Mechanisms of Transcriptional Silencing by the Nuclear Piwi Protein in *Drosophila* Germ Cells

Thesis by

Alicia Kathryn Rogers

In Partial Fulfillment of the Requirements for the

degree of

Doctor of Philosophy in Biology

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2018

(Defended March 29, 2018)

Alicia Kathryn Rogers
ORCID: 0000-0001-5525-6095

ACKNOWLEDGEMENTS

I am beyond excited to present my PhD thesis; however, I would be remiss not to acknowledge those who have given me support, love, and guidance along this journey. I would like to start by giving my deepest gratitude to my advisor, Katalin Fejes Tóth. She accepted me into her lab as an untrained biologist – didn't laugh *too* much when I kept holding the pipette incorrectly – and helped shape me into the scientist I am today. I am very grateful for all the time she spent in lab personally teaching me the techniques that have become the foundation of my molecular biology toolkit. Kata has always been encouraging and pushed me to be a critical thinker and creative trouble-shooter. I am incredibly fortunate to have had her as my advisor, and will strive to be as wonderful of a mentor for my future students. I would also like to thank my co-advisor, Alexei Aravin, for important discussions about my projects throughout my PhD and his continued support.

To my thesis committee members – thank you for your support and advice over the years. The chair of my committee, Bruce Hay, has always provided focused feedback and supportive words of encouragement. I am forever grateful to Barbara Wold for allowing nineteen-year-old me to spend a summer doing a bioinformatics internship in her lab. My time as an intern was when I decided I wanted to embark on a PhD at Caltech. I would like to thank her for always encouraging me and continuing to mentor me as a member of my committee. Without fail, Paul Sternberg has always provided encouragement, positive vibes, and support. I would like to thank him for always reassuring me when I doubted myself.

I thank my undergraduate advisors, Erich Baker and Sang-Chul Nam, and all of the faculty members in the Baylor School of Engineering and Computer Science for always supporting me during my time at Baylor University. Without their support, I would not be where I am today. I am especially grateful to Sang-Chul Nam for encouraging me to apply for the internship that led me to Caltech.

Also, I would also like to thank the individuals at the Caltech Core Facilities that made this work possible: Igor Antoshechkin for sequencing the many libraries I produced at the Millard and Muriel Jacobs Genetics and Genomics Laboratory, Andres Collazo and Steven Wilbert for the guidance that made the microscopy possible at the Caltech

Biological Imaging Facility, and the Caltech Cell Sorting Facility. I would also like to thank Justin Bois for taking time to discuss and help me with the colocalization analysis.

Thank you to all the former and current members of the AAA Team and KFT Lab: Ivan Olovnikov, Dubravka Pezic, Irina Meininger, Evelyn Stuwe, Adrien Le Thomas, Ariel Chen, Alex Webster, Ned Perkins, Sergei Manakov, Junho Hur, Hamada Masakazu, Ken Chan, Svetlana Ustyugova, Yicheng Luo, Chen-yin Ou, Jae Cho, Maria Ninova, Xiawei Huang, Riley Galton, April Jauhal, Sharan Prakash, Evita Varela, Zsofia Torok, and Qing Tang. Y'all have made the lab a fun and amazing place to work and I am grateful that I was able to work alongside each and every one of y'all. I am especially grateful to call some of y'all close friends.

I would like to thank my friends that helped celebrate every high, commiserate every low, and supported and encouraged me throughout grad school. There really are not enough words to express how appreciative I am for each of y'all being in my life. I would especially like to thank the original crew – Katie Schretter, David Hall, Sofi Quinodoz, Taso Dimitriadis, Larissa Charnsangavej, Naeem Husain, Harry Nunns, and Matteo Ronchi – my time at Caltech would not have been the same without y'all. I would also like to thank James Lee and Pei Shih for always being there – I truly appreciated our weekly post-Biolunch coffees and walks to the turtle pond.

No matter how the day in lab went, I was always greeted with unconditional love at home from my dogs, Gimbel and Uma. They will never know how much they helped get me through grad school or how much I appreciate them – thank you, ladies.

Scott Saunders has been a rock for me in the last stretch of grad school and I am truly thankful to call him my partner. Thank you for always being kind, supportive, and helping me laugh when I don't want to. I am so happy that grad school brought us together.

Finally, I will never be able to sufficiently express my gratitude to my family for all their support and love over the years. I want to start by thanking my late grandfather, Rudolph J Horsak, for being the one who first introduced me to the world of science and sparking this journey. I would also like to thank my sister, Julie Anne Lumley, for being the best little sister I could ever have and always supporting me. And last, but definitely not least, I would like to thank my mother, Lisa Kathryn Rogers, for continuously encouraging me in my dream to become a scientist and supporting me everyway she could. Her endless

love and confidence in me has been helped me push past every barrier that stood between me and this moment.

ABSTRACT

An important characteristic for life is the ability to persist – to reproduce and defend oneself against different stresses. The ability of a species to persist from one generation to the next heavily depends on the integrity of the genetic material being passed down, and thus organisms have developed strategies to ensure the integrity of their genomes remain in tact. In Metazoan germlines, piwi proteins and their associated piwi-interacting RNAs (piRNAs) provide a RNA-interference (RNAi) based defense system against the expression of transposable elements (TEs). TE expression is detrimental to an organism's genome – resulting in disruption of genes, double-stranded DNA breaks, and germ cell death – ultimately leading to the sterility of the organism. In *Drosophila melanogaster*, the piRNA pathway is composed of two cytoplasmic piwi clade Argonuate proteins, Aubergine (Aub) and Argonaute3 (Ago3), and a single nuclear piwi clade Argonuate protein, Piwi. The piwi clade Argonaute proteins bind piRNAs to form effector complexes that repress TE sequences.

The work presented in this thesis examines the role of the nuclear piwi clade Argonaute – Piwi – and the mechanisms by which Piwi accomplishes its functions. Chapter Two presents how Piwi/piRNA complexes identify genomic loci expressing TEs and direct the establishment of a repressive chromatin state to transcriptionally silence the loci. In Chapter Three, we explore the piRNA-induced transcriptional silencing (piRITS) pathway using a heterologous reporter based tethering system *in vivo*. We discuss how the recruitment of Piwi alone to a locus is not sufficient to induce repression, and establish a model for the connection bridging the Piwi/piRNA complex and effector silencing complex in the piRITS pathway. In Chapter Four, we employ our heterologous reporter based tethering system to explore the mechanism of piRNA precursor selection in the two cell types that make up *Drosophila* ovaries. We uncover a common mechanism of piRNA biogenesis in the two cell types and establish a unifying model of piRNA substrate selection. Finally, in Chapter Five, as essential step to understanding how Piwi achieves its nuclear function, we developed a heterologous two-hybrid system to identify factors that directly interact with Piwi. Overall, the work presented in this thesis provides a piece of the groundwork in understanding the mechanisms of transcriptional silencing of TEs in germ cells by Piwi. The work proposes that Piwi has dual functions in the nucleus. First, upon

target recognition, Piwi recruits the piRITS complex to target loci to accomplish Piwi-mediated transcriptional silencing by deposition of H3K9me3. Then, Piwi recruits the RDC complex to specifically bind H3K9me3 at target loci to allow piRNA-production from the locus.

PUBLISHED CONTENT AND CONTRIBUTIONS

1. Le Thomas, A. et al. (2013). "Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state". In: *Genes & Development* 27: 390-399. doi: 10.1101/gad.209841.112.

   AKR executed experiments and edited the manuscript.

2. Chen, A. C. et al. "Su(Var)2-10, a novel component of the piRNA-induced transcriptional silencing complex, controls co-transcriptional repression of piRNA targets". (Submitted).

   AKR generated flies and executed tethering experiments followed by RT-qPCR, ChIP-qPCR, or ChIP-seq.

3. Rogers, A. K. et al. (2017). "Zucchini-dependent piRNA processing is triggered by recruitment to the cytoplasmic processing machinery". In: *Genes & Development* 31: 1-12. doi: 10.1101/gad.303214.117.

   AKR generated flies, designed and executed the experiments, and edited the manuscript.

4. Rogers, A. K. et al. "Mago and Arp6 identified as novel *Drosophila* Piwi interactors using heterologous two-hybrid system". (Submitted)

   AKR generated constructs and cell line, designed and executed the experiments, and wrote the manuscript.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS AND/OR TABLES

Page

Chapter I

INTRODUCTION

The integrity of the genome that is passed from one generation to the next is of vital importance. Transposable elements (TEs) are selfish, mobile genetic elements that threaten the stability and integrity of the germline genome in all animals. If active, transposable elements can cause detrimental mutations, double stranded DNA breaks, and ultimately lead to sterility of the organism (Slotkin and Martienssen 2007). Organisms have evolved various mechanisms to defend the integrity of their genomes from active transposons. One mechanism that metazoans have developed for controlling the activity of transposable elements in the germline is the piRNA pathway (Siomi et al. 2011).

The piRNA pathway is a RNA silencing mechanism that acts specifically in the germline to repress the expression of transposable elements. The central components of the piRNA pathway are the Piwi-clade Argonautes and their associated small RNAs, termed piRNAs (Piwi-interacting RNAs). piRNAs are loaded into the Piwi-clade Argonautes to form an effector complex called pi-RISC (piRNA-induced silencing complex). The Argonaute protein provides the effector function, while the loaded small RNA provides the specific recognition of targets through sequence complementarity (Lin and Spradling 1997; Cox et al. 1998; Kuramochi-Miyagawa et al. 2001; Carmell et al. 2002; Houwing et al. 2007).

Piwi-clade Argonautes in *Drosophila melanogaster*

In *Drosophila melanogaster*, the model system used in our laboratory, there are three Piwi-clade Argonaute proteins – Piwi, Aubergine (Aub), and Argonaute3 (Ago3). The piwi proteins have a structure composed of four domains that is similar to other Argonaute proteins (Song et al. 2004; Wang et al. 2008). The mid domain anchors the 5' end of the small RNA while the PAZ domain anchors the 3' end of the small RNA. In addition, the PIWI domain provides the catalytic function of the piwi protein, allowing for cleavage of target transcripts. Piwi-clade Argonautes have a unique N-terminal domain that contains Arginine-rich motifs that can be symmetrically dimethylated. The symmetrically dimethylated Arginines (sDMAs) facilitate interactions between piwi proteins and Tudor-domain containing piRNA pathway components (Kirino et al. 2009; Nishida et al. 2009; Vagin et al. 2009; Liu et al. 2010). In *Drosophila*, the interactions between piwi proteins

and Tudor-domain containing proteins, which can act as scaffolds, generate distinct membraneless, perinuclear granules termed “nuage” (Brennecke et al. 2007a; Aravin et al. 2008; Malone et al. 2009).

Aub and Ago3 are expressed exclusively in the germ cells of *Drosophila* ovaries and localize within nuage in the cytoplasm. The endonucleolytic activity of Aub and Ago3 is essential for their function – Aub and Ago3 work together to post-transcriptionally repress transposable elements through cleavage of TE transcripts complementary to their associated piRNA (Vagin et al. 2006; Agger et al. 2007; Brennecke et al. 2007b; Gunawardane et al. 2007). In mammals and *Drosophila*, there is one Piwi-clade Argonaute that localizes to the nucleus. In *Drosophila*, this is the namesake of the Piwi-clade Argonautes – Piwi, which stands for P-element induced wimpy testis (Lin and Spradling 1997; Cox et al. 1998). Unlike Aub and Ago3 that are expressed exclusively in the nurse cells of the ovary, Piwi is expressed in the nurse cells as well as the somatic follicular cells that support the nurse cells (Cox et al. 2000; Harris and Macdonald 2001; Saito et al. 2006; Brennecke et al. 2007b; Gunawardane et al. 2007).When I began my thesis project, the role of Piwi in the nucleus was unknown.

The role of nuclear Piwi-piRNA complexes in regulating transposon loci

The chromatin structure of a locus’ environment plays a role in determining the transcriptional status of the locus. In *Drosophila melanogaster*, the properties of chromatin are largely defined by different histone modifications. Previous studies had shown that some TE loci are transcriptionally silenced by repressive chromatin marks (Bozzetti et al. 1995; Schmidt et al. 1999; Aravin et al. 2001; Aravin et al. 2004; Vagin et al. 2006). Furthermore, it had been indicated that loss of Piwi from the nucleus results in a change in histone modification marks over sites of TEs (Klenov et al. 2011; Pöyhönen et al. 2012). However, a genome wide analysis on the effect of Piwi on histone modifications and transcription of TEs had been lacking.

In Chapter Two of my thesis, we aimed to determine the function of nuclear Piwi in silencing transposable elements. We performed a genome-wide analysis of the effect of knocking down Piwi expression in the *Drosophila* germ cells. Transcriptionally silenced TE loci are marked with tri-methylated histone 3 lysine 9 (H3K9me3). Upon knockdown of Piwi, we found the repressive chromatin mark H3K9me3 at the loci of transposons is lost

and expression levels of transposons are increased. We showed that piRNAs are the essential guides for Piwi to recognize its targets, and that loading of Piwi with piRNAs is required for Piwi function whereas its endonucleolytic cleavage ability is not. Furthermore, we showed that Piwi interacts with RNA-associated factors in a RNA-dependent manner rather than with, as previously thought, chromatin-associated factors. Ultimately, we showed that Piwi functions in the nucleus to initiate Piwi-mediated transcriptional silencing of target loci by deposition of the repressive H3K9me3 chromatin mark (Le Thomas et al. 2013). Piwi-mediated transcriptional repression of TE loci by deposition of H3K9me3 was corroborated in two other independent studies (Sienski et al. 2012; Rozhkov et al. 2013).

Components of the piRNA pathway involved in Piwi-mediated transcriptional silencing

Piwi proteins are not capable of inducing chromatin modifications and transcriptional silencing by themselves, implying they must interact with other factors to successfully repress targets. As Piwi cannot "write" repressive chromatin marks at target loci itself, a fundamental question is what factors act downstream of Piwi-target recognition in establishing a repressive chromatin state. Previously, a handful of factors had been implicated in acting downstream of Piwi in the effector function of the piRITS (piRNA-induced transcriptional silencing) complex. One factor, Maelstrom (Mael), is a conserved factor implicated in TE silencing in both *Drosophila* and mice (Megosh et al. 2006; Aravin et al. 2009; Sienski et al. 2012). Similar to knockdown of Piwi, loss of Mael results in increased Pol II occupancy at TE transcriptional start sites (Sienski et al. 2012). However, knockdown of Mael results in only a minor reduction in H3K9me3 at TE loci, compared to loss of Piwi, implying that Mael is not involved in the establishment of H3K9me3 at Piwi-targeted loci (Sienski et al. 2012).

Two factors of unknown function, Panoramix (Panx) and Asterix (Arx), were found to be necessary for Piwi-mediated establishment of H3K9me3 at TE loci targets (Donertas et al. 2013; Muerdter et al. 2013; Sienski et al. 2015; Yu et al. 2015). piRNA-independent recruitment of Panx to a reporter construct is sufficient to transcriptionally silence the reporter loci by establishment of H3K9me3 (Sienski et al. 2015; Yu et al. 2015). Establishment of H3K9me3 by Panx requires the histone methyltransferase Eggless (Egg), also known as SetDB1 (Sienski et al. 2015; Yu et al. 2015). Egg "writes" the H3K9me3

mark and has been previously implicated in Piwi-mediated transcriptional silencing (Rangan et al. 2011; Sienski et al. 2015; Yu et al. 2015). While some of the players had been identified, the manner in which the piRITS complex factors associate with each other remains unclear.

In Chapter Three, we aimed to elucidate factors that act downstream of Piwi in establishing a repressive chromatin state. To study this, we utilized the RNA-binding domain of the lambdaN protein to piRNA-independently recruit factors to a heterologous reporter in *Drosophila* germ cells. Our laboratory had identified Su(Var)2-10 as a novel piRNA pathway component potentially involved in the piRITS complex (data unpublished). We found that recruitment of Piwi alone to the reporter transcript was not sufficient to induce a change in the transcriptional status of the locus – this was corroborated by two independent laboratories (Sienski et al. 2015; Yu et al. 2015). However, recruitment of Su(Var)2-10 or Panoramix to the reporter transcript resulted in a change in the chromatin state of the reporter locus. From the set of factors identified to be involved in the piRITs complex – Panoramix, Asterix, Su(Var)2-10, and the histone methyltransferase Eggless – we performed epistasis experiments to determine how these factors work together in Piwi-mediated transcriptional silencing. We found that repression by recruitment of Panoramix depended on Su(Var)2-10. Furthermore, we found that Su(Var)2-10 depends on Eggless to provide its repressive function.

piRNA biogenesis in *Drosophila melanogaster*

piRNA biogenesis is a multi-step process starting with transcription and early processing of precursor RNA in the nucleus, export, precursor cleavage followed by processing to mature piRNAs and loading into piwi proteins. There are two distinct mechansims of piRNA biogenesis in *Drosophila*: primary (also known as Zucchini-dependent) and secondary (also known as ping-pong-dependent) (Brennecke et al. 2007b; Gunawardane et al. 2007; Malone et al. 2009). Mature piRNAs are processed from longer transcripts, piRNA precursors (pre-piRNAs). The majority of pre-piRNAs are derived from piRNA clusters – long, discrete genomic regions with a high density of transposon sequences (Brennecke et al. 2007b; Gunawardane et al. 2007). In *Drosophila* ovaries, piRNA clusters are expressed in two cell types: germline cells and somatic follicular cells. The structure of piRNA clusters is largely dependent on the cell type it is expressed in.

Clusters that can be transcribed bi-directionally (dual stranded clusters) are found predominantly in the germline cells; whereas somatic clusters are largely uni-directionally transcribed (Brennecke et al. 2007a). Ultimately, the cluster structure dictates how a pre-piRNA is processed.

Pre-piRNAs derived from dual stranded clusters are predominantly processed by the ping-pong cycle. In the ping-pong cycle, processing of pre-piRNAs is initiated by the endonucleolytic (slicer) activity of cytoplasmic piwi proteins, Aub or Ago3, loaded with complementary piRNAs (Brennecke et al. 2007b; Gunawardane et al. 2007). The characteristic features of ping-pong biogenesis are the presence of complementary piRNA pairs whose 5' ends are separated by 10nt, and the generation of piRNAs with a bias for U at position 1 and for A at position 10 of their sequence (Wang et al. 2014).

Uni-strand cluster derived pre-piRNAs cannot undergo ping-pong processing. Instead, these pre-piRNAs must rely on the Zucchini-dependent biogenesis machinery for processing. In Zucchini-dependent biogenesis, endonucleolytic cleavage of pre-piRNAs by Zuc generates the 5', and in part the 3', ends of piRNAs. This process does not depend on the slicer activity of piwi proteins, nor the presence of complementary piRNAs (Malone et al. 2009). piRNAs generated through Zuc-dependent biogenesis show a bias for U at their 5' end but lack a bias for A at position 10.

Many of the cytoplasmic piRNA biogenesis factors, including Aub and Ago3 as well as numerous factors that were genetically identified to act in the pathway such as Vasa, Armitage (Armi), and Krimper, localize to membraneless perinuclear structures called nuage in germ cells and Yb bodies in the surrounding somatic follicular cells, leading to the assumption that these structures are the sites of piRNA biogenesis (Hay et al. 1988; Lasko and Ashburner 1990; Lim and Kai 2007; Kirino et al. 2009; Malone et al. 2009; Qi et al. 2011; Ipsaro et al. 2012; Nishimasu et al. 2012; Ohtani et al. 2013; Webster et al. 2015).

How piRNA precursors are identified by the biogenesis machinery remained unresolved. Other RNA processing events, such as miRNA or CRISPR RNA biogenesis, splicing, or polyadenylation, require specific sequence and/or structure motifs in precursor RNA that are recognized by the processing machinery (Zeng et al. 2005; Park et al. 2011; Barrangou and Marraffini 2014; Li and Patel 2016; Tsai and Joung 2016).

Despite efforts, common sequence or structural motifs that are shared by all piRNA precursors have not been identified (Muerdter et al. 2012). In lack of sequence motifs, two mutually non-exclusive models have been proposed to explain precursor selection in the germline, which we call the "persistent nuclear mark model" and the "selection by pre-existing piRNA model".

The "persistent nuclear mark model" proposes that specific proteins associate with piRNA precursors in the nucleus and remain associated with the transcripts in the cytoplasm, where they interact with the processing machinery. In the germline, genomic regions that give rise to the majority of piRNAs, called piRNA clusters, are marked by the histone 3 lysine 9 trimethylation (H3K9me3) mark and by the RDC complex. The RDC complex consists of the HP1 homolog, Rhino, and two additional proteins, Cutoff (Cuff) and Deadlock, and is required for transcription and early processing of cluster transcripts (Klattenhoff et al. 2009; Pane et al. 2011; Le Thomas et al. 2014; Mohn et al. 2014; Zhang et al. 2014; Chen et al. 2016). The well-conserved transcription export (TREX) complex is also enriched at cluster loci and binds piRNA precursors co-transcriptionally in an RDC-dependent fashion (Zhang et al. 2012; Hur and Chung 2016; Hur et al. 2016). Components of either the RDC or the TREX complex were proposed to constitute the mark that triggers processing in the cytoplasm. However, the nature of the mark and the mechanism by which it engages the processing machinery remain unclear. Evidence for RDC or TREX remaining associated with piRNA precursors after nuclear export is also lacking. Finally, tethering of Rhino to a single-stranded transgene (without its concomitant tethering to a complementary antisense transgene) does not trigger piRNA biogenesis (Zhang et al. 2014), arguing against the idea that binding of RDC by itself is sufficient to specify piRNA precursors.

