drosophila melanogaster*. Trends Genet. *33*, 882–894.

Ipsaro, J.J., Haase, A.D., Knott, S.R., Joshua-Tor, L., and Hannon, G.J. (2012). The structural biochemistry of Zucchini implicates it as a nuclease in piRNA biogenesis. Nature *491*, 279–283.

Ishizu, H., Iwasaki, Y.W., Hirakata, S., Ozaki, H., Iwasaki, W., Siomi, H., and Siomi, M.C. (2015). Somatic primary piRNA biogenesis driven by *cis*-acting RNA elements and *trans*-acting Yb. Cell Rep. *12*, 429–440.

Ishizu, H., Sumiyoshi, T., and Siomi, M.C. (2017). Use of the CRISPR-Cas9 system for genome editing in cultured *Drosophila* ovarian somatic cells. Methods *126*, 186–192.

Ishizu, H., Kinoshita, T., Hirakata, S., Komatsuzaki, C., and Siomi, M.C. (2019). Distinct and Collaborative Functions of Yb and Armitage in Transposon-Targeting piRNA Biogenesis. CellRep. *27*, 1822–1835.e8.

Iwasaki, Y.W., Siomi, M.C., and Siomi, H. (2015). PIWI-Interacting RNA: Its Biogenesis and Functions. Annu. Rev. Biochem. *84*, 405–433.

Iwasaki, Y.W., Murano, K., Ishizu, H., Shibuya, A., Iyoda, Y., Siomi, M.C., and Saito, K. (2016). Piwi modulates chromatin accessibility by regulating multiple factors including Histone H1 to repress transposons. Mol. Cell *63*, 408–419.

Iwasaki, Y.W., Ishino, K., and Siomi, H. (2017). Deep sequencing and high-throughput

analysis of PIWI-associated small RNAs. Methods *126*, 66–75.

Jankowsky, E. (2011). RNA helicases at work: binding and rearranging. Trends Biochem. Sci. *36*, 19–29.

Jankowsky, E., and Fairman, M.E. (2007). RNA helicases – one fold for many functions. Curr. Opin. Struct. Biol. *17*, 316–324.

Jankowsky, E., and Putnam, A. (2009). Duplex Unwinding with DEAD-Box Proteins. Methods Mol. Biol. *587*, 245–264.

Kainov, D.E., Tuma, R., and Mancini, E.J. (2006). Hexameric molecular motors: P4 packaging ATPase unravels the mechanism. Cell. Mol. Life Sci. *63*, 1095–1105.

Kawamura, Y., Saito, K., Kin, T., Ono, Y., Asai, K., Sunohara, T., Okada, T.N., Siomi, M.C., and Siomi, H. (2008). *Drosophila* endogenous small RNAs bind to Argonaute 2 in somatic cells. Nature *453*, 793–797.

Kim, V.N., Han, J., and Siomi, M.C. (2009). Biogenesis of small RNAs in animals. Nat. Rev. Mol. Cell Biol. *10*, 126–139.

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.

Lau, N.C., Seto, A.G., Kim, J., Kuramochi-Miyagawa, S., Nakano, T., Bartel, D.P., and Kingston, R.E. (2006). Characterization of the piRNA complex from rat testes. Science *313*, 363–367.

Lee, S.R., Pratt, G.A., Martinez, F.J., Yeo, G.W., and Lykke-Andersen, J. (2015). Target

discrimination in nonsense-mediated mRNA decay requires Upf1 ATPase activity. Mol. Cell *59*, 413–425.

Leeds, P., Peltz, S.W. Jacobson, A., and Culbertson, M.R. (1991). The product of the yeast *UPF1* gene is required for rapid turnover of mRNAs containing a premature translational termination codon. Genes Dev. *5*, 2303–2314.

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. (2009a). Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies. Cell *137*, 509–521.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and 1000 Genome Project Data Processing Subgroup (2009b). The sequence alignment/map format and SAMtools. Bioinformatics *25*, 2078–2079.

Lin, H., and Spradling, A.C. (1997). A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the *Drosophila* ovary. Development *124*, 2463–2476.

Meister, G. (2013). Argonaute proteins: functional insights and emerging roles. Nat. Rev. Genet. *14*, 447–459.

Mohn, F., Handler, D., and Brennecke, J. (2015). piRNA-guided slicing specifies transcripts for Zucchini-dependent, phased piRNA biogenesis. Science *348*, 812–817.