The "selection by pre-existing piRNA" model suggests that precursors are specified in the cytoplasm by complementary piRNAs associated with the cytoplasmic piwi proteins. Existing piRNAs can target pre-piRNAs and induce their processing via the ping-pong cycle (Brennecke et al. 2007b; Gunawardane et al. 2007). piRNA-dependent cleavage not only generates complementary piRNAs (ping-pong partners) but also leads to phased processing of the precursor downstream of

the initial cleavage (Han et al. 2015; Homolka et al. 2015; Mohn et al. 2015; Senti et al. 2015; Wang et al. 2015). This process generates Piwi-loaded piRNAs. Mutation of the two piwi proteins involved in the ping-pong cycle, Aub and Ago3, greatly reduces Piwi-bound piRNAs in the germline, leading to the suggestion that Zuc-mediated piRNA biogenesis in germ cells is triggered by the ping-pong cycle (Han et al. 2015; Homolka et al. 2015; Mohn et al. 2015; Senti et al. 2015; Wang et al. 2015).

Follicular cells of the fly ovary lack ping-pong biogenesis factors as well as the RDC complex; yet produce piRNAs, suggesting an alternative precursor selection mechanism. In Chapter Four, our piRNA-independent tethering assay allowed us to address the longstanding question of how piRNA precursors are recognized and channeled into the piRNA processing machinery. We used our *in vivo* piRNA-independent tethering assay to recruit piRNA pathway components to the reporter transcript in both cell types to probe reporter-derived piRNA production. We found that reporter-derived piRNAs could be made in both cell types by recruitment of Piwi and other piRNA biogenesis factors. By tethering different piRNA pathway components, we were able to determine which factors are sufficient to trigger reporter-derived piRNAs. We were also able to determine which piRNA pathway components are necessary for production of piRNAs from the reporter. Furthermore, we showed that upon Piwi tethering, the reporter transcript is sequestered in nuage, and that compartmentalization of the transcript is necessary and sufficient to place it within spatial proximity to the processing machinery and trigger piRNA production. Overall, this work provided a single unified model for piRNA precursor selection in both cell types of *Drosophila* ovaries.

Cytoplasmic loading of Piwi and translocation of the Piwi-piRNA complex to the nucleus

In steady-state, Piwi is found in the nucleus of both cell types of *Drosophila* (Cox et al. 2000; Harris and Macdonald 2001; Saito et al. 2006; Brennecke et al. 2007b; Gunawardane et al. 2007), but unloaded Piwi resides within the cytoplasmic nuage prior to translocation to the nucleus (Le Thomas et al. 2013). In nuage, Piwi is loaded with piRNA and then the Piwi-piRNA complex is translocated to the nucleus where Piwi functions. To

accomplish its function, Piwi must engage in many interactions with a distinct set of, mostly unidentified, proteins within the cytoplasmic and nuclear compartments.

In nuage Piwi must interact with the piRNA biogenesis machinery to be loaded. Currently, piRNA pathway components involved in loading of piwi proteins are not fully known. Nuage localized Tudor-domain containing factors Shutdown, Vreteno, Brother-of-Yb, and Sister-of-Yb have been identified as necessary for loading Piwi and Aub with piRNAs (Olivieri et al. 2010; Handler et al. 2013). In addition, loading of Piwi within nuage seems to require Armitage and Zucchini, while Yb is required for Piwi-loading in the Yb bodies of somatic cells (Szakmary et al. 2009; Olivieri et al. 2010).

After being loaded into Piwi, the pre-piRNA is further processed into a mature piRNA. The Tudor-domain containing factor, Papi, trims a subset of Piwi-loaded pre-piRNA at their 3' end to their mature length – 25 nucleotides (Hayashi et al. 2016). Another factor implicated in trimming the 3' end of Piwi-loaded pre-piRNAs is Nibbler (Hayashi et al. 2016; Wang et al. 2016). After trimming, Hen1, a methytransferase, catalyzes the characteristic 2'-*O*-methylation found at the 3' terminal nucleotide of the Piwi-loaded piRNA (Horwich et al. 2007; Saito et al. 2007; Kawaoka et al. 2011). It remains to be seen if more factors involved in pre-piRNA processing into mature piRNA will be identified.

While many factors have been genetically identified as part of the piRNA pathway, it is unclear what role many of these factors serve in the pathway. Currently, factors involved in the translocation of loaded Piwi to the nucleus, if any, are unknown. As discussed in Chapters Two, Three, and Four of this thesis, there have been several factors implicated in piRNA biogenesis as well as in the piRITS complex. However, to date, only a few of the piRNA pathway factors have been biochemically dissected. Further, the ability to dissect the interactome is complicated by the complexity of the interactions. While studies have been carried out to determine whether the factors that interact with Piwi are involved in loading Piwi with piRNA in nuage or recruiting the silencing machinery to target loci, it is still unclear exactly whether Piwi engages in direct or indirect biochemical interactions with these factors. The ability to discern direct biochemical interactions with Piwi would provide important insights into understanding the sequence of events that must occur for Piwi to be loaded in the cytoplasm, translocated to the nucleus, and accomplish

its function in the nucleus. In Chapter Five, we aimed to identify novel interactions with Piwi and determine which interactions were most likely occurring directly. To probe the Piwi interactome, we developed a heterologous cell culture assay to tether lacI-mKate2-tagged Piwi to lacO arrays inserted in the genome of a mammalian cell line (U2OS). Tethering of Piwi to three distinct foci in the cell allowed us to screen fifty-five CFP-tagged factors to look for colocalization between the mKate2 and CFP signal, indicating a possible direct interaction. We found several new Piwi interactors and identified fourteen possible direct interactions.

References

Agger K, Cloos P, Christensen J, Pasini D, Rose S, Rappsilber J, Issaeva I, Canaani E, Salcini A, Helin K. 2007. UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. *Nature* **449**: 731-734.

Aravin AA, Klenov MS, Vagin VV, Bantignies F, Cavalli G, Gvozdev VA. 2004. Dissection of a natural RNA silencing process in the Drosophila melanogaster germ line. *Molecular and cellular biology* **24**: 6742-6750.

Aravin AA, Naumova NM, Tulin AV, Vagin VV, Rozovsky YM, Gvozdev VA. 2001. Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the D. melanogaster germline. *Curr Biol* **11**: 1017-1027.

Aravin AA, Sachidanandam R, Bourc'his D, Schaefer C, Pezic D, Toth KF, Bestor T, Hannon GJ. 2008. A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. *Molecular cell* **31**: 785-799.

Aravin AA, van der Heijden GW, Castaneda J, Vagin VV, Hannon GJ, Bortvin A. 2009. Cytoplasmic compartmentalization of the fetal piRNA pathway in mice. *PLoS Genet* **5**: e1000764.

Barrangou R, Marraffini LA. 2014. CRISPR-Cas Systems: Prokaryotes Upgrade to Adaptive Immunity. *Molecular cell* **54**: 234-244.

Bozzetti MP, Massari S, Finelli P, Meggio F, Pinna LA, Boldyreff B, Issinger OG, Palumbo G, Ciriaco C, Bonaccorsi S et al. 1995. The Ste locus, a component of the parasitic cry-Ste system of Drosophila melanogaster, encodes a protein that forms crystals in primary spermatocytes and mimics properties of the beta subunit of casein kinase 2. *Proc Natl Acad Sci U S A* **92**: 6067-6071.

Brennecke J, Aravin A, Stark A, Dus M, Kellis M, Sachidanandam R, Hannon G. 2007a. Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. *Cell* **128**: 1089-1103.

Brennecke J, Aravin AA, Stark A, Dus M, Kellis M, Sachidanandam R, Hannon GJ. 2007b. Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. *Cell* **128**: 1089-1103.

Carmell MA, Xuan Z, Zhang MQ, Hannon GJ. 2002. The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. *Genes Dev* **16**: 2733-2742.

Chen YC, Stuwe E, Luo Y, Ninova M, Le Thomas A, Rozhavskaya E, Li S, Vempati S, Laver JD, Patel DJ et al. 2016. Cutoff Suppresses RNA Polymerase II Termination to Ensure Expression of piRNA Precursors. *Molecular cell* **63**: 97-109.

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

Cox DN, Chao A, Lin H. 2000. piwi encodes a nucleoplasmic factor whose activity modulates the number and division rate of germline stem cells. *Development* **127**: 503-514.

Donertas D, Sienski G, Brennecke J. 2013. Drosophila Gtsf1 is an essential component of the Piwi-mediated transcriptional silencing complex. *Genes Dev* **27**: 1693-1705.

Gunawardane LS, Saito K, Nishida KM, Miyoshi K, Kawamura Y, Nagami T, Siomi H, Siomi MC. 2007. A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in Drosophila. *Science* **315**: 1587-1590.

Han BW, Wang W, Li CJ, Weng ZP, Zamore PD. 2015. piRNA-guided transposon cleavage initiates Zucchini-dependent, phased piRNA production. *Science* **348**: 817-821.

Handler D, Meixner K, Pizka M, Lauss K, Schmied C, Gruber FS, Brennecke J. 2013. The genetic makeup of the Drosophila piRNA pathway. *Molecular cell* **50**: 762-777.

Harris AN, Macdonald PM. 2001. Aubergine encodes a Drosophila polar granule component required for pole cell formation and related to eIF2C. *Development* **128**: 2823-2832.

Hay B, Jan LY, Jan YN. 1988. A protein component of Drosophila polar granules is encoded by vasa and has extensive sequence similarity to ATP-dependent helicases. *Cell* **55**: 577-587.

Hayashi R, Schnabl J, Handler D, Mohn F, Ameres SL, Brennecke J. 2016. Genetic and mechanistic diversity of piRNA 3'-end formation. *Nature* **539**: 588-592.

Homolka D, Pandey RR, Goriaux C, Brasset E, Vaury C, Sachidanandam R, Fauvarque MO, Pillai RS. 2015. PIWI Slicing and RNA Elements in Precursors Instruct Directional Primary piRNA Biogenesis. *Cell Rep* **12**: 418-428.

Horwich MD, Li C, Matranga C, Vagin V, Farley G, Wang P, Zamore PD. 2007. The Drosophila RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC. *Curr Biol* **17**: 1265-1272.

Houwing S, Kamminga LM, Berezikov E, Cronembold D, Girard A, van den Elst H, Filippov DV, Blaser H, Raz E, Moens CB et al. 2007. A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish. *Cell* **129**: 69-82.

Hur JK, Chung YD. 2016. A novel model of THO/TREX loading onto target RNAs in metazoan gene expression. *BMB reports* **49**: 355-356.

Hur JK, Luo Y, Moon S, Ninova M, Marinov GK, Chung YD, Aravin AA. 2016. Splicing-independent loading of TREX on nascent RNA is required for efficient expression of dual-strand piRNA clusters in Drosophila. *Genes Dev* **30**: 840-855.

Ipsaro JJ, Haase AD, Knott SR, Joshua-Tor L, Hannon GJ. 2012. The structural biochemistry of Zucchini implicates it as a nuclease in piRNA biogenesis. *Nature* **491**: 279-283.

Kawaoka S, Izumi N, Katsuma S, Tomari Y. 2011. 3' end formation of PIWI-interacting RNAs in vitro. *Molecular cell* **43**: 1015-1022.

Kirino Y, Kim N, de Planell-Saguer M, Khandros E, Chiorean S, Klein PS, Rigoutsos I, Jongens TA, Mourelatos Z. 2009. Arginine methylation of Piwi proteins catalysed by dPRMT5 is required for Ago3 and Aub stability. *Nature cell biology* **11**: 652-658.

Klattenhoff C, Xi H, Li C, Lee S, Xu J, Khurana JS, Zhang F, Schultz N, Koppetsch BS, Nowosielska A et al. 2009. The Drosophila HP1 homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters. *Cell* **138**: 1137-1149.

Klenov M, Sokolova O, Yakushev E, Stolyarenko A, Mikhaleva E, Lavrov S, Gvozdev V. 2011. Separation of stem cell maintenance and transposon silencing functions of Piwi protein. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 18760-18765.

Kuramochi-Miyagawa S, Kimura T, Yomogida K, Kuroiwa A, Tadokoro Y, Fujita Y, Sato M, Matsuda Y, Nakano T. 2001. Two mouse piwi-related genes: miwi and mili. *Mechanisms of development* **108**: 121-133.

Lasko PF, Ashburner M. 1990. Posterior localization of vasa protein correlates with, but is not sufficient for, pole cell development. *Genes Dev* **4**: 905-921.

Le Thomas A, Rogers AK, Webster A, Marinov GK, Liao SE, Perkins EM, Hur JK, Aravin AA, Toth KF. 2013. Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. *Genes Dev* **27**: 390-399.

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

Li SS, Patel DJ. 2016. Drosha and Dicer: Slicers cut from the same cloth. *Cell Res* **26**: 511-512.

Lim AK, Kai T. 2007. Unique germ-line organelle, nuage, functions to repress selfish genetic elements in Drosophila melanogaster. *Proc Natl Acad Sci U S A* **104**: 6714-6719.

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

Liu K, Chen C, Guo Y, Lam R, Bian C, Xu C, Zhao DY, Jin J, MacKenzie F, Pawson T et al. 2010. Structural basis for recognition of arginine methylated Piwi proteins by the extended Tudor domain. *Proc Natl Acad Sci U S A* **107**: 18398-18403.

Malone CD, Brennecke J, Dus M, Stark A, McCombie WR, Sachidanandam R, Hannon GJ. 2009. Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary. *Cell* **137**: 522-535.

Megosh HB, Cox DN, Campbell C, Lin H. 2006. The role of PIWI and the miRNA machinery in Drosophila germline determination. *Curr Biol* **16**: 1884-1894.

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

Mohn F, Sienski G, Handler D, Brennecke J. 2014. The rhino-deadlock-cutoff complex licenses noncanonical transcription of dual-strand piRNA clusters in Drosophila. *Cell* **157**: 1364-1379.

Muerdter F, Guzzardo PM, Gillis J, Luo Y, Yu Y, Chen C, Fekete R, Hannon GJ. 2013. A genome-wide RNAi screen draws a genetic framework for transposon control and primary piRNA biogenesis in Drosophila. *Molecular cell* **50**: 736-748.

Muerdter F, Olovnikov I, Molaro A, Rozhkov NV, Czech B, Gordon A, Hannon GJ, Aravin AA. 2012. Production of artificial piRNAs in flies and mice. *Rna* **18**: 42-52.

Nishida KM, Okada TN, Kawamura T, Mituyama T, Kawamura Y, Inagaki S, Huang H, Chen D, Kodama T, Siomi H et al. 2009. Functional involvement of Tudor and dPRMT5 in the piRNA processing pathway in Drosophila germlines. *EMBO J* **28**: 3820-3831.

Nishimasu H, Ishizu H, Saito K, Fukuhara S, Kamatani MK, 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 YW, Shibuya A, Siomi H, Siomi MC, 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, Sachidanandam R, Mechtler K, 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.

Pane A, Jiang P, Zhao DY, Singh M, Schupbach T. 2011. The Cutoff protein regulates piRNA cluster expression and piRNA production in the Drosophila germline. *EMBO J* **30**: 4601-4615.

Park JE, Heo I, Tian Y, Simanshu DK, Chang H, Jee D, Patel DJ, Kim VN. 2011. Dicer recognizes the 5 ' end of RNA for efficient and accurate processing. *Nature* **475**: 201-U107.

Pöyhönen M, de Vanssay A, Delmarre V, Hermant C, Todeschini A, Teysset L, Ronsseray S. 2012. Homology-dependent silencing by an exogenous sequence in the Drosophila germline. *G3 (Bethesda, Md)* **2**: 331-338.

Qi H, Watanabe T, Ku HY, Liu N, Zhong M, 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. *The Journal of biological chemistry* **286**: 3789-3797.

Rangan P, Malone CD, Navarro C, Newbold SP, Hayes PS, Sachidanandam R, Hannon GJ, Lehmann R. 2011. piRNA production requires heterochromatin formation in Drosophila. *Curr Biol* **21**: 1373-1379.

Rozhkov NV, Hammell M, Hannon GJ. 2013. Multiple roles for Piwi in silencing Drosophila transposons. *Genes Dev* **27**: 400-412.

Saito K, Nishida KM, Mori T, Kawamura Y, Miyoshi K, Nagami T, Siomi H, Siomi MC. 2006. Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the Drosophila genome. *Genes Dev* **20**: 2214-2222.

Saito K, Sakaguchi Y, Suzuki T, Suzuki T, Siomi H, Siomi MC. 2007. Pimet, the Drosophila homolog of HEN1, mediates 2'-O-methylation of Piwi- interacting RNAs at their 3' ends. *Genes Dev* **21**: 1603-1608.

Schmidt A, Palumbo G, Bozzetti MP, Tritto P, Pimpinelli S, Schafer 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.

Senti KA, Jurczak D, Sachidanandam R, Brennecke J. 2015. piRNA-guided slicing of transposon transcripts enforces their transcriptional silencing via specifying the nuclear piRNA repertoire. *Genes Dev* **29**: 1747-1762.

Sienski G, Batki J, Senti KA, Donertas D, Tirian L, Meixner K, Brennecke J. 2015. Silencio/CG9754 connects the Piwi-piRNA complex to the cellular heterochromatin machinery. *Genes Dev* **29**: 2258-2271.

Sienski G, Donertas D, Brennecke J. 2012. Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression. *Cell* **151**: 964-980.

Siomi M, Sato K, Pezic D, Aravin A. 2011. PIWI-interacting small RNAs: the vanguard of genome defence. *Nature reviews Molecular cell biology* **12**: 246-258.

Slotkin RK, Martienssen R. 2007. Transposable elements and the epigenetic regulation of the genome. *Nat Rev Genet* **8**: 272-285.

Song JJ, Smith SK, Hannon GJ, Joshua-Tor L. 2004. Crystal structure of Argonaute and its implications for RISC slicer activity. *Science* **305**: 1434-1437.

Szakmary A, Reedy M, Qi H, Lin H. 2009. The Yb protein defines a novel organelle and regulates male germline stem cell self-renewal in Drosophila melanogaster. *The Journal of cell biology* **185**: 613-627.

Tsai SQ, Joung JK. 2016. Defining and improving the genome-wide specificities of CRISPR-Cas9 nucleases. *Nat Rev Genet* **17**: 300-312.

Vagin VV, Sigova A, Li C, Seitz H, Gvozdev V, Zamore PD. 2006. A distinct small RNA pathway silences selfish genetic elements in the germline. *Science* **313**: 320-324.

Vagin VV, Wohlschlegel J, Qu J, Jonsson Z, Huang X, Chuma S, Girard A, Sachidanandam R, Hannon GJ, Aravin AA. 2009. Proteomic analysis of murine Piwi proteins reveals a role for arginine methylation in specifying interaction with Tudor family members. *Genes Dev* **23**: 1749-1762.

Wang H, Ma Z, Niu K, Xiao Y, Wu X, Pan C, Zhao Y, Wang K, Zhang Y, Liu N. 2016. Antagonistic roles of Nibbler and Hen1 in modulating piRNA 3' ends in Drosophila. *Development* **143**: 530-539.

Wang W, Han BW, Tipping C, Ge DT, Zhang Z, Weng ZP, Zamore PD. 2015. Slicing and Binding by Ago3 or Aub Trigger Piwi-Bound piRNA Production by Distinct Mechanisms. *Molecular cell* **59**: 819-830.

Wang W, Yoshikawa M, Han BW, Izumi N, Tomari Y, Weng Z, Zamore PD. 2014. The initial uridine of primary piRNAs does not create the tenth adenine that Is the hallmark of secondary piRNAs. *Molecular cell* **56**: 708-716.

Wang Y, Juranek S, Li H, Sheng G, Tuschl T, Patel DJ. 2008. Structure of an argonaute silencing complex with a seed-containing guide DNA and target RNA duplex. *Nature* **456**: 921-926.

Webster A, Li S, Hur JK, Wachsmuth M, Bois JS, Perkins EM, Patel DJ, Aravin AA. 2015. Aub and Ago3 Are Recruited to Nuage through Two Mechanisms to Form a Ping-Pong Complex Assembled by Krimper. *Molecular cell* **59**: 564-575.

Yu Y, Gu J, Jin Y, Luo Y, Preall JB, Ma J, Czech B, Hannon GJ. 2015. Panoramix enforces piRNA-dependent cotranscriptional silencing. *Science* **350**: 339-342.

Zeng Y, Yi R, Cullen BR. 2005. Recognition and cleavage of primary microRNA precursors by the nuclear processing enzyme Drosha. *Embo Journal* **24**: 138-148.

Zhang F, Wang J, Xu J, Zhang Z, Koppetsch BS, Schultz N, Vreven T, Meignin C, Davis I, Zamore PD et al. 2012. UAP56 couples piRNA clusters to the perinuclear transposon silencing machinery. *Cell* **151**: 871-884.

Zhang Z, Wang J, Schultz N, Zhang F, Parhad SS, Tu S, Vreven T, Zamore PD, Weng Z, Theurkauf WE. 2014. The HP1 homolog rhino anchors a nuclear complex that suppresses piRNA precursor splicing. *Cell* **157**: 1353-1363.

Chapter II

PIWI INDUCES PIRNA-GUIDED TRANSCRIPTIONAL SILENCING AND ESTABLISHMENT OF A REPRESSIVE CHROMATIN STATE

This work was first published as:

Le Thomas, A. et al. (2013). "Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state". In: *Genes & Development* 27: 390-399. doi: 10.1101/gad.209841.112.

Abstract

In the Metazoan germline, piwi proteins and associated piwi-interacting RNAs (piRNAs) provide a defense system against the expression of transposable elements. In the cytoplasm, piRNA sequences guide piwi complexes to destroy complementary transposon transcripts by endonucleolytic cleavage. However, some piwi family members are nuclear, raising the possibility of alternative pathways for piRNA-mediated regulation of gene expression. We found that *Drosophila* Piwi is recruited to chromatin, co-localizing with RNA polymerase II on polytene chromosomes. Knockdown of Piwi in the germline increases expression of transposable elements that are targeted by piRNAs, whereas protein-coding genes remain largely unaffected. De-repression of transposons upon Piwi depletion correlates with increased occupancy of RNA polymerase II on their promoters. Expression of piRNAs that target a reporter construct results in a decrease in Pol II occupancy and an increase in repressive H3K9me3 marks and HP1 on the reporter locus. Our results indicate that Piwi identifies targets complementary to the associated piRNA and induces transcriptional repression by establishing a repressive chromatin state when correct targets are found.

Introduction

Diverse small RNA pathways function in all kingdoms of life, from bacteria to higher eukaryotes. In eukaryotes, several classes of small RNA associate with members of the Argonaute protein family forming effector complexes in which the RNA provides target recognition by sequence complementarity, and the Argonaute provides the repressive function. Argonaute-small RNA complexes have been shown to regulate gene expression both transcriptionally and post-transcriptionally. Post-transcriptional repression involves either cleavage of target RNA through the endonucleolytic activity of Argonautes, or through sequestering targets into cytoplasmic ribonucleoprotein (RNP) granules (Hutvagner and Simard 2008).

The mechanism of transcriptional repression by small RNAs has been extensively studied in fission yeast and plants. Several studies showed that Argonaute-small RNA complexes induce transcriptional repression by tethering chromatin modifiers to target loci. In fission yeast, the effector complex containing the Argonaute and the bound small interfering (si)RNA associates with the histone H3K9 methyltransferase, Clr4, to install repressive H3K9-dimethyl marks at target sites (Nakayama et al. 2001; Maison and Almouzni 2004; Sugiyama et al. 2005; Grewal and Jia 2007). Methylation of histone H3 lysine 9 (H3K9) leads to recruitment of the heterochromatin protein 1 (HP1) homolog, Swi6, enhancing silencing and further promoting interaction with the Argonaute complex. The initial association of Ago with chromatin, however, requires active transcription (Ameyar-Zazoua et al. 2012; Keller et al. 2012). Plants also utilize siRNAs to establish repressive chromatin at repetitive regions. Contrary to yeast, heterochromatin in plants is marked by DNA methylation, although repression also depends on histone methylation by a Clr4 homolog (Soppe et al. 2002; Onodera et al. 2005). Although siRNA-mediated gene silencing is predominant on repetitive sequences, it is not limited to these sites. Constitutive expression of double-stranded RNA mapping to promoter regions results in production of corresponding siRNAs, *de novo* DNA methylation, and gene silencing (Mette et al. 2000; Matzke et al. 2004).

In metazoans, small RNA pathways are predominantly associated with post-transcriptional silencing. One class of small RNA, microRNA, regulates expression of a

large fraction of protein-coding genes (Friedman et al. 2009). In *Drosophila*, siRNAs silence expression of transposable elements in somatic cells (Chung et al. 2008; Ghildiyal et al. 2008) and target viral genes upon infection (Galiana-Arnoux et al. 2006; Wang et al. 2006; Zambon et al. 2006). Another class of small RNAs, Piwi-interacting RNAs (piRNAs), associates with the Piwi clade of Argonautes and acts to repress mobile genetic elements in the germline of both *Drosophila* and mammals (Siomi et al. 2011). Analysis of piRNA sequences in *Drosophila* revealed a very diverse population of small RNAs that primarily maps to transposon sequences and is derived from a number of heterochromatic loci called piRNA clusters, which serve as master regulators of transposon repression (Brennecke et al. 2007). Additionally, a small fraction of piRNAs seems to be processed from the mRNA of several host protein-coding genes (Robine et al. 2009; Saito et al. 2009). The *Drosophila* genome encodes three piwi proteins: Piwi, Aubergine (AUB) and Argonaute3 (AGO3). In the cytoplasm, AUB and AGO3 work together to repress transposons through cleavage of transposon transcripts, which are recognized through sequence complementarity by the associated piRNAs (Vagin et al. 2006; Agger et al. 2007; Brennecke et al. 2007; Gunawardane et al. 2007).

In both *Drosophila* and mammals, one member of the Piwi clade proteins localizes to the nucleus. Analogously to small RNA pathways in plants, the mouse piRNA pathway is required for *de novo* DNA methylation and silencing of transposable elements (TEs) (Carmell et al. 2007; Aravin et al. 2008; Kuramochi-Miyagawa et al. 2008), however, the exact mechanism of this process is unknown. In *Drosophila*, DNA methylation is absent; however several studies indicate that elimination of Piwi from the nucleus causes changes in histone marks on transposable elements (Klenov et al. 2011; Pöyhönen et al. 2012), yet a genome-wide analysis of Piwi's effect on chromatin marks and transcription is lacking.

Here we show that Piwi interacts with chromatin on polytene chromosomes in nurse cell nuclei. We found that Piwi exclusively represses loci that are targeted by piRNAs. We show that Piwi-mediated silencing occurs through repression of transcription and correlates with installment of repressive chromatin marks at targeted loci.

Results

To analyze the role of Piwi in the nucleus, we generated transgenic flies expressing a GFP-tagged Piwi protein (GFP-Piwi) under the control of its native regulatory region. GFP-Piwi was expressed in the ovary and testis in a pattern indistinguishable from the localization of native Piwi and was able to rescue the piwi null phenotype as indicated by ovarian morphology, fertility, transposon expression, and piRNA levels. GFP-Piwi was deposited into the mature egg and localized to the pole plasm; however, contrary to a previous observation (Brower-Toland et al. 2007) we did not detect Piwi expression outside of the ovary and testis in third instar larvae or adult flies. We also did not observe association of Piwi with polytene chromosomes in salivary gland cells of third instar larvae. In both follicular and germline cells of the *Drosophila* ovary, GFP-Piwi localized exclusively in the nucleus with slightly higher concentrations apparent in regions enriched for DAPI, indicating a possible interaction with chromatin. To gain further insight into Piwi localization in the nucleus, we took advantage of the fact that nurse cell chromosomes are polytenized and can be visualized on the *otu* mutant background (Mal'ceva et al. 1997). Analysis of polytene chromosomes from nurse cells demonstrated that GFP-Piwi associates with chromatin in a specific banding pattern. Interestingly, co-immunostaining showed that GFP-Piwi signal on polytene chromosomes generally overlaps with RNA polymerase II signal, which marks sites of active transcription (Fig. 1A).

In order to identify factors that might be responsible for targeting Piwi to chromatin, we immunoprecipitated Piwi complexes from the *Drosophila* ovary and analyzed Piwi interaction partners by mass spectrometry. We have purified Piwi complexes from ovaries of three different transgenic lines expressing GFP-Piwi, myc-Piwi, or FLAG-Piwi, using antibodies against each respective tag. As a control we used flies expressing free GFP in the ovary. We have identified over fifty factors that showed significant enrichment in all three Piwi purifications but were absent in the control. We were unable to identify chromatin-associated factors that directly associate with Piwi, but identified several RNA-binding proteins that associate with nascent transcripts such as splicing (Rm62, Pep, Ref1, Yps, CG9684, CG31368, CG5728, Mago) and nuclear export (Tho2,

Hpr1) factors (Fig 1B). Upon RNase A treatment prior to immunoprecipitation, the presence of most of these RNA-binding proteins in purified Piwi complexes was eliminated.

Piwi proteins are believed to find their targets through sequence complementarity of the associated piRNA. In fact, it has been proposed that lack of the associated piRNA leads to destabilization of piwi proteins and to Piwi's inability to localize to the nucleus (Saito et al. 2009; Haase et al. 2010; Olivieri et al. 2010; Handler et al. 2011; Ishizu et al. 2011). On the other hand, Piwi has been proposed to have functions that are independent of its role in transposon control by regulating stem-cell niche development (Cox et al. 1998; Klenov et al. 2011). To address the role of piRNA in translocation of Piwi into the nucleus and its function, we generated transgenic flies expressing a point-mutant Piwi, referenced as Piwi-YK, that is deficient in piRNA binding due to a substitution of two conserved amino acid residues (Y551L, K555E) in the 5' phosphate-binding pocket (Kiriakidou et al. 2007; Djuranovic et al. 2010). The Piwi-YK mutant was expressed in *Drosophila* follicular and germ cells at levels similar to wild-type Piwi, but was completely devoid of associated piRNA (Fig. 2A). In contrast to wild-type Piwi, Piwi-YK could be found in the cytoplasm, supporting the existence of a quality-control mechanism that prevents entrance of unloaded Piwi into the nucleus (Ishizu et al. 2011). Nevertheless, a significant amount of piRNA-deficient Piwi localized to the nucleus (Fig. 2B). Similar to wild-type Piwi, Piwi-YK seemed to associate with chromatin, as indicated by its localization in DAPI-stained regions of the nuclei, and this is consistent with fluorescence loss in photobleaching (FLIP) experiments that demonstrated reduced nuclear mobility compared to free diffusion (Fig. S1). Based on sterility and ovarian morphology, the Piwi-YK transgene was unable to rescue the piwi null phenotype, despite its nuclear localization (Fig 2C), indicating that while piRNA binding is not absolutely essential for stability and nuclear localization of Piwi, it is required for Piwi function.

To directly test the function of Piwi in the nucleus, we analyzed the effect of Piwi deficiency on gene expression and chromatin state on a genome-wide scale. Piwi mutant females have atrophic ovaries caused by Piwi deficiency in somatic follicular cells (Lin and Spradling 1997; Cox et al. 1998), which precludes analysis of Piwi function in null

mutants. Instead, we used RNAi knockdown to deplete Piwi in germ cells, while leaving it functionally intact in somatic follicular cells. The Piwi knockdown flies did not exhibit gross morphological defects in the ovary; however, they showed drastic reduction in GFP-Piwi expression in germ cells and were sterile (Fig. 3A-B). To analyze the effect of Piwi deficiency on the steady-state transcriptome, as well as the transcription machinery, we performed RNA-seq and Pol II ChIP-seq experiments from Piwi knockdown and control flies.

In agreement with previous observations that implicated Piwi in transposon repression (Saito et al. 2006; Aravin et al. 2007; Brennecke et al. 2007), we found that steady-state transcript levels of several transposable elements were increased upon Piwi knockdown in germ cells (Fig. 3C, 3D, S2). We found little to no change of RNA levels for transposons whose activity is restricted to follicular cells of the ovary, indicating that the observed changes are indeed due to loss of Piwi in the germline (Fig. S2). The analysis of Pol II ChIP-seq showed that Pol II occupancy increased over promoters of multiple transposable elements (Fig. 3D-F, Fig. S3). Indeed, the change in steady-state levels of transposon transcripts upon Piwi depletion correlated with changes of RNA polymerase II occupancy (Fig. 3F). This result demonstrates that Piwi ensures low level of transposon transcripts through a repressive effect on the transcription machinery.

To test if Piwi-mediated transcriptional repression is accompanied by a corresponding change in chromatin state, we used ChIP-seq to analyze the genome-wide distribution of the repressive H3K9me3 mark in the ovary upon Piwi knockdown. We identified 705 genomic loci at which the level of H3K9me3 significantly decreased. More than 90% of regions that show a decrease in the H3K9me3 mark upon Piwi depletion overlapped transposable element sequences, compared to 33% that is expected from random genome sampling (Fig. 4A). Furthermore, these regions tend to be located in the heterochromatic portions of the genome that are not assembled on the main chromosomes (Fig. 4B). Only twenty of the identified regions localized to the euchromatic parts of the genome. Of these, fifteen (75%) contained potentially active annotated copies of transposons. Taken together, our results indicate that Piwi is required for installment of repressive H3K9me3 chromatin marks on transposable element sequences of the genome.

While the vast majority of protein-coding host genes did not show significant changes in transcript level or RNA polymerase II occupancy upon Piwi knockdown, expression of a small set of protein coding genes (150 genes with a p-value <0.05), was significantly increased (Fig.5A, Supplementary Table 1). There are several possible explanations for Piwi's effect on host gene expression. First, failure in the piRNA pathway might cause upregulation of several genes that generate piRNAs in wild-type ovaries (Robine et al. 2009; Saito et al. 2009). However, the genes upregulated in Piwi-deficient ovaries were not enriched in piRNAs compared to other genes. Second, H3K9me3 marks installed on transposable element sequences in a Piwi-dependent manner might spread into neighboring host genes and repress their transcription, as was recently demonstrated in a follicular cell culture model (Sienski et al. 2012). To address this possibility we analyzed genomic positions of the genes whose expression was increased upon Piwi knockdown relative to genomic regions that showed a decrease in H3K9me3 marks. We found that upregulated genes did not show a significant change in the H3K9me3 mark (fig 5B, Fig. S4). Furthermore, the few genes located close to the regions that show decrease in H3K9me3 signal had unaltered expression levels upon Piwi knockdown. Next, we analyzed the functions of upregulated genes using GO term classifications and found significant enrichment for proteins involved in protein turnover and stress- and DNA-damage response pathways (Fig. 5C). Particularly, we found that thirty-one subunits of the proteasome complex were overexpressed. Therefore, our analysis indicates that upregulation of specific host genes is likely a secondary response to elevated transposon levels and genomic damage.

In contrast to host genes, transcripts of transposable elements are targeted by piRNA. To directly address the role of piRNA in Piwi-mediated transcriptional silencing, we took advantage of a fly strain that expresses artificial piRNAs against the *lacZ* gene, which are loaded into Piwi complexes and are able to repress lacZ reporter expression in germ cells (Fig. 6A)(Josse et al. 2007; Muerdter et al. 2012). Expression of piRNAs that are antisense to the reporter gene caused transcriptional silencing of the *lacZ* gene as measured by Pol II occupancy (Fig. 6B). Furthermore, we found that piRNA-induced silencing of the reporter gene was associated with an increase in the repressive H3K9me3 mark and HP1 occupancy, and a decrease in the abundance of the active H3K4me2/3 marks at the reporter

locus (Fig. 6C). This result is in good agreement with the genome-wide effect of Piwi depletion on distribution of the H3K9me3 mark and suggests that transcriptional silencing correlates with the establishment of a repressive chromatin structure and is mediated by piRNAs that match the target locus.

Discussion

Little is known about the function of nuclear piwi proteins. The nuclear piwi in mouse (Miwi2) affects DNA methylation of transposable elements (Carmell et al. 2007; Aravin et al. 2008; Kuramochi-Miyagawa et al. 2008). Several recent reports implicate *Drosophila* Piwi in regulation of chromatin marks on transposon sequences (Lin and Yin 2008; Klenov et al. 2011; Wang and Elgin 2011; Sienski et al. 2012). The mechanism of these processes is unknown in both organisms. Previously, Piwi was shown to associate with polytene chromosomes in salivary gland cells and to co-localize with HP1, a chromodomain protein that binds to heterochromatin and a few loci in euchromatin, suggesting that HP1 mediates Piwi's interaction with chromatin (Brower-Toland et al. 2007). However, recent results showed that the putative HP1 binding site on Piwi is dispensable for Piwi-mediated transposon silencing (Wang and Elgin 2011).

We did not detect Piwi expression outside of the ovary and testis, including in salivary gland cells, using a GFP-Piwi transgene expressed under native regulatory elements. We detected GFP-Piwi on polytene chromosomes in ovarian nurse cells that have a germline origin, however, it localizes in a pattern that largely does not overlap with HP1. FLIP experiments with GFP-Piwi indicated a relatively fast rate of fluorescence redistribution as compared to histone H2A (Fig. S1), implying a transient interaction of Piwi with chromatin. Our proteomic analysis of Piwi complexes isolated from *Drosophila* ovary did not identify chromatin-associated factors but revealed several RNA-binding proteins such as splicing and nuclear export factors that bind nascent RNA transcripts (Fig 1B). Importantly, the interaction of most of these RNA-binding proteins with Piwi was dependent on RNA, indicating that Piwi associates with nascent transcripts. As Piwi itself lacks DNA and RNA-binding domains (beyond the piRNA binding domain) it is likely that the recruitment of Piwi to chromatin is through interactions with

other RNA-binding proteins or through sequence-specific interactions between Piwi-bound piRNA and nascent transcripts.

Using specific Piwi knockdown in germ cells of the *Drosophila* ovary, we analyzed the effect of Piwi depletion on gene expression, the transcription machinery and H3K9me3 chromatin marks genome-wide. In agreement with previous results (Klenov et al. 2011) we found upregulation of several transposable elements upon Piwi knockdown (Fig. 3C). The transposable elements that did not change their expression upon germline knockdown of Piwi might be expressed exclusively in somatic follicular cells of the ovary, such as the gypsy retrotransposon. Alternatively, some elements present in the genome might not have transcriptionally active copies, or the cytoplasmic AUB/AGO3 proteins may efficiently silence them at the post-transcriptional level.

The increase in steady-state levels of RNA upon Piwi depletion strongly correlates with an increase in Pol II occupancy on the promoters of transposons (Fig 3D, 3F, Fig S2). This result suggests that Piwi represses transposon expression at the transcriptional level, although we cannot completely exclude the possibility of an additional post-transcriptional effect. It was shown previously that depletion or mutation of Piwi leads to depletion of the repressive H3K9me3 mark and an increase in the active H3K4me2/3 marks on several transposon sequences (Klenov et al. 2011; Wang and Elgin 2011). Our ChIP-seq data extends these results to a genome-wide scale, proving that transposons are indeed the sole targets of Piwi, and demonstrates that changes in histone marks directly correlate with transcriptional repression.

Piwi depletion in the germline does not affect expression of the majority of host genes, although a small fraction of genes change expression (Fig 5A). One possible mechanism of the effect Piwi has on host genes is spreading of repressive chromatin structure from transposon sequences to adjacent host genes. Indeed, such a spreading and resulting repression of host gene transcription was observed in an ovarian somatic cell (OSC) culture model (Sienski et al. 2012). However, we did not find significant changes in the H3K9me3 mark for genes that are upregulated upon germline depletion of Piwi, arguing against this mechanism playing a major role in host gene regulation. Instead we found that the majority of host genes whose expression is increased as a result of Piwi depletion participate in

protein turnover (e.g., proteasome subunits) and stress- and DNA-damage response pathways, indicating that they might be activated as a secondary response to cellular damage induced by transposon activation. The different effect of Piwi depletion on host gene expression in ovary and cultured cells might be explained by the fact that silencing of host genes due to transposon insertion would likely have a strong negative effect on the fitness of the organism, but could be tolerated in cultured cells. Accordingly, new transposon insertions that cause repression of adjacent host genes should be eliminated from the fly population, but can be detected in cultured cells. In agreement with this explanation, the majority of cases of repressive chromatin spreading in OSC were observed for new transposon insertions that are absent in the sequenced *Drosophila* genome. Indeed, it was shown that the vast majority of new transposon insertions is present at a low frequency in the *Drosophila* population likely due to strong negative selection (Petrov et al. 2003). Such selection was primarily attributed to the ability of TE sequences to cause recombination and genomic rearrangements. We propose that in addition to the effects on recombination, the selection against transposons can be driven by their negative impact on host gene expression in the germline linked to Piwi-mediated chromatin silencing.

How does Piwi discriminate its proper targets, transposons, from host genes? In the case of cytoplasmic Piwi proteins AUB and AGO3, recognition and post-transcriptional destruction of TE transcripts is guided by associated piRNAs. Our results indicate that piRNAs provide guidance for transcriptional silencing by the nuclear Piwi protein as well. First, in contrast to host genes that are not targeted by piRNAs, transposable element transcripts, which are regulated by Piwi, are recognized by antisense Piwi-bound piRNA (Brennecke et al. 2007). Second, a Piwi mutant that is unable to bind piRNA failed to rescue the piwi null mutation despite its ability to enter the nucleus. Finally, expression of artificial piRNAs that target a reporter locus induced transcriptional silencing associated with an increase in repressive H3K9me3 and HP1 chromatin marks and decrease in the active H3K4me2/3 marks (Fig. 6B-C). In contrast, the tethering of Piwi to chromatin in a piRNA-independent fashion by fusing Piwi with the lacI DNA binding domain that recognizes lacO sequences inserted upstream of a reporter gene did not lead to silencing of the reporter (data not shown). Together, our results demonstrate that piRNAs are the essential guides of Piwi to recognize its targets for transcriptional repression.