Munafò, M., Manelli, V., Falconio, F.A., Sawle, A., Kneuss, E., Eastwood, E.L., Seah, J.W.E., Czech, B., and Hannon, G.J. (2019). Daedalus and Gasz recruit Armitage to mitochondria, bringing piRNA precursors to the biogenesis machinery. Genes Dev. *33*, 844–856.

Murota, Y., Ishizu, H., Nakagawa, S., Iwasaki, Y.W., Shibata, S., Kamatani, M.K., Saito, K., Okano, H., Siomi, H., and Siomi, M.C. (2014). Yb integrates piRNA intermediates and processing factors into perinuclear bodies to enhance piRISC assembly. Cell Rep. *8*, 103–113.

Niki, Y., Yamaguchi, T., and Mahowald, A.P. (2006). Establishment of stable cell lines of *Drosophila* germ-line stem cells. Proc. Nat. Acad. Sci. USA *103*, 16325–16330.

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.

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.

Nishimasu, H., Ishizu, H., Saito, K., Fukuhara, S., Kamatani, M.K., Bonnefond, L., Matsumoto, N., Nishizawa, T., Nakanaga, K., Aoki, J., et al. (2012). Structure and function of Zucchini endoribonuclease in piRNA biogenesis. Nature *491*, 284–287.

Ohtani, H., Iwasaki, Y.W., Shibuya, A., Siomi, H., Siomi, M.C., and Saito, K. (2013). DmGTSF1 is necessary for Piwi–piRISC-mediated transcriptional transposon silencing in the *Drosophila* ovary. Genes Dev. *27*, 1656–1661.

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.

Ozata, D.M., Gainetdinov, I., Zoch, A., O'Carroll, D., and Zamore, P.D. (2019). PIWI-interacting RNAs: small RNAs with big functions. Nat. Rev. Genet. *20*, 89–108.

Pall, G.S., and Hamilton, A.J. (2008). Improved northern blot method for enhanced detection of small RNA. Nat. Protoc. *3*, 1077–1084.

Pandey, R.R., Tokuzawa, Y., Yang, Z., Hayashi, E., Ichisaka, T., Kajita, S., Asano, Y., Kunieda, T., Sachidanandam, R., Chuma, S., et al. (2013). Tudor domain containing 12 (TDRD12) is essential for secondary PIWI interacting RNA biogenesis in mice. Proc. Nat. Acad. Sci. USA *110*, 16492–16497.

Pandey, R.R., Homolka, D., Chen, K.M., Sachidanandam, R., Fauvarque, M.O., and Pillai, R.S. (2017). Recruitment of Armitage and Yb to a transcript triggers its phased processing into primary piRNAs in *Drosophila* ovaries. PLoS Genet. *13*, e1006956.

Preall, J.B., Czech, B., Guzzardo, P.M., Muerdter, F., and Hannon, G.J. (2012). *shutdown* is a component of the *Drosophila* piRNA biogenesis machinery. RNA *18*, 1446–1457.

Pyle, A.M. (2008). Translocation and unwinding mechanisms of RNA and DNA helicases. Annu. Rev. Biophys. *37*, 317–336.

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.

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.

Rogers, A.K., Situ, K., Perkins, E.M., and Toth, K.F. (2017). Zucchini-dependent piRNA processing is triggered by recruitment to the cytoplasmic processing machinery. Genes Dev. *31*, 1858–1869.

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.

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.

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.

Schmidt, A., Palumbo, G., Bozzetti, M.P., Tritto, P., Pimpinelli, S., and Schäfer, U. (1999). Genetic and molecular characterization of *sting*, a gene Involved in crystal formation and meiotic drive in the male germ line of *Drosophila melanogaster*. Genetics *151*, 749–760.

Schüpbach, T., and Wieschaus, E. (1991). Female sterile mutations on the second chromosome of *Drosophila melanogaster*. II. Mutations blocking oogenesis or altering egg morphology. Genetics *129*, 1119–1136.

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

Sienski, G., Batki, J., Senti, K., Dönertas, D., Tirian, L., Meixner, K., and Brennecke, J. (2015). Silencio/CG9754 connects the Piwi–piRNA complex to the cellular heterochromatin machinery. Genes Dev. *29*, 1–14.

Singleton, M.R., Dillingham, M.S., and Wigley, D.B. (2007). Structure and mechanism of

helicases and nucleic acid translocases. Annu. Rev. Biochem. *76*, 23–50.

Siomi, M.C., Sato, K., Pezic, D., and Aravin, A.A. (2011). PIWI-interacting small RNAs: the vanguard of genome defence. Nat. Rev. Mol. Cell Biol. *12*, 246–258.