It is tempting to propose that, similar to Argonautes in fission yeast, *Drosophila* Piwi directly recruits the enzymatic machinery that establishes the repressive H3K9me3 mark on its targets. Establishment of repressive marks can lead to stable chromatin-based transcriptional silencing that does not require further association of Piwi with target loci. This model explains why we found that Piwi is relatively mobile in the nucleus, indicative of only a transient interaction with chromatin. The Piwi-mediated transcriptional silencing has an interesting parallel in *C. elegans* where the Piwi protein, PRG-1, and associated 21U RNAs are able to induce stable transgenerational repression that correlates with formation of silencing chromatin marks on target loci. Interestingly, PRG-1 and 21U RNAs are necessary only for initial establishment of silencing, while continuing repression depends on siRNA and the WAGO-group of Argonautes (Ashe et al. 2012; Bagijn et al. 2012; Buckley et al. 2012; Shirayama et al. 2012). Future studies should reveal the pathway that leads to transcriptional repression downstream of Piwi in *Drosophila* and the differences and similarities to other species.

Materials and Methods

Drosophila stocks:

Nanos-Gal4-VP16 (BL4937); UASp-shWhite (BL33623); UASp-shPiwi (BL 33724); Chr. I & II Balancer (BL7197) were purchased from the Bloomington Stock Center. GFP-Piwi expressing flies (see below) were backcrossed onto the *piwi1/piwi2* (available from Bloomington) background or the *otu7/otu11* (available from Bloomington) background respectively. LacZ reporter lines were a generous gift from S. Ronsseray.

Generation of Transgenic Fly Lines:

The GFP-Piwi, 3xFLAG-HA-Piwi and myc-Piwi constructs were generated using bacterial recombineering (Gene Bridges Counter Selection Kit) to insert the respective tag after the start codon of the Piwi genomic region cloned in BAC clone BACN04M10. The KpnI-XbaI genomic fragment that contains the Piwi gene and flanking sequences

was transferred to corresponding sites of the pCasper4 vector to create pCasper4/tagged-Piwi.

The pCasper4/GFP-Piwi construct was used to generate pCasper4/GFP-Piwi-YK with two point mutations, Y551I and K555E. Mutations were introduced by PCR, amplifying products corresponding to a 3.1kb upstream fragment and a 2.58kb downstream fragment. The upstream fragment included a unique XbaI site at the 5' end of the amplicon and overlapped 39bp with the downstream fragment, which included a unique BamHI site at its 3' end. The single XbaI-BamHI fragment was generated by overlap PCR with outside primers and cloned into corresponding sites of pCasper4/GFP-Piwi to replace the wild-type fragment. Transgenic flies were generated by P-element mediated transformation (BestGene).

IP of Piwi proteins and RNA gel of piRNA:

Dissected ovaries were lysed in lysis buffer (20mM HEPES, pH 7.0; 150mM KCl; 2.5mM MgCl; 0.5% Triton-X100; 0.5% Igepal; 100U/mL RNasin [Promega]; EDTA-free Complete Protease Inhibitor Cocktail [Roche]) and supernatant clarified by centrifugation. Supernatant was incubated with anti-eGFP polyclonal antibody (Covance) conjugated to Protein-G Dynabeads at 4°C. Beads were spiked with 5 pmol of synthesized 42 nucleotide RNA oligomer to assess purification efficiency, proteinase-K digested and phenol extracted. Isolated RNA was CIP treated, radiolabeled using PNK and γ-P32 labeled ATP, and run on a 15% urea-PAGE gel. Western blots of ovary lysate and anti-eGFP immunoprecipitates were obtained from 8% SDS-PAGE gels and probed with polyclonal rabbit anti-eGFP antibody.

Mass spectrometric analysis of Piwi interaction partners:

Lysis and clarification of ovary samples were performed as described above, using lysis buffer with reduced detergent (0.1% Triton-X100; 0.1% Igepal). Piwi proteins with Flag, Myc, or GFP tag were purified from *Drosophila* ovaries using corresponding antibodies covalently coupled to M-270 epoxy Dynabeads (Invitrogen) (Cristea et al. 2005). Immunoprecipitation of free GFP from GFP-expressing ovaries was used as a negative control. Immunoprecipitations were performed in the presence or absence of RNase A (100 ug/ml, 30 minutes at 25C). Piwi and co-purified interacting proteins were

resolved on NuPAGE Novex 4-12% Bis-Tris gels and stained with colloidal Coomassie blue. Gel fragments that contained protein bands were excised, in-gel trypsinized, and the peptides extracted, following the standard protocol of the Proteome Exploration Lab at Caltech. Peptide analyses were performed on an LTQ-FT Ultra (Thermo Fisher Scientific) equipped with a nanoelectrospray ion source (Thermo Fisher Scientific) connected to an EASY-nLC. Fractionation of peptides was performed on a 15 cm reversed phase analytical column (75 μm ID) in-house packed with 3 μm C18 beads (ReproSil-Pur C18-AQ medium; Dr. Maisch GmbH). Acquired spectra were searched against the *Drosophila melanogaster* proteome using the search engine Mascot (Matrix Science, v2.2.06) and protein inferences were performed using Scaffold (Proteome Software, v3).

Chromatin Immunoprecipitation, ChIP-seq and RNA-seq:

Chromatin immunoprecipitation was carried out using standard protocols (Moshkovich and Lei 2010). ChIP-seq and RNA-seq library construction and sequencing were carried out using standard protocols following the general principles described in (Johnson et al. 2007; Mortazavi et al. 2008) respectively. Data analysis was carried out using a combination of publicly available software tools and custom-written python scripts. Additional details regarding high-throughput data analysis are described in the supplementary methods section. For qPCR primers see Supplementary Table 2. GO term analysis of genes upregulated upon Piwi knockdown was performed using The **D**atabase for **A**nnotation, **V**isualization and **I**ntegrated **D**iscovery (**DAVID**) (Huang et al. 2009a; Huang et al. 2009b) and Flybase for additional assignment of GO terms.

Antibodies:

eGFP antibody (rabbit polyclonal serum, Covance) was affinity purified in our laboratory. Anti-myc (Millipore), anti-FLAG (Sigma), RNA polymerase II (Ab5408) and RNA polymerase II pSer5 (Ab5131) are commercially available.

Imaging of Ovaries:

Ovaries were fixed in 4% PFA in PBS for 20 minutes, permeabilized in 1% Triton-X100 in PBS, DAPI stained (Sigma-Aldrich), washed, and mounted in 50% glycerol/PBS.

Images were captured using an AxioImager microscope; an Apotome structured illumination system was used for optical sections (Carl Zeiss).

Fluorescence Loss In Photobleaching (FLIP):

FLIP time series were captured on an LSM510 confocal microscope equipped with a 40x/0.9 N.A. Imm Corr multi-immersion objective. Ovaries were dissected into halocarbon 700 oil (Sigma) and mounted under a 0.17 mm coverslip (Carl Zeiss) immediately before imaging. Two initial baseline images were captured, followed by 80 to 100 iterations consisting of 2 bleach iterations at 100% laser power (488nm or 543nm for GFP and RFP tagged proteins, respectively) followed by 2 images with reduced illumination intensity. FLIP series were cropped and median filtered with a 2 pixel radius to reduce noise using FIJI (Schindelin et al. 2012) and the "Rigid Body" function of the StackReg plugin (Thévenaz et al. 1998) to correct drift when needed. Using MATLAB software (The Mathworks, Natick, MA), images were background-subtracted and corrected for acquisition bleaching. A value representing the true loss of intensity relative to the initial pre-bleach images, where 0 indicates no change in intensity and 1 represents complete photobleaching, was calculated for each pixel, for each bleach/capture cycle, and plotted with a color lookup table and calibration bar. Scale bars and annotations were made in Inkscape (http://inkscape.org).

Preparation of Polytene squashes for immunofluorescence:

Flies carrying the GFP-Piwi BAC construct were backcrossed onto the *otu[7]* and *otu[11]* background. Progeny from the cross of the two lines were grown at 18°C. Stages 7-12 egg chambers were separated and transferred to a poly-lysine coated microscopic slide into PBST. From here the "smush" protocol was followed (Johansen et al. 2009) but PFA crosslinking was reduced to 10 minutes. Slides were imaged using an AxioImager microscope and a 63x oil immersion objective (Carl Zeiss, Thornwood, NY).

Acknowledgements

We are grateful to Evelyn Stuwe from the Aravin lab for purifying the GFP antibody, I. Antoshechkin of the Millard and Muriel Jacobs Genetics and Genomics Laboratory for sequencing, D. Trout, H. Amrhein and S. Upchurch for computational assistance, the Bloomington Stock Center for fly stocks and S. Hess, B. Graham and M Sweredoski from the Proteome Exploration Lab at the Beckmann Institute, Caltech for assistance with the mass spectrometry experiments. We thank members of the Aravin lab for critical comments on the manuscript. We thank Barbara Wold and members of the Wold lab for helpful discussions on ChIP protocols and analysis. AKR and EMP are supported by the Institutional Training Grant NIH/NRSA 5T32 GM07616 and EMP additionally by the Gordon Ross Medical Foundation. GKM is supported by The Beckman Foundation, the Donald Bren Endowment, and NIH grant U54 HG004576. This work was supported by grants from the National Institutes of Health (R01 GM097363, R00 HD057233 and DP2 OD007371A to AAA.), by the Searle Scholar Award to AAA and the Ellison Medical Foundation New Scholar in Aging Award to KFT.

References

Agger K, Cloos P, Christensen J, Pasini D, Rose S, Rappsilber J, Issaeva I, Canaani E, Salcini A, Helin K. 2007. UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. *Nature* **449**: 731-734.

Ameyar-Zazoua M, Rachez C, Souidi M, Robin P, Fritsch L, Young R, Morozova N, Fenouil R, Descostes N, Andrau J-C et al. 2012. Argonaute proteins couple chromatin silencing to alternative splicing. *Nature structural & molecular biology* **19**: 998-1004.

Aravin A, Hannon G, Brennecke J. 2007. The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. *Science (New York, NY)* **318**: 761-764.

Aravin AA, Sachidanandam R, Bourc'his D, Schaefer C, Pezic D, Toth KF, Bestor T, Hannon GJ. 2008. A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. *Mol Cell* **31**: 785-799.

Ashe A, Sapetschnig A, Weick E-M, Mitchell J, Bagijn M, Cording A, Doebley A-L, Goldstein L, Lehrbach N, Le Pen J et al. 2012. piRNAs can trigger a multigenerational epigenetic memory in the germline of C. elegans. *Cell* **150**: 88-99.

Bagijn M, Goldstein L, Sapetschnig A, Weick E-M, Bouasker S, Lehrbach N, Simard M, Miska E. 2012. Function, targets, and evolution of Caenorhabditis elegans piRNAs. *Science (New York, NY)* **337**: 574-578.

Brennecke J, Aravin A, Stark A, Dus M, Kellis M, Sachidanandam R, Hannon G. 2007. Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. *Cell* **128**: 1089-1103.

Brower-Toland B, Findley S, Jiang L, Liu L, Yin H, Dus M, Zhou P, Elgin S, Lin H. 2007. Drosophila PIWI associates with chromatin and interacts directly with HP1a. *Genes & development* **21**: 2300-2311.

Buckley B, Burkhart K, Gu S, Spracklin G, Kershner A, Fritz H, Kimble J, Fire A, Kennedy S. 2012. A nuclear Argonaute promotes multigenerational epigenetic inheritance and germline immortality. *Nature* **489**: 447-451.

Carmell M, Girard A, van de Kant H, Bourc'his D, Bestor T, de Rooij D, Hannon G. 2007. MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. *Developmental cell* **12**: 503-514.

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

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

Cristea IM, Williams R, Chait BT, Rout MP. 2005. Fluorescent proteins as proteomic probes. *Molecular & Cellular Proteomics* **4**: 1933-1941.

Djuranovic S, Zinchenko M, Hur J, Nahvi A, Brunelle J, Rogers E, Green R. 2010. Allosteric regulation of Argonaute proteins by miRNAs. *Nature structural & molecular biology* **17**: 144-150.

Friedman R, Farh K, Burge C, Bartel D. 2009. Most mammalian mRNAs are conserved targets of microRNAs. *Genome research* **19**: 92-105.

Galiana-Arnoux D, Dostert C, Schneemann A, Hoffmann J, Imler J-L. 2006. Essential function in vivo for Dicer-2 in host defense against RNA viruses in Drosophila. *Nature immunology* **7**: 590-597.

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

Grewal S, Jia S. 2007. Heterochromatin revisited. *Nature reviews Genetics* **8**: 35-46.

Gunawardane L, Saito K, Nishida K, Miyoshi K, Kawamura Y, Nagami T, Siomi H, Siomi M. 2007. A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in Drosophila. *Science (New York, NY)* **315**: 1587-1590.

Haase A, Fenoglio S, Muerdter F, Guzzardo P, Czech B, Pappin D, Chen C, Gordon A, Hannon G. 2010. Probing the initiation and effector phases of the somatic piRNA pathway in Drosophila. *Genes & development* **24**: 2499-2504.

Handler D, Olivieri D, Novatchkova M, Gruber F, Meixner K, Mechtler K, Stark A, Sachidanandam R, Brennecke J. 2011. A systematic analysis of Drosophila TUDOR domain-containing proteins identifies Vreteno and the Tdrd12 family as essential primary piRNA pathway factors. *The EMBO journal* **30**: 3977-3993.

Huang DW, Sherman B, Lempicki R. 2009a. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic acids research* **37**: 1-13.

Huang DW, Sherman B, Lempicki R. 2009b. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols* **4**: 44-57.

Hutvagner G, Simard M. 2008. Argonaute proteins: key players in RNA silencing. *Nature reviews Molecular cell biology* **9**: 22-32.

Ishizu H, Nagao A, Siomi H. 2011. Gatekeepers for Piwi-piRNA complexes to enter the nucleus. *Current opinion in genetics & development* **21**: 484-490.

Johansen K, Cai W, Deng H, Bao X, Zhang W, Girton J, Johansen J. 2009. Polytene chromosome squash methods for studying transcription and epigenetic chromatin modification in Drosophila using antibodies. *Methods (San Diego, Calif)* **48**: 387-397.

Johnson D, Mortazavi A, Myers R, Wold B. 2007. Genome-wide mapping of in vivo protein-DNA interactions. *Science (New York, NY)* **316**: 1497-1502.

Josse T, Teysset L, Todeschini A-L, Sidor C, Anxolabéhère D, Ronsseray S. 2007. Telomeric trans-silencing: an epigenetic repression combining RNA silencing and heterochromatin formation. *PLoS genetics* **3**: 1633-1643.

Keller C, Adaixo R, Stunnenberg R, Woolcock KJ, Hiller S, Buhler M. 2012. HP1(Swi6) mediates the recognition and destruction of heterochromatic RNA transcripts. *Mol Cell* **47**: 215-227.

Kiriakidou M, Tan G, Lamprinaki S, De Planell-Saguer M, Nelson P, Mourelatos Z. 2007. An mRNA m7G cap binding-like motif within human Ago2 represses translation. *Cell* **129**: 1141-1151.

Klenov M, Sokolova O, Yakushev E, Stolyarenko A, Mikhaleva E, Lavrov S, Gvozdev V. 2011. Separation of stem cell maintenance and transposon silencing functions of Piwi protein. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 18760-18765.

Kuramochi-Miyagawa S, Watanabe T, Gotoh K, Totoki Y, Toyoda A, Ikawa M, Asada N, Kojima K, Yamaguchi Y, Ijiri T et al. 2008. DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. *Genes & development* **22**: 908-917.

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

Lin H, Yin H. 2008. A novel epigenetic mechanism in Drosophila somatic cells mediated by Piwi and piRNAs. *Cold Spring Harbor symposia on quantitative biology* **73**: 273-281.

Maison C, Almouzni G. 2004. HP1 and the dynamics of heterochromatin maintenance. *Nature reviews Molecular cell biology* **5**: 296-304.

Mal'ceva N, Belyaeva E, King R, Zhimulev I. 1997. Nurse cell polytene chromosomes of Drosophila melanogaster otu mutants: morphological changes accompanying interallelic complementation and position effect variegation. *Developmental genetics* **20**: 163-174.

Matzke M, Aufsatz W, Kanno T, Daxinger L, Papp I, Mette M, Matzke A. 2004. Genetic analysis of RNA-mediated transcriptional gene silencing. *Biochimica et biophysica acta* **1677**: 129-141.

Mette M, Aufsatz W, van der Winden J, Matzke M, Matzke A. 2000. Transcriptional silencing and promoter methylation triggered by double-stranded RNA. *EMBO J* **19**: 5194-5201.

Mortazavi A, Williams B, McCue K, Schaeffer L, Wold B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature methods* **5**: 621-628.

Moshkovich N, Lei E. 2010. HP1 recruitment in the absence of argonaute proteins in Drosophila. *PLoS genetics* **6**.

Muerdter F, Olovnikov I, Molaro A, Rozhkov N, Czech B, Gordon A, Hannon G, Aravin A. 2012. Production of artificial piRNAs in flies and mice. *RNA (New York, NY)* **18**: 42-52.

Nakayama J, Rice J, Strahl B, Allis C, Grewal S. 2001. Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. *Science (New York, NY)* **292**: 110-113.

Olivieri D, Sykora M, Sachidanandam R, Mechtler K, 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.

Onodera Y, Haag J, Ream T, Costa Nunes P, Pontes O, Pikaard C. 2005. Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. *Cell* **120**: 613-622.

Petrov DA, Aminetzach YT, Davis JC, Bensasson D, Hirsh AE. 2003. Size matters: non-LTR retrotransposable elements and ectopic recombination in Drosophila. *Mol Biol Evol* **20**: 880-892.

Pöyhönen M, de Vanssay A, Delmarre V, Hermant C, Todeschini A, Teysset L, Ronsseray S. 2012. Homology-dependent silencing by an exogenous sequence in the Drosophila germline. *G3 (Bethesda, Md)* **2**: 331-338.

Robine N, Lau N, Balla S, Jin Z, Okamura K, Kuramochi-Miyagawa S, Blower M, Lai E. 2009. A broadly conserved pathway generates 3'UTR-directed primary piRNAs. *Current biology : CB* **19**: 2066-2076.

Saito K, Inagaki S, Mituyama T, Kawamura Y, Ono Y, Sakota E, Kotani H, Asai K, Siomi H, Siomi M. 2009. A regulatory circuit for piwi by the large Maf gene traffic jam in Drosophila. *Nature* **461**: 1296-1299.

Saito K, Nishida K, Mori T, Kawamura Y, Miyoshi K, Nagami T, Siomi H, Siomi M. 2006. Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the Drosophila genome. *Genes & development* **20**: 2214-2222.

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B et al. 2012. Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**: 676-682.

Shirayama M, Seth M, Lee H-C, Gu W, Ishidate T, Conte D, Mello C. 2012. piRNAs initiate an epigenetic memory of nonself RNA in the C. elegans germline. *Cell* **150**: 65-77.

Sienski G, Dönertas D, Brennecke J. 2012. Transcriptional silencing of transposons by piwi and maelstrom and its impact on chromatin state and gene expression. *Cell* **151**: 964-980.

Siomi M, Sato K, Pezic D, Aravin A. 2011. PIWI-interacting small RNAs: the vanguard of genome defence. *Nature reviews Molecular cell biology* **12**: 246-258.

Soppe WJ, Jasencakova Z, Houben A, Kakutani T, Meister A, Huang M, Jacobsen S, Schubert I, Fransz P. 2002. DNA methylation controls histone H3 lysine 9 methylation and heterochromatin assembly in Arabidopsis. *EMBO J* **21**: 6549-6559.

Sugiyama T, Cam H, Verdel A, Moazed D, Grewal S. 2005. RNA-dependent RNA polymerase is an essential component of a self-enforcing loop coupling heterochromatin assembly to siRNA production. *Proceedings of the National Academy of Sciences of the United States of America* **102**: 152-157.

Thévenaz P, Ruttimann U, Unser M. 1998. A pyramid approach to subpixel registration based on intensity. *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society* **7**: 27-41.

Vagin V, Sigova A, Li C, Seitz H, Gvozdev V, Zamore P. 2006. A distinct small RNA pathway silences selfish genetic elements in the germline. *Science (New York, NY)* **313**: 320-324.

Wang S, Elgin S. 2011. Drosophila Piwi functions downstream of piRNA production mediating a chromatin-based transposon silencing mechanism in female germ line. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 21164-21169.

Wang X-H, Aliyari R, Li W-X, Li H-W, Kim K, Carthew R, Atkinson P, Ding S-W. 2006. RNA Interference Directs Innate Immunity Against Viruses in Adult Drosophila. *Science* **312**: 452-454.

Zambon RA, Vakharia VN, Wu LP. 2006. RNAi is an antiviral immune response against a dsRNA virus in Drosophila melanogaster. *Cellular microbiology* **8**: 880-889.

Figures and Figure Legends

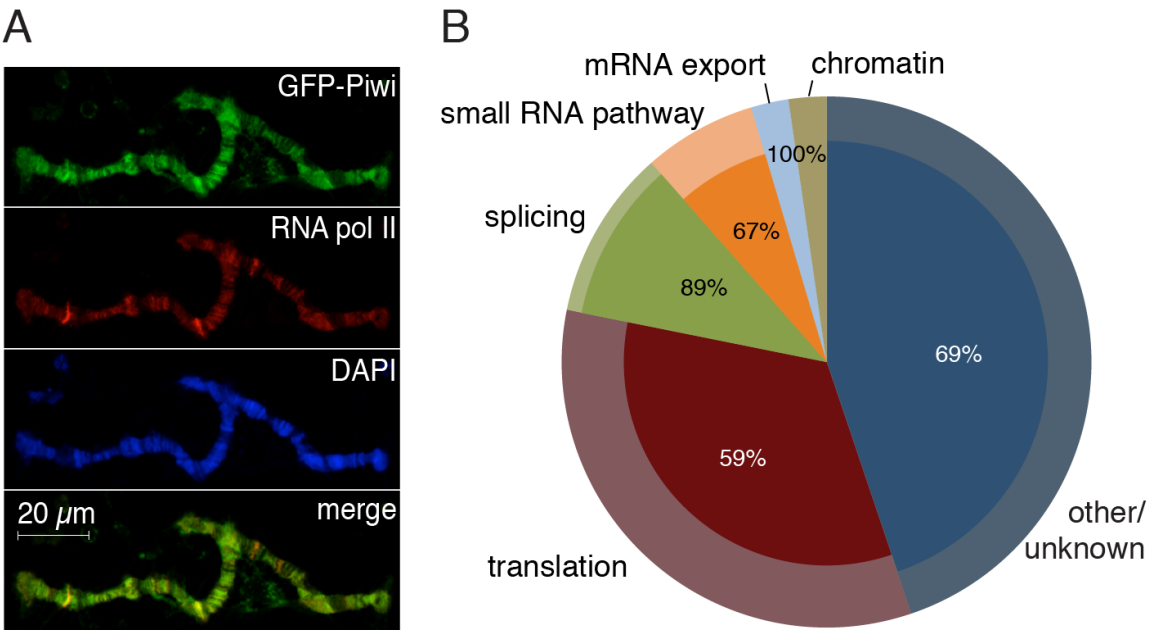

Figure 1. Piwi associates with chromatin and nuclear transcripts.

(**A**) Polytene chromosomes from *Drosophila* nurse cells expressing GFP-Piwi on the *otu[7]/otu[11]* background. Piwi pattern on chromosomes correlates with RNA polymerase II staining. (**B**) Mass spectrometry analysis of Piwi interaction partners. Piwi complexes were precipitated in the presence and absence of RNase A. Outer circle represents classification of Piwi-associated proteins based on GO term analysis. Inner pies represent the fraction of each group whose association with Piwi depends on RNA (% indicated). Note that chromatin, splice, and mRNA export factors are virtually absent after RNase A treatment.

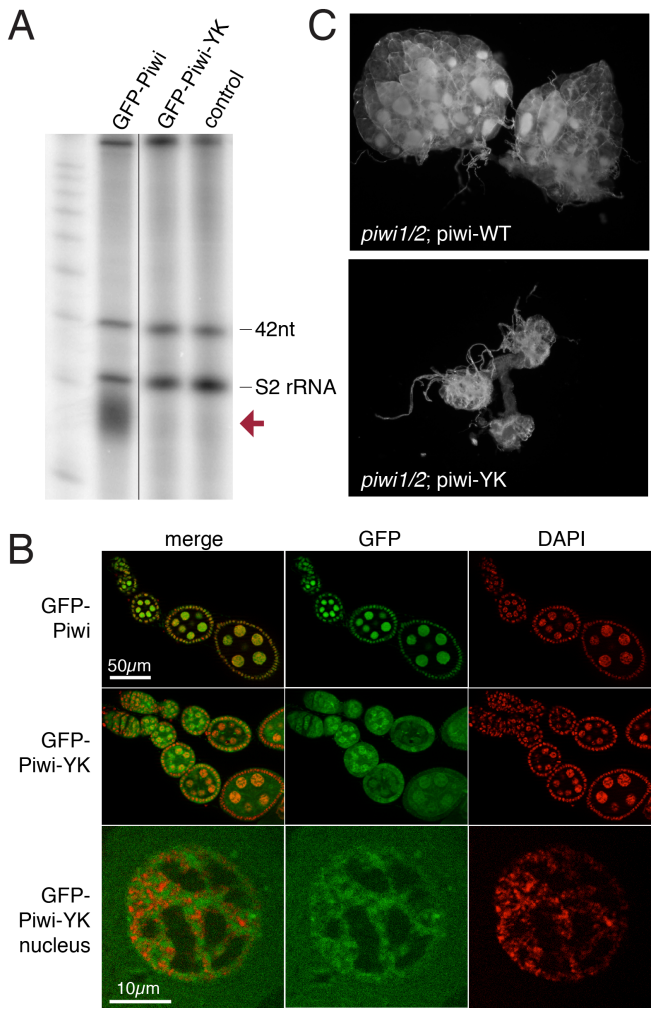

Figure 2. Piwi function, but not its nuclear localization, requires piRNA association.

(**A**) The Piwi-YK mutant does not associate with piRNA. Immunoprecipitation of Piwi-piRNA complexes was performed with GFP antibody on ovaries from GFP-Piwi and GFP-Piwi-YK transgenic flies and a control strain. Small RNAs were isolated, 5'-labeled, and resolved on a denaturing gel. Same amount of 42 nt RNA oligonucleotides were spiked into all samples prior to RNA isolation to control for loss of RNA during isolation and labeling. piRNAs (red arrow) are absent in the Piwi-YK complex. (**B**) GFP-Piwi-YK is present in the nuclei of nurse cells and associates with chromatin (DAPI-stained areas). (**C**) Piwi-YK mutant does not rescue the morphological changes caused by the piwi null mutation. Dark field images of ovaries where either the wild-type piwi or the piwi-YK transgene has been backcrossed onto the piwi null background.

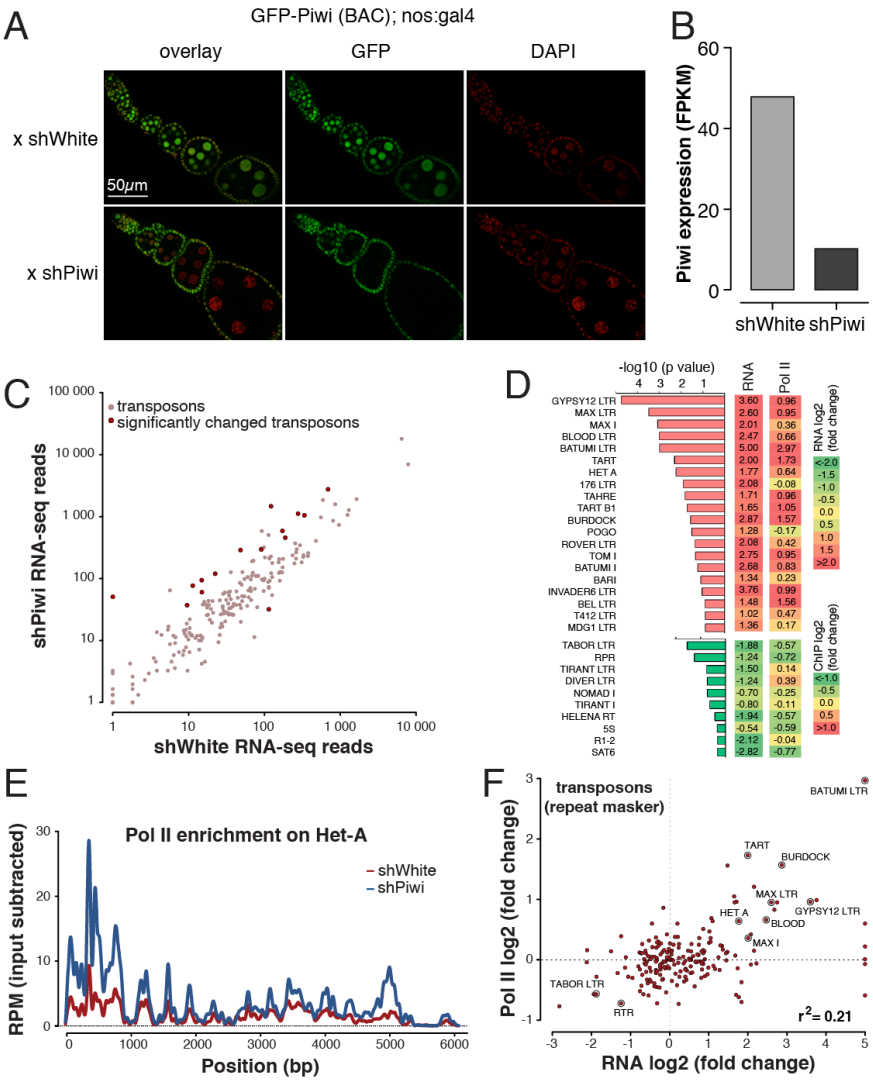

Figure 3. Piwi transcriptionally represses transposable elements.

(**A**) Piwi knockdown is efficient and specific to ovarian germ cells, as indicated by GFP-Piwi localization. GFP-Piwi; Nanos-Gal4-VP16 flies were crossed to control shRNA (shWhite) or shPiwi lines. Piwi is specifically depleted in germ cells and not in follicular cells consistent with expression of the Nanos-Gal4-VP16 driver. **(B)** Piwi expression as measured by RNA-seq in the Piwi knockdown and control lines. Note that Piwi expression is unaffected in follicular cells leading to relatively weak apparent knockdown in RNA-seq libraries from whole ovaries. (**C**) Effect of Piwi knockdown on the expression of transposable elements. Two biological replicate RNA-seq experiments were carried out and differential expression was assessed using DESeq. Transposons that show significant change ($p < 0.05$) are indicated by dark red circles. Out of 217 individual RepeatMasker-

annotated transposable elements, 15 show significant increase in expression upon Piwi knockdown. (**D**) The change in the levels of transposable element transcripts and RNA polymerase II occupancy on their promoters upon Piwi knockdown. 20 upregulated and 10 downregulated transposons with the most significant changes in expression level are shown. Note the low statistical significance for downregulated transposons. For a complete list of transposons see Fig S2. (**E**) RNA polymerase II signal over the Het-A retrotransposon in control flies (shWhite, red) and upon Piwi knockdown (shPiwi, blue). (**F**) Increased abundance of transposon transcripts upon Piwi depletion correlates with increased Pol II occupancy over their promoters (r^2 = 0.21). Note that the majority of elements do not show significant change in either RNA abundance or Pol II occupancy.

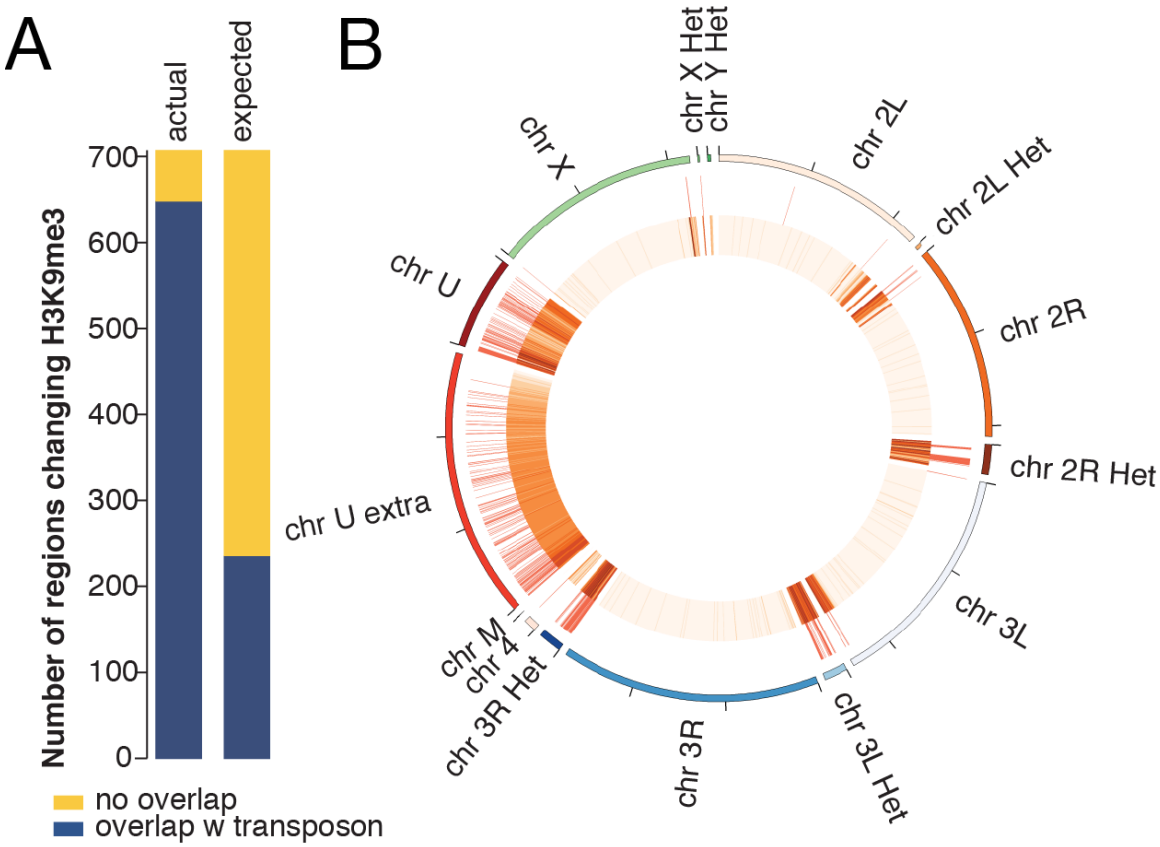

Figure 4. Piwi-induced transcriptional repression correlates with establishment of a repressive chromatin state.

(**A**) Overlap between genomic regions of H3K9me3 depletion upon Piwi knockdown and transposable elements. Two replicates of H3K9me3 ChIP-seq experiments were carried out on control and Piwi-depleted ovaries and enriched regions were identified using DESeq (see methods for details). A total of 705 regions show significant ($p < 0.05$) decrease in H3K9me3 occupancy upon Piwi knockdown, while only 30 regions showed a similarly significant increase. Out of the 705 regions that show a decrease in H3K9me3 mark upon Piwi knockdown, 91% (646) overlap with transposable element sequences compared to 33% expected from random genome sampling. (**B**) Genomic positions of H3K9me3-depleted regions upon Piwi depletion (outer circle) and RepeatMasker-annotated transposons (inner circle). Note that almost all regions are localized in heterochromatic and repeat-rich portions of the genome ('Het', chrU, and chrUExtra chromosomes).

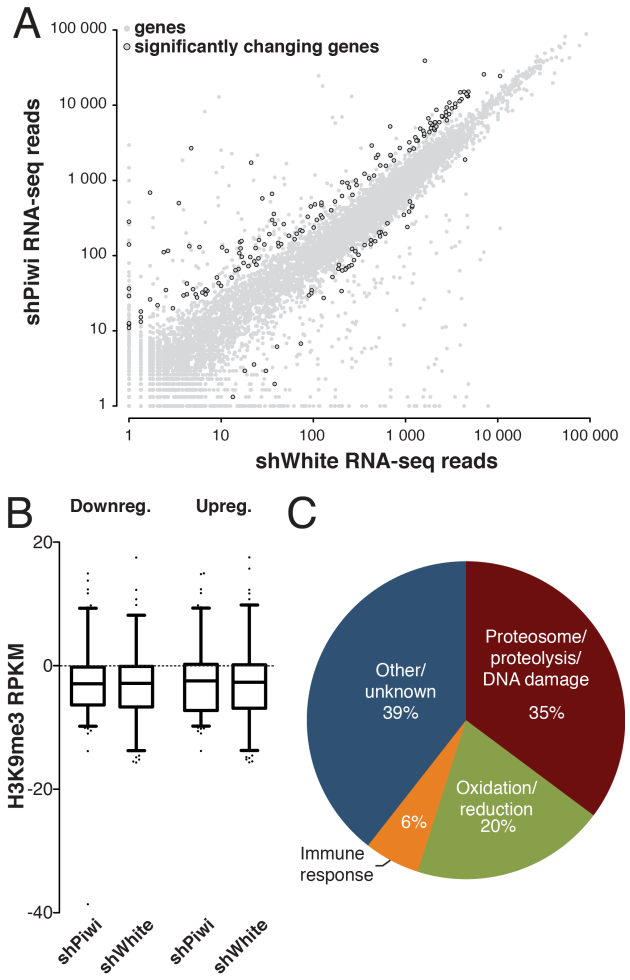

Figure 5. Piwi does not directly repress protein-coding genes.

(**A**) Effect of Piwi knockdown on the expression of genes. Two replicate RNA-seq experiments were carried out and differential expression was assessed using DESeq. Genes that show significant change ($p < 0.05$) are indicated by black circles. The vast majority of genes do not change significantly upon germline Piwi knockdown (shPiwi) compared to control (shWhite). (**B**) H3K9me3 mark density does not change over genes that show significant change in expression upon Piwi knockdown (see fig. 3C). Up- and downregulated genes are plotted separately. Signal indicated is after background subtraction. (**C**) Functional analysis of upregulated genes by DAVID reveals activation of the protein degradation and DNA-damage response pathways. Percentages of all upregulated genes are indicated.

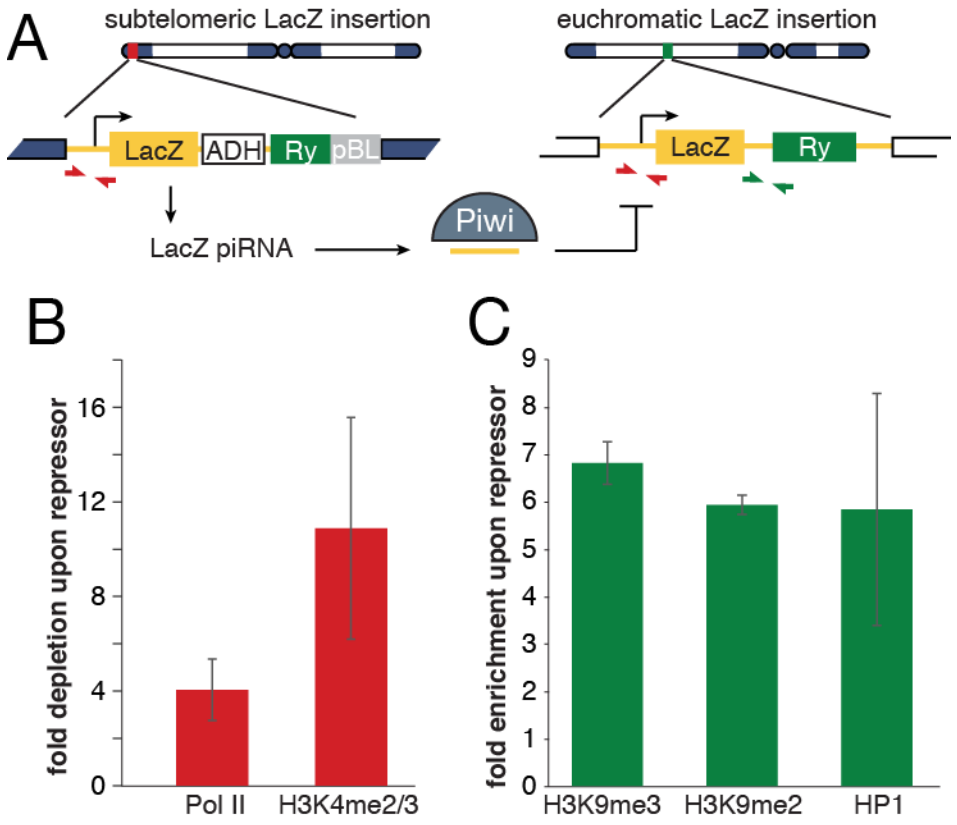

Figure 6. piRNA-dependent targeting of Piwi to a reporter locus leads to establishment of repressive chromatin state and transcriptional silencing.

(**A**) The mechanism of trans-silencing mediated by artificial piRNA and schematic representation of the repressor and reporter lacZ constructs. The repressor construct is inserted in a subtelomeric piRNA cluster leading to generation of piRNA from its sequence. Primers mapping to both constructs used for the RNA pol II and H3K4me2/3 ChIP-qPCR are shown by red arrows, and primers specific to the reporter locus used for the H3K9me3, H3K9me2, and HP1 ChIP-qPCR are indicated by green arrows. (**B**) piRNAs induce transcriptional repression of the lacZ reporter. Pol II and H3K4me2/3 signals decreased on the lacZ promoter in the presence of artificial piRNAs as measured by ChIP-qPCR. Shown is the fold depletion of signal in flies that carry both repressor and reporter constructs compared to control flies that have only the reporter construct. The signal was normalized to RP49. (**C**) piRNAs induce an increase in H3K9me3 and H3K9me2 marks, and HP1 binding as measured by ChIP-qPCR. Shown is the fold increase of corresponding ChIP signals downstream of the lacZ reporter in flies that carry both repressor and reporter

constructs compared to control flies that have only reporter construct. The signal was normalized to RP49.