Song, S.U., Kurkulos, M., Boeke, J.D., and Corces, V.G. (1997). Infection of the germ line by retroviral particles produced in the follicle cells: a possible mechanism for the mobilization of the *gypsy* retroelement of *Drosophila*. Development *124*, 2789–2798.

Sun, X., Perlick, H.A., Dietz, H.C., and Maquat, L.E. (1998). A mutated human homologue to yeast Upf1 protein has a dominant-negative effect on the decay of nonsense-containing mRNAs in mammalian cells. Proc. Nat. Acad. Sci. USA *95*, 10009–10014.

Szakmary, A., Reedy, M., Qi, H., and Lin, H. (2009). The Yb protein defines a novel organelle and regulates male germline stem cell self-renewal in *Drosophila melanogaster*. J. Cell Biol. *185*, 613–627.

Tanner, N.K., and Linder, P. (2001). DExD/H box RNA helicases: from generic motors to specific dissociation functions. Mol. Cell *8*, 251–262.

Tijerina, P., Bhaskaran, H., and Russell, R. (2006). Nonspecific binding to structured RNA and preferential unwinding of an exposed helix by the CYT-19 protein, a DEAD-box RNA chaperone. Proc. Nat. Acad. Sci. USA *103*, 16698–16703.

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.

Vagin, V. V, Yu, Y., Jankowska, A., Luo, Y., Wasik, K.A., Malone, C.D., Harrison, E.,

Rosebrock, A., Wakimoto, B.T., Fagegaltier, D., et al. (2013). Minotaur is critical for primary piRNA biogenesis. RNA *19*, 1064–1077.

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

Wang, L., Dou, K., Moon, S., Tan, F.J., and Zhang, Z.Z. (2018). Hijacking oogenesis enables massive propagation of LINE and retroviral transposons. Cell *174*, 1082–1094.e12.

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

Wilson, R.C., and Doudna, J.A. (2013). Molecular mechanisms of RNA interference. Annu. Rev. Biophys. *42*, 217–239.

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.

Yamashiro, H., and Siomi, M.C. (2018). PIWI-Interacting RNA in *Drosophila*: biogenesis, transposon regulation, and beyond. Chem. Rev. *118*, 4404–4421.

Yamashiro, H., Negishi, M., Kinoshita, T., Ishizu, H., Ohtani, H., and Siomi, M.C. (2019). Armitage determines Piwi–piRISC processing from precursor formation and quality control to inter-organelle translocation. EMBO Rep. e48769.

Yang, Q., and Jankowsky, E. (2006). The DEAD-box protein Ded1 unwinds RNA duplexes by a mode distinct from translocating helicases. Nat. Struct. Mol. Biol. *13*, 981–986.

Yang, Q., Del Campo, M., Lambowitz, A.M., and Jankowsky, E. (2007). DEAD-Box Proteins Unwind Duplexes by Local Strand Separation. Mol. Cell *28*, 253–263.

Yashiro, R., Murota, Y., Nishida, K.M., Yamashiro, H., Fujii, K., Ogai, A., Yamanaka, S., Negishi, L., Siomi, H., and Siomi, M.C. (2018). Piwi Nuclear Localization and Its Regulatory Mechanism in *Drosophila* Ovarian Somatic Cells. Cell Rep. *23*, 3647–3657.

Yu, Y., Gu, J., Jin, Y., Luo, Y., Preall, J.B., Ma, J., Czech, B., and Hannon, G.J. (2015). Panoramix enforces piRNA-dependent cotranscriptional silencing. Science *350*, 339–342.

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.

Acknowledgement

First of all, I would like to express my deepest appreciation to Dr. Mikiko C. Siomi for her constructive comments, helps in writings, and precious opportunity as a Ph.D. student in her laboratory. I appreciate Dr. Soichiro Yamanaka, Dr. Hirotsugu Ishizu, Dr. Kaoru Sato, Dr. Kazumichi M. Nishida and Dr. Ryo Murakami. They always technically advised me. I am also grateful to Dr. Hirotsugu Ishizu, Dr. Shigeki Hirakata and Chihiro Komatsuzaki for their considerable guidance and technical helps.

I thank Dr. Yukihide Tomari (The University of Tokyo) for kindly providing λN-HA-LacZ plasmids. I also acknowledge other members of Siomi laboratory for discussions and comments on my research. Without any of their help, my research would have not progress. I was financially supported by the Graduate Program for Leaders in Life Innovation.