Chapter III

PIRNA-INDEPENDENT RECRUITMENT OF PIWI ALONE TO RNA IS NOT SUFFICIENT TO TRIGGER ESTABLIMENT OF A REPRESSIVE CHROMATIN STATE AND INDUCE TRANSCRIPTIONAL SILENCING

Part of this work has been submitted for publication as:

Chen, A. C. et al. "Su(Var)2-10, a novel component of the piRNA-induced transcriptional silencing complex, controls co-transcriptional repression of piRNA targets". (Submitted).

Introduction

While the function of nuclear piwi proteins is not fully resolved, there have been several advances in our understanding of the role of *Drosophila* Piwi. The sole nuclear piwi protein in *Drosophila* has been implicated in the regulation of chromatin marks on transposon sequences (Lin and Yin 2008; Klenov et al. 2011; Wang and Elgin 2011; Sienski et al. 2012; Le Thomas et al. 2013; Rozhkov et al. 2013). However, the mechanism by which Piwi accomplishes deposition of repressive chromatin marks specifically on transposon sequences is unknown.

Piwi proteins are not capable of inducing chromatin modifications and transcriptional silencing by themselves, implying they must interact with other factors to successfully repress targets. Furthermore, the only DNA or RNA-binding domains Piwi contains is a piRNA binding domain, suggesting two potential mechanisms for Piwi recruitment to target loci: (1) Piwi is recruited to chromatin through interactions with other factors and (2) Piwi is recruited to target loci via sequence-specific recognition by Piwi-loaded piRNA. Despite previous evidence that suggested that HP1 mediates Piwi's interaction with chromatin (Brower-Toland et al. 2007), Piwi's putative HP1 binding site was shown to be dispensable for Piwi-mediated transposon silencing (Wang and Elgin 2011). Furthermore, we showed that GFP-Piwi does not localize in a pattern that overlaps with HP1 on polytene chromosomes in ovarian nurse cells (Le Thomas et al. 2013), suggesting that HP1 is does not mediate Piwi's interaction with chromatin. In an attempt to identify chromatin-associated factors that interact with Piwi, we performed proteomic analysis of Piwi complexes isolated from *Drosophila* ovaries. Our analyses did not

identify chromatin-associated factors, but rather revealed Piwi interacts with RNA-binding factors, that associate with nascent transcripts, in an RNA-dependent manner (Le Thomas et al. 2013). This data suggested that recruitment of Piwi to transposon loci is achieved by Piwi's interactions with other RNA-binding factors or through the sequence-specific recognition of nascent transcripts by Piwi-bound piRNAs.

While genetically many factors have been implicated in the piRNA pathway, *in vivo* analysis is hindered because disruption of factors involved in the early processes of the pathway results in the inability to determine if these factors may also act further downstream. Previously, we established that piRNAs are the essential guides for recruitment of Piwi to loci that are targeted for transcriptional silencing (Le Thomas et al. 2013). We showed that piRNA-recruitment of Piwi to a reporter sequence resulted in transcriptional silencing of the reporter, accompanied by the establishment of a repressive chromatin state at the reporter locus (Le Thomas et al. 2013). In an effort to reveal the pathway that leads to transcriptional repression of loci downstream of Piwi, we wanted to piRNA-independently recruit Piwi to a reporter locus. We reasoned that piRNA-independent recruitment of Piwi to a reporter should result in the establishment of a repressive chromatin state along the reporter locus and transcriptional silencing of the reporter. This reporter system, in combination of systematic knockdown of piRNA pathway factors, would provide an *in vivo* tool to determine if factors act downstream of Piwi in the piRNA-induced transcriptional silencing complex (piRITS).

Results

Artificial tethering of Piwi to RNA does not lead to transcriptional repression of a reporter

Previously, we determined that piRNA-independent recruitment of Piwi to chromatin, using the lacI DNA binding domain that recognizes lacO sequences inserted upstream of a reporter gene, did not result in establishment of a repressive chromatin state at the reporter (Le Thomas et al. 2013). This result is consistent with our data suggesting Piwi interacts with nascent transcripts rather than with chromatin. To determine if piRNA-independent recruitment of Piwi via RNA is sufficient to induce transcriptional silencing of a reporter *in vivo*, we generated a heterologous tethering system to a reporter that does

not have sequence homology with the *Drosophila* genome (Fig. 1A). The sequence of the reporter mRNA contains a fluorophore (mKate2) followed by four BoxB sequences in its 3'UTR, which are bound by the λN domain of antiterminator protein N fused to Piwi (Fig. 1A). Expression of both λN-GFP-tagged Piwi and the reporter was driven in the germ cells of the fly ovary using a maternal tubulin-GAL4 driver. Like wild type Piwi, the majority of the λN-GFP-tagged Piwi is localized to the nucleus (Fig. 1B). Analysis of the reporter transcript level by RT-qPCR revealed recruitment of Piwi to the transcript does not significantly affect reporter mRNA levels compared to recruitment of the λN-GFP control (Fig. 1C). Furthermore, profiling of RNA polymerase II over the reporter locus by ChIP-qPCR indicated that Pol II occupancy is not affected upon Piwi tethering (Fig. 1D). Accordingly, analysis of the repressive H3K9me3 mark along the reporter locus by ChIP-qPCR and ChIP-seq showed that H3K9me3 levels were unaffected by recruitment of Piwi (Fig. 1E-F). These results are consistent with previous findings (Sienski et al. 2015; Yu et al. 2015) that indicate artificial tethering of Piwi to RNA is not sufficient to induce transcriptional silencing and changes in chromatin state of the reporter.

Recruitment of factors that are known to act downstream of Piwi to RNA is sufficient to induce silencing of the reporter at the transcript level

Previously, genetic screens identified two factors of unknown function, Asterix (Arx) and Panoramix (Panx), which associate with Piwi and are necessary for transcriptional silencing of targets (Donertas et al. 2013; Muerdter et al. 2013; Sienski et al. 2015; Yu et al. 2015). Analysis of the reporter transcript level by RT-qPCR revealed recruitment of Panx to the transcript results in significant repression of reporter mRNA levels compared to recruitment of the λN-GFP control (Fig. 2A). This result is consistent with previous findings (Sienski et al. 2015; Yu et al. 2015). Reporter transcript levels are also significantly reduced upon tethering of Arx to the reporter mRNA (Fig. 2A).

Eggless (Egg), a histone methyltransferase that deposits the H3K9me3 mark, has been implicated in Piwi-mediated transcriptional silencing (Rangan et al. 2011; Sienski et al. 2015; Yu et al. 2015). We recruited Egg to the reporter transcript and analyzed the transcript level by RT-qPCR. Our analysis showed that tethering of Egg to the reporter is also sufficient to significantly reduce the reporter transcript level (Fig. 2A).

Recruitment of Panx is sufficient to establish a repressive chromatin state at the reporter locus

While recruitment of Panx, Arx, and Egg to the reporter showed reduced transcript levels, we wanted to determine if their recruitment established a repressive chromatin state at the locus of the reporter. Consistent with previous reports, ChIP-qPCR analysis of the H3K9me3 mark along the reporter locus revealed that recruitment of Panx, but not Arx, is sufficient to establish a repressive chromatin state (Sienski et al. 2015; Yu et al. 2015) (Fig. 2B-C). Interestingly, ChIP-qPCR analysis revealed recruitment of the H3K9me3 "writer" Egg is not sufficient to induce a repressive chromatin state along the reporter (Fig. 2D).

Recruitment of Su(Var)2-10, a novel component of the piRNA-induced transcriptional silencing complex, is sufficient to transcriptionally silence the reporter locus

The molecular mechanism that connects Piwi and its associated factors involved in transcriptional silencing (Panx, Arx, and Egg) is unknown. Efforts in our lab identified Su(Var)2-10, a conserved protein whose homologs have been demonstrated to act as E3 ligases for SUMOylation of substrates (Hari et al. 2001), as a candidate factor that links Piwi and the rest of the piRITS. We tethered λN-tagged Su(Var)2-10 to our reporter to study Su(Var)2-10's role in the piRITS. Analysis of the reporter transcript level by RT-qPCR revealed recruitment of Su(Var)2-10 to the transcript results in significant repression of reporter mRNA levels compared to recruitment of the λN-GFP control (Fig. 3A). Furthermore, ChIP-qPCR and ChIP-seq analysis showed recruitment of Su(Var)2-10 to the reporter transcript results in the establishment of the repressive H3K9me3 mark along the reporter transcript (Fig. 3B). To determine if transcriptional silencing by Su(Var)2-10 is dependent on the activity of Egg, we tethered Su(Var)2-10 to the reporter and concomitantly knocked-down Egg. RT-qPCR and ChIP-qPCR analyses revealed de-repression of the reporter at the transcript level and a loss of H3K9me3 along the reporter locus (Fig. 3C-D), suggesting that transcriptional silencing of the reporter by Su(Var)2-10 is Egg-dependent.

To determine if Su(Var)2-10 acts upstream or downstream of Panx in the piRITS, we concomitantly tethered Panx and knocked-down Su(Var)2-10. In this case, RT-qPCR analysis revealed a modest increase in reporter expression (Fig. 3E). Accordingly, ChIP-qPCR analysis showed recruitment of Panx to our mKate2 reporter with concomitant knockdown of Su(Var)2-10 resulting in an insignificant loss of H3K9me3 along the

reporter locus (Fig. 3F). However, when Panx is recruited to a reporter that contains 10 BoxB sequences (allowing for stronger recruitment of Panx) we are able to see a significant, but not complete, loss of H3K9me3 along the reporter locus by ChIP-qPCR (Fig. 3G). Importantly, in the converse situation – recruitment of Su(Var)2-10 with concomitant knockdown of Panx – RT-qPCR and ChIP-qPCR show no significant change in reporter transcript levels or H3K9me3 occupancy along the reporter locus (Fig. 3H-I). Taken together, these data suggest a model in which Su(Var)2-10 acts in the piRITS downstream of Panx (and accordingly Piwi and Arx) and upstream of Egg (Fig. 4).

Discussion

In Metazoa, the piRNA-induced transcriptional silencing complex (piRITS) plays an essential role in genome defense against transposable elements (TEs). In *Drosophila*, targeting of TE loci by Piwi/piRNA complexes is required for deposition of the repressive chromatin mark H3K9me3 at TE loci (Sienski et al. 2012; Le Thomas et al. 2013; Rozhkov et al. 2013). The existing model proposes that piRNA-loaded Piwi binds to nascent transcripts in search of complementary targets; however, how recognition of nascent RNA by Piwi/piRNA complexes recruits the machinery necessary for H3K9me3 deposition is unclear. With the intent of determining factors involved in the piRITS downstream of Piwi, we generated a heterologous tethering system expressed in *Drosophila* ovaries (Fig. 1A-B). Our tethering system consists of a fluorescent reporter (mKate2) followed by 4BoxB sequences, allowing for λN-GFP-tagged Piwi to be recruited in a piRNA-independent manner to reporter transcripts.

Unfortunately, piRNA-independent recruitment of Piwi to a reporter transcript is not sufficient to induce transcriptional silencing or a change in the chromatin state of the reporter locus (Fig. 1C-F); however this data is corroborated by previous findings (Sienski et al. 2015; Yu et al. 2015). The fact that piRNA-independent tethering of Piwi alone is not sufficient to recruit the silencing effector complex implies that recognition of a target by the piRNA loaded within Piwi is essential. Upon loading or after target recognition a Piwi/piRNA complex may be modified or undergo a conformational change that permits the formation of the bridge between the Piwi/piRNA complex and the silencing effector complex.

Despite the heterologous tethering system not working as expected for Piwi, we were still able to use our heterologous tethering system to probe factors downstream of Piwi in the piRITS. Asterix (Arx) and Panoramix (Panx) were previously identified as piRITS factors that associate with Piwi and are necessary for transcriptional silencing of targets (Sienski et al. 2015; Yu et al. 2015). Consistent with previous findings (Sienski et al. 2015; Yu et al. 2015), our assay revealed that recruitment of λN-Panx and, to a lesser extent, recruitment of λN-Arx results in repression of the reporter transcript (Fig. 2A). The histone methyltransferase Eggless (Egg) acts further downstream of Panx and Arx in the piRITS to deposit the H3K9me3 at the target locus (Rangan et al. 2011; Sienski et al. 2015; Yu et al. 2015). Similar to Panx and Arx, recruitment of λN-Egg resulted in repression of the reporter transcript (Fig. 2A). Taken together, this data illustrates that we can induce repression of our reporter by tethering factors that act downstream of Piwi in the piRITS, suggesting we can use this system to further probe the interactions between these factors.

While the piRITS factors, known to act downstream of Piwi, tested were all able to induce repression of the reporter transcript, we found that only recruitment of Panx was able to alter the reporter chromatin state and establish occupancy of H3K9me3 along the reporter locus (Fig. 2B-D). The fact that recruitment of Egg, the H3K9me3 "writer", was not sufficient to induce a repressive chromatin state along the reporter locus suggests that Egg may undergo modification prior to being able to deposit H3K9me3. Previous publications have shown that Egg is SUMOylated (Rosendorff et al. 2006; Koch et al. 2009). Taken with our results, this implies that recruitment of non-SUMOylated Egg does not result in the establishment of a repressive chromatin state along the reporter.

In our lab's search to uncover the molecular mechanism that connects Piwi and the the rest of the piRITS, we identified Su(Var)2-10 – an E3 ligase for SUMOylation of substrates (Hari et al. 2001) – as a factor that may provide a link between Piwi and Egg. Recruitment of λN-Su(Var)2-10 to our reporter resulted in repression of the reporter transcript and H3K9me3 occupancy along the reporter locus (Fig. 3A-B). Furthermore, recruitment of λN-Su(Var)2-10 with concomitant knockdown of Egg indicated that repression by Su(Var)2-10 is dependent on Egg (Fig 3C-D). This data places Su(Var)2-10 upstream of Egg in the piRITS.

Similar to Su(Var)2-10, previous studies have show that transcriptional silencing by Panx requires Egg (Sienski et al. 2015; Yu et al. 2015). In a modular fashion, we used our heterologous tethering assay to determine whether Su(Var)2-10 acted upstream or downstream of Panx in the piRITS. We tethered Panx with concomitant knockdown of Su(Var)2-10 and found a modest increase in reporter expression (Fig. 3E). To see a significant loss of H3K9me3 occupancy along the reporter locus, we had to tether Panx to a reporter with 10 BoxB sequences (Fig 3F-G). In the reciprocal experiment, knockdown of Panx did not affect repression of the reporter locus by Su(Var)2-10. Taken together, this data indicates that transcriptional repression by Panx is dependent on Su(Var)2-10, whereas repression by Su(Var)2-10 is not dependent on Panx.

Our heterologous tethering system has allowed for the beginning steps of determining factors that act downstream of Piwi in the piRITS. Currently, the function of Su(Var)2-10 remains unknown; however, we were able to illustrate that Su(Var)2-10 can directly participate in co-transcriptional silencing. Previous studies have reported that Panx and Arx are able to form a complex with Piwi, but not with Egg (Sienski et al. 2015; Yu et al. 2015). Taken together with our data, this suggests a model in which Su(Var)2-10 provides an essential link between the Piwi/piRNA complex and the silencing effector complex by acting downstream of Piwi (and its associated factors Panx and Arx) and upstream of Egg in the piRITS. Furthermore, the putative SUMO E3 ligase activity of Su(Var)2-10 may be essential for Egg's ability to deposit H3K9me3 at target loci (Fig. 4). The next steps in further understanding the mechanisms of the piRITS will be to study the biochemical activities and physical interactions between all factors involved.

Methods and Materials

Drosophila stocks:

The sh-Su(Var)2-10 and sh-White stocks were obtained from the Bloomington Stock Center (BDSC #32956 and BDSC #35182). The sh-Egg and sh-Panoramix lines were a generous gift from the Brennecke lab. The Luciferase-10xBoxB reporter line was a generous gift from the Hannon lab. The UASp-λN-eGFP-Piwi, UASp-λN-eGFP-GFP, UASp-λN-eGFP-Asterix, UASp-λN-eGFP-Panoramix, UASp-λN-eGFP-Egg, UASp-λN-eGFP-Su(Var)2-10, and UASp-mKate2-4xBoxB-K10polyA lines were generated in lab as

previously described (Chen et al. 2016). The expression of all constructs was driven by maternal alpha-tubulin67C-Gal4 (MT-Gal4) obtained from the Bloomington Stock Center (BDSC #7063). Flies carrying the λN-transgenes and the alpha-tubulin67C-Gal4 driver were crossed to flies carrying the UASp-mKate2-4xBoxB-K10polyA reporter and where applicable the respective hairpin, to obtain flies that express all components of the system in the female germline. Flies were put on yeast for 3 to 4 days prior to dissection and were 5-8 days old at the time of dissection.

ChIP-qPCR and ChIP-seq:

H3K9me3 ChIPs were carried out using anti-H3K9me3 from abcam (ab8898) and RNA Polymerase II ChIP was carried out using anti-RNA Pol II from abcam (ab5408) with the protocol described in (Le Thomas et al., 2014). One hundred pairs of ovaries were used for each ChIP experiment. qPCR was performed on a Mastercycler® ep realplex PCR thermal cycler machine (Eppendorf). ChIP-qPCR data for H3K9me3 and Pol II ChIPs was calculated as IP to input ratios of the target region, normalized to IP to input ratios for the *rp49* locus. PCR primers are listed in Table S1.

ChIP-seq library construction was done using the NEBNext ChIP-Seq Library Prep Master Mix Set (#E6240) and the NEBNext Multiplex Oligos for Illumina (#E7335S). Libraries were sequenced on the Illumina HiSeq 2500 (SE 50bp reads) platform. The resulting reads were mapped against the reporter sequence and against the DM3 genome using Bowtie 0.12.7 (Langmead et al. 2009) with the following settings: " –v 0 --best –strata". Reads that specifically mapped to the reporter sequence and not to the DM3 genome with zero mismatches were extracted from the libraries. The reads were plotted along the reporter sequence and normalized to number of reads mapping to the DM3 genome using IGV 2.3.68 (Robinson et al. 2011).

RT-qPCR:

Total RNA was isolated from 20 ovaries with Ribozol (Amresco, N580), followed by DNAse treatment with Amplification grade DNAseI (Invitrogen, 18068-015) according to the manual. Reverse transcription was carried out using Superscript III (Invitrogen) with oligod$(T)_{15}$. qPCR was performed on a Mastercycler® ep realplex PCR thermal cycler machine (Eppendorf). RT-qPCR data target expression was normalized to the *rp49* locus expression. PCR primers are listed in Table S1.

References

Brower-Toland B, Findley S, Jiang L, Liu L, Yin H, Dus M, Zhou P, Elgin S, Lin H. 2007. Drosophila PIWI associates with chromatin and interacts directly with HP1a. *Genes & development* **21**: 2300-2311.

Chen YC, Stuwe E, Luo Y, Ninova M, Le Thomas A, Rozhavskaya E, Li S, Vempati S, Laver JD, Patel DJ et al. 2016. Cutoff Suppresses RNA Polymerase II Termination to Ensure Expression of piRNA Precursors. *Molecular cell* **63**: 97-109.

Donertas D, Sienski G, Brennecke J. 2013. Drosophila Gtsf1 is an essential component of the Piwi-mediated transcriptional silencing complex. *Genes & development* **27**: 1693-1705.

Hari KL, Cook KR, Karpen GH. 2001. The Drosophila Su(var)2-10 locus regulates chromosome structure and function and encodes a member of the PIAS protein family. *Genes & development* **15**: 1334-1348.

Klenov M, Sokolova O, Yakushev E, Stolyarenko A, Mikhaleva E, Lavrov S, Gvozdev V. 2011. Separation of stem cell maintenance and transposon silencing functions of Piwi protein. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 18760-18765.

Koch CM, Honemann-Capito M, Egger-Adam D, Wodarz A. 2009. Windei, the Drosophila homolog of mAM/MCAF1, is an essential cofactor of the H3K9 methyl transferase dSETDB1/Eggless in germ line development. *PLoS genetics* **5**: e1000644.

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology* **10**: R25.

Le Thomas A, Rogers AK, Webster A, Marinov GK, Liao SE, Perkins EM, Hur JK, Aravin AA, Toth KF. 2013. Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. *Genes & development* **27**: 390-399.

Lin H, Yin H. 2008. A novel epigenetic mechanism in Drosophila somatic cells mediated by Piwi and piRNAs. *Cold Spring Harbor symposia on quantitative biology* **73**: 273-281.

Muerdter F, Guzzardo PM, Gillis J, Luo Y, Yu Y, Chen C, Fekete R, Hannon GJ. 2013. A genome-wide RNAi screen draws a genetic framework for transposon control and primary piRNA biogenesis in Drosophila. *Molecular cell* **50**: 736-748.

Rangan P, Malone CD, Navarro C, Newbold SP, Hayes PS, Sachidanandam R, Hannon GJ, Lehmann R. 2011. piRNA production requires heterochromatin formation in Drosophila. *Current biology : CB* **21**: 1373-1379.

Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011. Integrative genomics viewer. *Nature biotechnology* **29**: 24-26.

Rosendorff A, Sakakibara S, Lu S, Kieff E, Xuan Y, DiBacco A, Shi Y, Shi Y, Gill G. 2006. NXP-2 association with SUMO-2 depends on lysines required for transcriptional repression. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 5308-5313.

Rozhkov NV, Hammell M, Hannon GJ. 2013. Multiple roles for Piwi in silencing Drosophila transposons. *Genes & development* **27**: 400-412.

Sienski G, Batki J, Senti KA, Donertas D, Tirian L, Meixner K, Brennecke J. 2015. Silencio/CG9754 connects the Piwi-piRNA complex to the cellular heterochromatin machinery. *Genes & development* **29**: 2258-2271.

Sienski G, Dönertas D, Brennecke J. 2012. Transcriptional silencing of transposons by piwi and maelstrom and its impact on chromatin state and gene expression. *Cell* **151**: 964-980.

Wang S, Elgin S. 2011. Drosophila Piwi functions downstream of piRNA production mediating a chromatin-based transposon silencing mechanism in female germ line. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 21164-21169.

Yu Y, Gu J, Jin Y, Luo Y, Preall JB, Ma J, Czech B, Hannon GJ. 2015. Panoramix enforces piRNA-dependent cotranscriptional silencing. *Science* **350**: 339-342.

Figures and Figure Legends

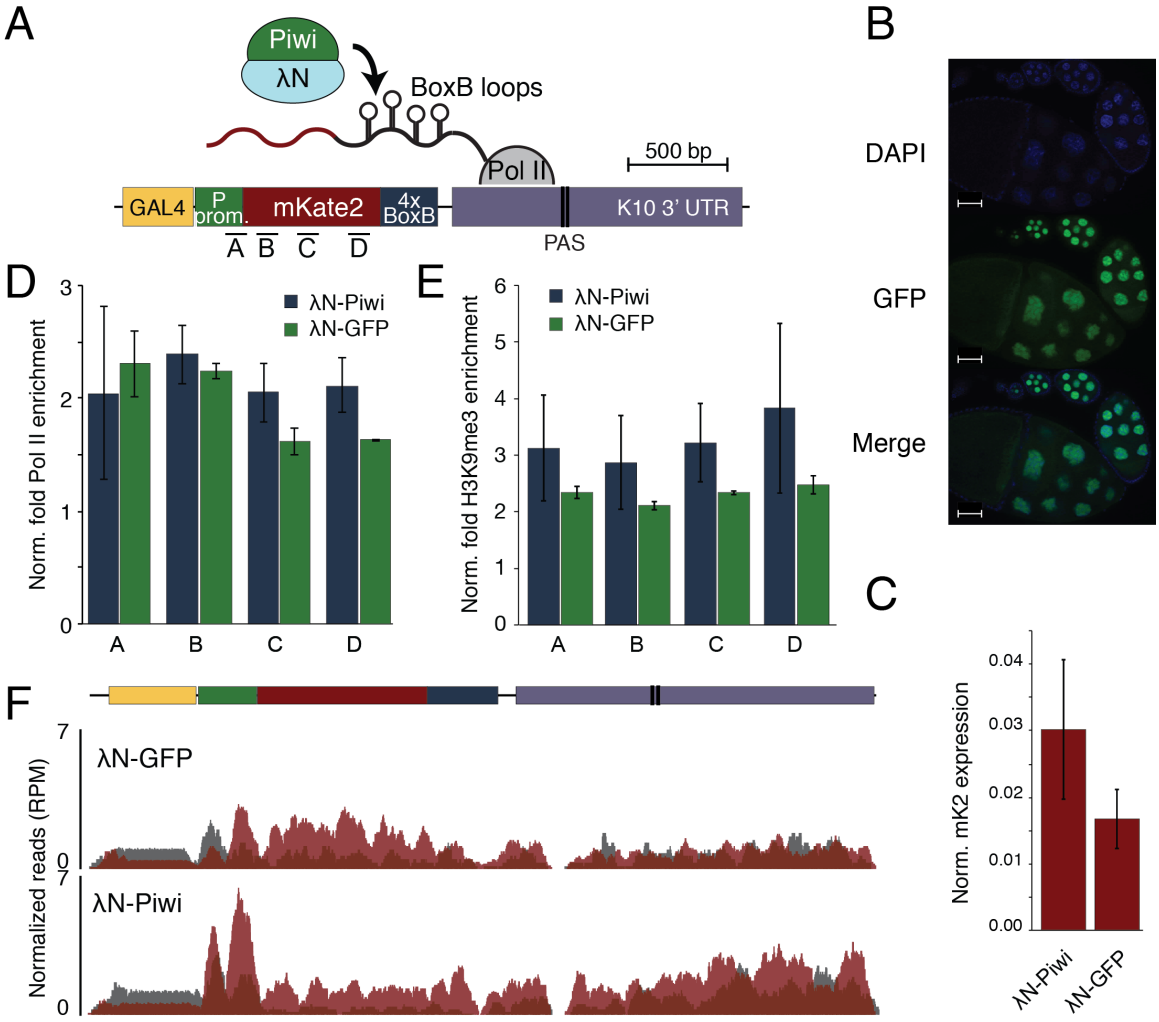

Figure 1. Piwi recruitment to RNA does not change transcription or chromatin structure of the locus.

(A) Schematic diagram of the reporter construct used to study the effect of Piwi recruitment to RNA. λN-GFP-Piwi or λN-GFP-eGFP was co-expressed with the reporter in ovaries driven by MT-Gal4. **(B)** Like wildtype Piwi, in steady-state, λN-GFP-Piwi is nuclear in localization. Scale bars represent 10μM. **(C)** Piwi tethering does not lead to an affect on reporter expression level compared to tethering of the control, λN-GFP-eGFP. Relative expression level of mKate2 was measured using RT-qPCR and primer pair C. The error bars represent the standard deviation of ten biological replicates. **(D)** Tethering of Piwi results in no change in RNA Pol II occupancy at the locus of the reporter as measured by ChIP-qPCR. The error bars represent the standard deviation of two biological replicates. **(E)** Tethering of Piwi results in no change in H3K9me3

occupancy at the locus of the reporter as measured by ChIP-qPCR. The error bars represent the standard deviation of two biological replicates. **(F)** H3K9me3 ChIP-seq indicates that tethering of Piwi does not result in a change in H3K9me3 occupancy across the reporter locus. Input is shown in gray and H3K9me3 IP is shown in red. Shown are all reads mapped to the reporter, normalized by the total number of reads mapping to the DM3 genome.

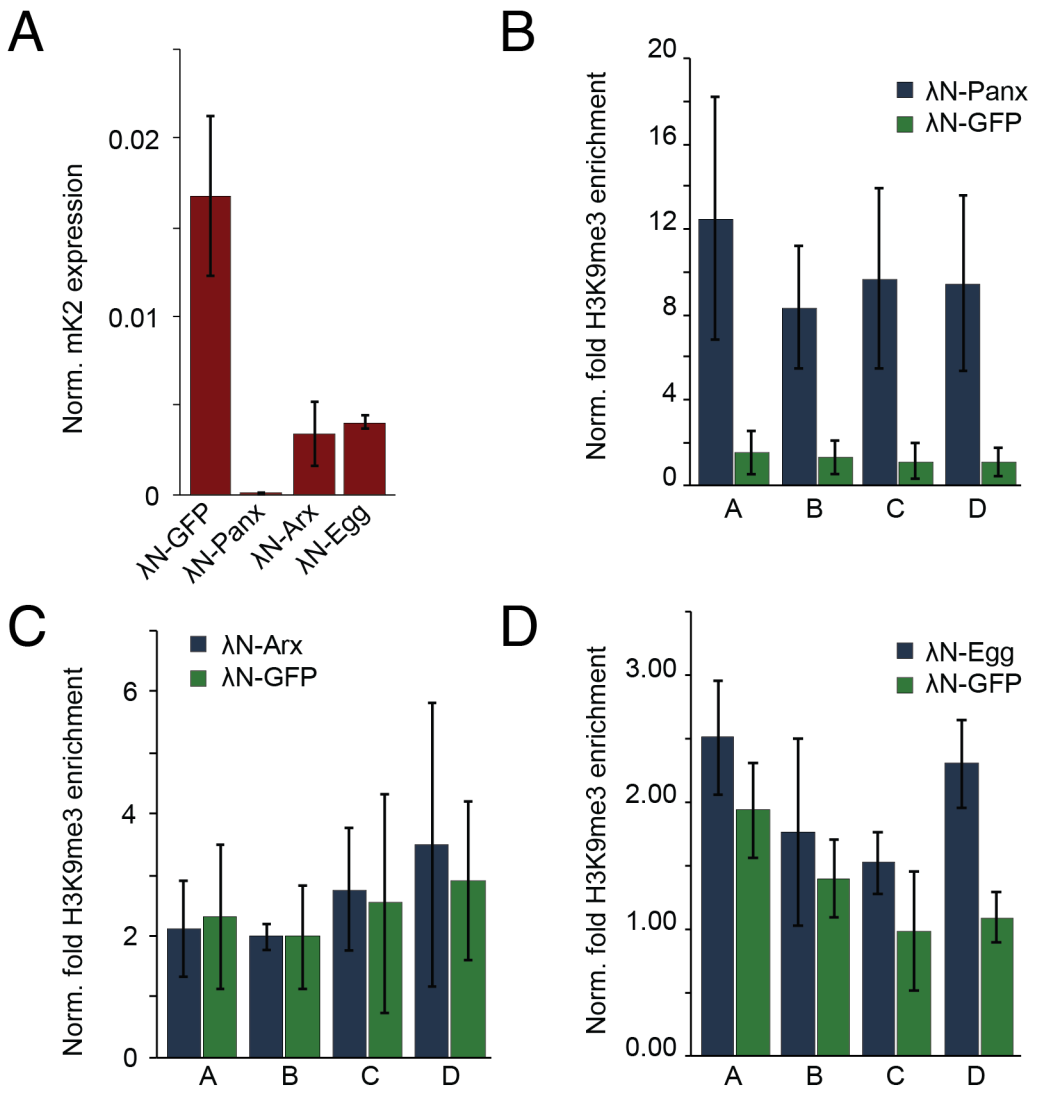

Figure 2. Recruitment of factors downstream of Piwi in the piRITS to the reporter transcript induces changes in transcription status of the locus, but only Panx induces changes in chromatin structure of the locus.

(A) Tethering of Panx, Arx, and Egg leads to an affect on reporter expression level compared to tethering of the control, λN-GFP-eGFP. Relative expression level of mKate2 was measured using RT-qPCR and primer pair C. The error bars represent the standard deviation of two biological replicates. **(B)** Tethering of Panx results in a change in H3K9me3 occupancy at the locus of the reporter as measured by ChIP-qPCR. The error bars represent the standard deviation of two biological replicates. **(C)** Tethering of Arx results in no change in H3K9me3 occupancy at the locus of the reporter as measured by ChIP-qPCR. The error bars represent the standard deviation of two biological replicates. **(D)** Tethering of Egg results in a change in H3K9me3

occupancy at the locus of the reporter as measured by ChIP-qPCR. The error bars represent the standard deviation of two biological replicates.

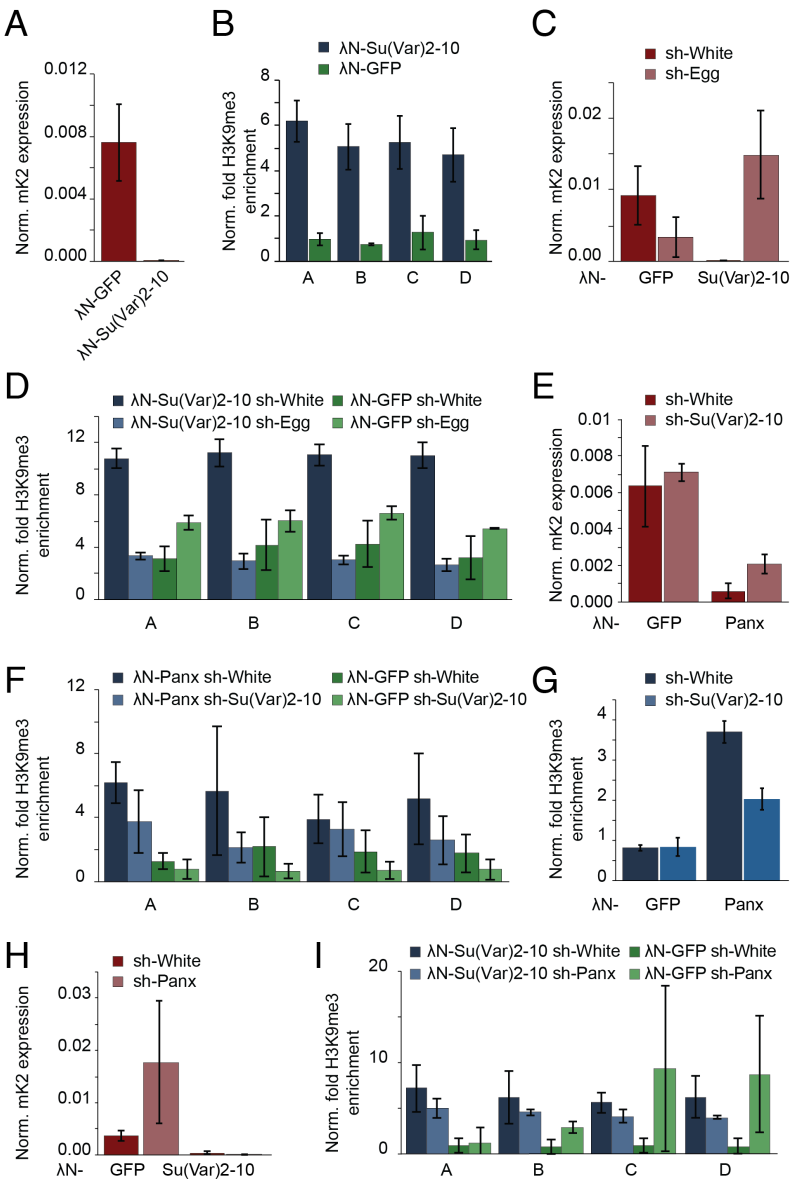

Figure 3. Su(Var)2-10 acts downstream of Piwi and Panx and upstream of Egg in the piRITS. **(A)** Tethering of Su(Var)2-10 leads to repression of reporter expression level compared to tethering of the control, λN-GFP-eGFP. Relative expression level of mKate2 was measured using RT-qPCR and primer pair C. The error bars represent the standard deviation of two biological replicates. **(B)** Recruitment of Su(Var)2-10 results in an increase in H3K9me3 occupancy at the locus of the reporter as measured by ChIP-qPCR. The error bars represent the standard deviation of two biological replicates. **(C)** Tethering of Su(Var)2-10 results in no change in H3K9me3 occupancy at the locus of the reporter as measured by ChIP-qPCR. The error bars represent the

standard deviation of two biological replicates. **(D)** Tethering of Egg results in a change in H3K9me3 occupancy at the locus of the reporter as measured by ChIP-qPCR. The error bars represent the standard deviation of two biological replicates.

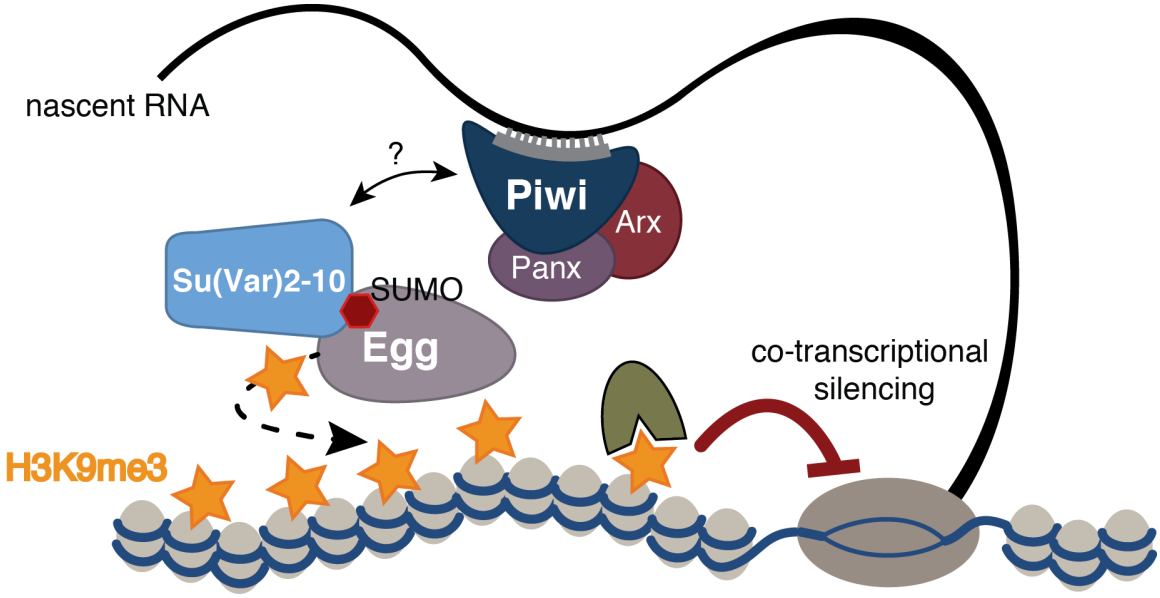

Figure 4. Model of the piRNA-induced transcriptional silencing complex.

Arx and Panx associate with piRNA-loaded Piwi and work to recognize RNA targets. Downstream, H3K9me3 is deposited on Piwi targets by the histone methyltransferase Egg, further recruiting silencing factors that induce transcriptional silencing. Su(Var)2-10 interacts with Egg downstream of Piwi, providing a bridge between the piRNA-Piwi complex and the silencing effector complex. The biochemical link between Piwi and Su(Var)2-10 remains elusive.

Chapter IV

ZUCCHINI-DEPENDENT PIRNA PROCESSING IS TRIGGERED BY RECRUITMENT TO THE CYTOPLASMIC PROCESSING MACHINERY

This work was first published as:

Rogers, A. K. et al. (2017). "Zucchini-dependent piRNA processing is triggered by recruitment to the cytoplasmic processing machinery". In: *Genes & Development* 31: 1-12. doi: 10.1101/gad.303214.117.

Abstract

The piRNA pathway represses transposable elements in the gonads and thereby plays a vital role in protecting the integrity of germline genomes of animals. Mature piRNAs are processed from longer transcripts, piRNA precursors (pre-piRNAs). In Drosophila, processing of pre-piRNAs is initiated by piRNA-guided Slicer cleavage, or by the endonuclease Zucchini (Zuc). As Zuc does not have any sequence or structure preferences *in vitro*, it is not known how piRNA precursors are selected and channeled into the Zuc-dependent processing pathway. We show that a heterologous RNA that lacks complementary piRNAs is processed into piRNAs upon recruitment of several piRNA pathway factors. This processing requires Zuc and the helicase Armitage (Armi). Aubergine (Aub) and Argonaute 3 (Ago3), as well as components of the nuclear RDC complex, which are required for piRNA biogenesis in germ cells, are dispensable. Our approach allows discrimination of proteins involved in transcription and export of piRNA precursors from components required for the cytoplasmic processing steps. piRNA processing correlates with localization of the substrate RNA to nuage, a distinct membraneless cytoplasmic compartment, which surrounds the nucleus of germ cells, suggesting that sequestration of RNA to this subcellular compartment is both necessary and sufficient for selecting piRNA biogenesis substrates.

Introduction

Mature piRNAs are processed from longer transcripts, piRNA precursors (pre-piRNAs), most of which are derived from piRNA clusters – genomic regions with a high density of transposon sequences (Brennecke et al. 2007; Gunawardane et al. 2007). Insertion of a heterologous sequence in piRNA clusters leads to its processing into piRNAs, indicating that piRNAs can be processed from any sequence (Todeschini et al. 2010; Muerdter et al. 2012). piRNA biogenesis is a multi-step process starting with transcription and early processing of precursor RNA in the nucleus, export, precursor cleavage followed by processing to mature piRNAs and loading into piwi proteins. In the cytoplasm, piRNA precursors are processed by two mechanisms: the ping-pong cycle and Zucchini-dependent piRNA biogenesis (Brennecke et al. 2007; Gunawardane et al. 2007; Malone et al. 2009). In the ping-pong cycle, processing of pre-piRNAs is initiated by the endonucleolytic (slicer) activity of cytoplasmic piwi proteins, Aubergine (Aub) or Argonaute3 (Ago3), loaded with complementary piRNAs (Brennecke et al. 2007; Gunawardane et al. 2007). The characteristic features of ping-pong biogenesis are the presence of complementary piRNA pairs whose 5' ends are separated by 10nt, and the generation of piRNAs with a bias for U at position 1 and for A at position 10 of their sequence (Wang et al. 2014). Zucchini-dependent biogenesis, in which endonucleolytic cleavage of pre-piRNAs by Zuc generates the 5', and in part the 3', ends of piRNAs, neither depends on the slicer activity of piwi proteins, nor the presence of complementary piRNAs (Malone et al. 2009). piRNAs generated through Zuc-dependent biogenesis show a bias for U at their 5' end but lack a bias for A at position 10. Many of the cytoplasmic piRNA biogenesis factors, including Aub and Ago3 as well as numerous factors that were genetically identified to act in the pathway such as Vasa, Armitage (Armi), and Krimper, localize to membraneless perinuclear structures called nuage in germ cells and Yb bodies in the surrounding somatic follicular cells, leading to the assumption that these structures are the site of piRNA biogenesis (Hay et al. 1988; Lasko and Ashburner 1990; Lim and Kai 2007; Kirino et al. 2009; Malone et al. 2009; Qi et al. 2011; Ipsaro et al. 2012; Nishimasu et al. 2012; Ohtani et al. 2013; Webster et al. 2015).

How piRNA precursors are identified by the biogenesis machinery remains unresolved. Other RNA processing events, such as miRNA or CRISPR RNA biogenesis, splicing, or polyadenylation, require specific sequence and/or structure motifs in precursor RNA that are recognized by the processing machinery (Zeng et al. 2005; Park et al. 2011; Barrangou and Marraffini 2014; Li and Patel 2016; Tsai and Joung 2016). Common sequence or structural motifs that are shared by all piRNA precursors were not identified (Muerdter et al. 2012). In lack of sequence motifs, two mutually non-exclusive models have been proposed to explain precursor selection in the germline, which we call the "persistent nuclear mark model" and the "selection by pre-existing piRNA model".

The "persistent nuclear mark model" proposes that specific proteins associate with piRNA precursors in the nucleus and remain associated with the transcripts in the cytoplasm, where they interact with the processing machinery. In the germline, genomic regions that give rise to the majority of piRNAs, called piRNA clusters, are marked by the histone 3 lysine 9 trimethylation (H3K9me3) mark and by the RDC complex. RDC consists of the HP1 homolog, Rhino, and two additional proteins, Cutoff (Cuff) and Deadlock, and is required for transcription and early processing of cluster transcripts (Klattenhoff et al. 2009; Pane et al. 2011; Le Thomas et al. 2014; Mohn et al. 2014; Zhang et al. 2014; Chen et al. 2016). The well-conserved transcription export (TREX) complex is also enriched at cluster loci and binds piRNA precursors co-transcriptionally in an RDC-dependent fashion (Zhang et al. 2012; Hur and Chung 2016; Hur et al. 2016). Components of either the RDC or the TREX complex were proposed to constitute the mark that triggers processing in the cytoplasm. However, the nature of the mark and the mechanism by which it engages the processing machinery remain unclear. Evidence for RDC or TREX remaining associated with piRNA precursors after nuclear export is also lacking. Finally, tethering of Rhino to a single-stranded transgene (without its concomitant tethering to a complementary antisense transgene) does not trigger piRNA biogenesis (Zhang et al. 2014), arguing against the idea that binding of RDC by itself is sufficient to specify piRNA precursors.

The "selection by pre-existing piRNA" model suggests that precursors are specified in the cytoplasm by complementary piRNAs associated with the

cytoplasmic piwi proteins. Existing piRNAs can target pre-piRNAs and induce their processing via the ping-pong cycle (Brennecke et al. 2007; Gunawardane et al. 2007). piRNA-dependent cleavage not only generates complementary piRNAs (ping-pong partners) but also leads to phased processing of the precursor downstream of the initial cleavage (Han et al. 2015; Homolka et al. 2015; Mohn et al. 2015; Senti et al. 2015; Wang et al. 2015). This process generates Piwi-loaded piRNAs. Mutation of the two piwi proteins involved in the ping-pong cycle, Aub and Ago3, greatly reduces Piwi-bound piRNAs in the germline, leading to the suggestion that Zuc-mediated piRNA biogenesis in germ cells is triggered by the ping-pong cycle (Han et al. 2015; Homolka et al. 2015; Mohn et al. 2015; Senti et al. 2015; Wang et al. 2015).

Follicular cells of the fly ovary lack ping-pong biogenesis factors as well as the RDC complex, yet produce piRNAs, suggesting an alternative precursor selection mechanism. Some piRNA precursors specific to follicular cells, such as the mRNA of *traffic-jam (tj)* and transcripts from the *flamenco* piRNA cluster, harbor sequences that are bound by the RNA binding protein Yb (Ishizu et al. 2015) and can trigger processing into piRNAs (Homolka et al. 2015; Ishizu et al. 2015). Whether other somatic piRNA precursors also harbor sequence motifs that are bound by Yb, and whether binding of Yb leads to piRNA processing, remains to be elucidated.

Genetically many factors have been identified to be essential for piRNA biogenesis; however, a mechanistic dissection of which steps these factors are involved in (i.e., transcription, export, recognition or cytoplasmic processing) is largely missing. Assigning many of the identified factors to specific steps of the biogenesis pathway is hindered by the absence of an *in vitro* system that allows independent analysis of different stages of the pathway. *In vivo* analysis is impaired because disrupting proteins acting in early steps of the pathway, such as transcription of pre-piRNAs, eliminates piRNA precursors and thereby the possibility to test whether these factors also act in downstream steps of processing.

The core of the cytoplasmic processing machinery, including the endonuclease Zuc, is shared by nurse and follicular cells, raising the question of how the same processing machinery has evolved to recognize its targets by two distinct – sequence-specific in follicular cells and sequence-independent in nurse cells – mechanisms. We

show that direct recruitment of a heterologous sequence to the cytoplasmic piRNA processing machinery can bypass the necessity for sequence-specific recognition, as well as the need of both complementary piRNAs and a unique chromatin architecture. Our work reveals that sequestration of RNA to the cytoplasmic processing machinery is both necessary and sufficient for selecting substrates for piRNA biogenesis, leading to a unified model of piRNA specification in the two cell types. The described experimental approach bypasses nuclear biogenesis steps, enabling identification and functional dissection of factors that act in Zuc-mediated processing downstream of nuclear export.

Results

Recruitment of Yb and Piwi to RNA in follicular cells leads to piRNA processing

Two transcripts that are processed into piRNAs in follicular cells, the mRNA of *tj* and transcripts from the *flamenco* piRNA cluster, were shown to harbor sequences that associate with the Yb protein and trigger piRNA processing when transferred to heterologous transcripts (Homolka et al. 2015; Ishizu et al. 2015). To determine if the interaction of a transcript with Yb is sufficient to induce processing, we tested whether we could trigger piRNA production in follicular cells by tethering a heterologous sequence that does not have sequence homology with natural piRNA substrates to Yb (Fig. 1A). The sequence of the reporter mRNA contains four BoxB sequences in its 3'UTR, which are bound by the λN domain of antiterminator protein N fused to Yb (Fig. 1A). We co-expressed λN-tagged Yb with the reporter in follicular cells of the fly ovary and sequenced ovarian small RNAs. Tethering of the transcript to Yb resulted in production of reporter-derived small RNAs, while small RNAs were absent upon recruitment of the control GFP protein (Fig. 1B-C). The majority of small RNAs produced from the reporter upon Yb tethering were 25-28nt long and had a U bias at the 1st nucleotide (82.2% of all reporter-derived reads had a 5'U) consistent with these sequences being piRNAs (Fig. 1D-E). Thus, recruitment of Yb to a novel transcript is sufficient to trigger piRNA processing in follicular cells in the absence of any sequence or structure motifs in the RNA.

In follicular cells, piRNAs are loaded into Piwi, the only member of the piwi clade expressed in these cells. Therefore we asked whether recruitment of Piwi itself to

RNA is sufficient to trigger processing. Tethering of Piwi also induced generation of piRNAs from the reporter (Fig. 1B-E). We note, that the piRNA profiles along the reporter differ depending on the factor recruited. Piwi recruitment leads to processing of the entire transcript, while Yb recruitment leads to piRNA biogenesis only downstream of the BoxB sites, which the Yb fusion protein binds to (Fig. 1C). This is similar to the piRNAs observed in the *tj* mRNA and when using the recognition sequence from flamenco, which both lead to processing only downstream of the recognition sequence (Homolka et al. 2015; Ishizu et al. 2015). Our data indicate that Yb binding not only identifies piRNA precursors but also specifies the entry site for processing. Thus, in follicular cells, recruitment of either Yb or Piwi is sufficient to trigger piRNA processing from a novel substrate and Yb might be responsible for specifying where processing initiates within the RNA.

Recruitment of the nuclear RDC and TREX complexes or Zucchini is not sufficient to trigger piRNA generation in germ cells

In the germline piRNA biogenesis is more complex than in somatic follicular cells and depends on numerous nuclear and cytoplasmic factors. Two nuclear complexes, RDC and TREX, that localize to chromatin of piRNA loci and nascent pre-piRNA, respectively, were proposed to play a role in selecting piRNA precursors and guiding them to the processing machinery in germ cells. To test if the recruitment of the RDC or TREX to the heterologous transcript is sufficient to trigger piRNA processing in germ cells, we tethered the RDC component Cuff and the TREX component Thoc5 to the reporter in the female germline and profiled ovarian small RNAs. Tethering of these factors did not trigger piRNA generation (Fig. 2A-B), suggesting that recruitment of RDC or TREX is not sufficient to trigger biogenesis from novel RNA substrates in germ cells.

Next we tested if recruitment of Zuc, the endonuclease responsible for generating the 5’ end and - to some extent - the 3’ end of piRNAs (Malone et al. 2009; Ipsaro et al. 2012; Nishimasu et al. 2012; Han et al. 2015; Mohn et al. 2015) is sufficient to trigger piRNA biogenesis. Small RNA sequencing upon Zuc tethering revealed the presence of abundant small RNAs arising from the reporter transcript (Fig. 2A-B). However, small RNA sequences were predominantly 22-23nt in length (mostly from the 5’ region of the

reporter) (Fig. S1) and were depleted of 5'U reads (13.9% of all reporter-mapped reads had a 5'U; standard deviation (std) = 1.3; note that the reporter is depleted of uracils with an overall U content of 21.3%) (Fig. 2C-F). Small RNAs produced upon Zuc tethering were sensitive to $NaIO_4$ treatment (Fig. 2A, F), further confirming that the majority of reporter-derived reads upon tethering Zuc are not *bona fide* piRNAs. Thus, Zuc tethering induces processing of the transcript, but not the efficient production of piRNAs. A closer look at the size profile revealed a smaller peak (27.4% of all reporter-derived reads; std = 0.01) around the 25-28nt piRNA size range (Fig. 2C). This small RNA population had slightly more 5'U RNAs (28.4% had 5'U; std = 0.03) and was derived from the entire length of the reporter (Fig. 2E, Supplemental Fig. S1).

To test whether any genuine piRNAs are formed upon recruitment of Zuc, we immunoprecipitated Piwi/piRNA complexes and cloned the isolated piRNAs. Analysis of Piwi-associated RNAs yielded reporter-derived small RNAs with a size profile and 5'U bias characteristic of piRNAs (41.8% had 5'U) (Fig. 2A, G-H). Thus, tethering of Zuc leads to reporter processing into small RNAs, however, only a small fraction of the formed small RNAs are genuine piRNAs.

Recruitment of Piwi and cytoplasmic piRNA pathway components triggers piRNA biogenesis in germ cells

Recent studies suggested that in germ cells both slicer-dependent and Zuc-mediated processing is triggered by initial Aub- or Ago3-mediated cleavage guided by complementary piRNAs (Han et al. 2015; Mohn et al. 2015; Senti et al. 2015; Wang et al. 2015). Our reporter lacks complementary piRNAs based on both bioinformatics analysis (Supplemental Fig. S2) and the fact that no processing of the reporter was detected in control flies upon recruitment of GFP. Thus, our inability to trigger piRNA biogenesis in germ cells might be due to the absence of piRNAs that are complementary to the reporter. We tested whether the recruitment of Piwi – which unlike Yb, is expressed in both cell types (Cox et al. 2000) – can trigger piRNA generation in the absence of cognate piRNAs in germ cells. Small RNA profiling revealed that – like in follicular cells – recruitment of Piwi in germ cells triggers small RNA processing from the reporter (Fig. 3A-B). Similar results were obtained with a reporter inserted in a different genomic location and containing a different sequence (except for the BoxB sites)(Sienski et al. 2015)

(Supplemental Fig. S3), indicating that neither the genomic location of the target, nor its sequence was crucial for processing.

We tested if other components of the piRNA pathway can trigger processing in germ cells upon their recruitment to the reporter. Recruitment of nuclear proteins that work together with, or downstream of, Piwi to repress Piwi targets, Asterix (Arx), Panoramix (Panx) and the histone methyltransferase SetDB1/Egg, (Donertas et al. 2013; Muerdter et al. 2013; Sienski et al. 2015; Yu et al. 2015) did not trigger piRNA biogenesis (Fig. 3A-B). Recruitment of the cytoplasmic components Aub and Armi, two nuage components involved in pong-pong and Zuc-dependent processing respectively, resulted in small RNA biogenesis (Fig. 3A-B). Small RNAs generated upon Piwi, Aub or Armi tethering were resistant to $NaIO_4$ treatment, confirming that they are generated by the canonincal piRNA biogenesis machinery (Supplemental Fig. S4A).

The bulk of the reporter-mapped small RNAs upon Piwi, Aub, and Armi tethering was 25-28nt in length (Fig. 3C), and the majority (60.4% (std = 1.8), 55.9% and 78.4% (std = 0.06), respectively) had a strong bias for U at the first position (Fig. 3D, Supplemental Fig. S4B), both features of piRNAs. Cloning of small RNAs isolated from immunoprecipitated Piwi complex upon Piwi or Aub tethering confirmed that they are indeed *bona-fide* piRNAs (Fig. 3E-G). Reporter piRNAs are derived only from the sense strand of the reporter (99.4% (std = 0.001), 99.4% and 99.8% (std = 0.001), respectively) and lack a bias for A at position 10 (a feature of ping-pong piRNAs) (Fig. 3D, G). This, together with the lack of complementary piRNAs, suggests that these piRNAs are formed through a ping-pong-independent mechanism. We note that neither of our tethering experiments resulted in phased small RNAs from the reporter, although we did detect phasing in these libraries when we analyzed cluster 42AB (Supplemental Fig. S4C). We conclude that tethering of several different factors such as Piwi, Aub, and Armi to the reporter in germ cells triggers piRNA biogenesis. The mechanism by which reporter piRNAs are generated does not require complementary piRNAs and ping-pong processing, and thus seems similar, if not identical, to the mechanism observed in follicular cells.

Identification of factors required for piRNA processing in germ cells

Having shown that tethering of Piwi, Aub or Armi is sufficient for piRNA biogenesis, we set out to identify factors that are required for piRNA processing of the reporter mRNA. We combined tethering of Piwi to the reporter with germline knockdown (Supplemental Fig. S5) of the respective gene, followed by small RNA cloning.

Factors that work with Piwi to induce transcriptional repression (Panx and Egg) (Donertas et al. 2013; Muerdter et al. 2013; Sienski et al. 2015; Yu et al. 2015) are not required for piRNA biogenesis upon Piwi tethering (Fig. 4A-B). Similarly, Cuff and Thoc5, components of the RDC and TREX complex, respectively, were dispensable for generation of reporter piRNAs (Fig. 4A-B). These results indicate that nuclear Piwi partners as well as RDC and TREX are not required for piRNA biogenesis triggered by Piwi recruitment.

In contrast, reporter piRNAs were strongly (7-fold) reduced upon knockdown of Zuc (Fig. 4A-B). Similarly, strong (14-fold) reduction of piRNA biogenesis was observed upon knockdown of Armi (a putative RNA helicase). These data confirm that reporter-derived small RNAs are genuine piRNAs formed through the canonical Zuc-dependent biogenesis pathway. A somewhat weaker effect on piRNA biogenesis was seen upon knockdown of other cytoplasmic piRNA processing proteins, such as the nuage components Vasa, an RNA helicase, and the Tudor-domain protein Krimper, which reduced reporter piRNA levels ~2-fold (Fig. 4A-B). The cytoplasmic piwi proteins, Aub and Ago3, were dispensable for piRNA generation from the reporter upon Piwi tethering. This result further confirms that reporter piRNAs are generated independently of the ping-pong cycle (Fig. 4A-B). Knockdown of Cuff and Ago3 resulted in an apparent increase in reporter-derived small RNAs when depicted as RPMs of all mapping reads (Fig. 4A-B), likely due to the massive loss of endogenous piRNAs mapping to the DM3 genome. Overall, our results indicate that reporter RNA is processed by the canonical piRNA processing machinery, which includes Zuc and Armi, and that tethering bypasses the requirement of nuclear proteins.

Tethering to Piwi leads to transcript accumulation in nuage

We used RNA-FISH to test if piRNA biogenesis triggered by tethering of Piwi leads to a change in the subcellular localization of the reporter RNA. Reporter RNA was uniformly distributed in the cytoplasm in control cells upon GFP tethering. Accumulation of the transcript in distinct perinuclear foci, which also show vasa enrichment, was observed upon Piwi tethering, indicating that Piwi recruits the bound RNA to nuage (Fig. 5A). Thus, induction of piRNA biogenesis from a heterologous reporter correlates with sequestration of the reporter mRNA to nuage.

Discussion

Compared to the processing of other classes of small RNAs, piRNA biogenesis is very complex and requires numerous proteins with, to date, unspecified roles in piRNA processing. How the precursor RNAs are discriminated from other cellular RNA and directed to the processing machinery is also not clear. As standard genetic and biochemical approaches failed to provide clear answers to these questions we used an alternative approach, which is based on tethering of several piRNA pathway components to a heterologous reporter, to gain a better understanding of piRNA biogenesis. We found that tethering of several proteins – Piwi, Aub, and Armi – to an RNA that lacks homology to other piRNA precursors triggers its processing into piRNAs (Fig. 3A-B). The piRNAs produced by this approach have all features of canonical piRNAs: they have the correct size, a bias for U at the first nucleotide position, and they are loaded into Piwi (Fig. 3C-G). Biogenesis of these piRNAs is mediated by the canonical processing machinery (Fig. 4).

Experimental approach to dissect the piRNA biogenesis pathway

Most of the numerous proteins required for piRNA biogenesis were identified through genetic screens (Handler et al. 2011; Handler et al. 2013; Muerdter et al. 2013). Mutations in these genes diminish or eliminate mature piRNAs; however, it was difficult to assign specific molecular functions to these proteins, largely due to the fact that we do not know intermediate steps of piRNA processing that could be analyzed in these mutants. For example, the nuclear RDC and TREX complexes are required for piRNA biogenesis and were proposed to direct pre-piRNAs to the processing machinery

(Klattenhoff et al. 2009; Pane et al. 2011; Le Thomas et al. 2014; Mohn et al. 2014; Zhang et al. 2014; Chen et al. 2016; Hur and Chung 2016; Hur et al. 2016); however, the molecular mechanism by which these complexes might bring precursors to the cytoplasmic machinery remains unresolved.

The approach employed in this study, tethering combined with knock-down of piRNA biogenesis factors, provides a valuable tool to dissect steps of the piRNA biogenesis pathway and discriminate proteins involved in processing and precursor localization *per se* from factors involved in upstream steps of biogenesis such as transcription and nuclear export of pre-piRNAs. As expected, processing absolutely requires Zuc, the endonuclease that mediates formation of both the 5' and the 3' end of piRNAs (Fig. 4) (Malone et al. 2009; Ipsaro et al. 2012; Nishimasu et al. 2012; Han et al. 2015; Mohn et al. 2015). While recruitment of Zuc to RNA leads to processing of the transcript, the majority of the obtained small RNAs are not piRNAs, indicating that Zuc by itself is not sufficient for efficient piRNA processing (Fig. 2). We show that another factor that is required for piRNA biogenesis upon tethering is the putative RNA helicase Armi (Fig. 4). Thus, efficient piRNA processing requires the interaction between Zuc and other proteins such as Armi.

We found that components of the nuclear RDC and TREX complexes, which associate with chromatin of piRNA clusters and nascent pre-piRNAs respectively, are not required for piRNA processing upon tethering (Fig. 4), suggesting that their requirement for piRNA biogenesis can be by-passed. Previous studies show that both RDC and TREX are required for transcription of pre-piRNAs (Klattenhoff et al. 2009; Pane et al. 2011; Le Thomas et al. 2014; Mohn et al. 2014; Zhang et al. 2014; Chen et al. 2016; Hur and Chung 2016; Hur et al. 2016). TREX is likely also involved in nuclear export of pre-piRNAs. It is possible that RDC and TREX are only necessary in the nucleus for transcription and export of pre-piRNAs and they do not interact with the cytoplasmic processing machinery. Alternatively, these complexes might be required to channel natural precursors to the processing machinery, but this function can be effectively by-passed by recruitment of several different proteins to the RNA. Either way, RDC and TREX are not part of the core processing machinery. The same approach can be used to assign roles to other – including yet unidentified – piRNA biogenesis factors.

Zucchini-dependent, ping-pong independent piRNA processing exists in germ cells

Tethering of several proteins to an RNA that has no complementarity to mature piRNAs triggers its processing to piRNAs. This processing is observed in both follicular and germ cells (Fig. 1,3), and is independent of pre-existing piRNAs (Supplemental Fig. S2) and ping-pong. Historically, two types of piRNA biogenesis were proposed: the ping-pong cycle, which relies on the catalytic (slicer) activity of Aub and Ago3 to generate the 5' end of new piRNAs (Brennecke et al. 2007; Gunawardane et al. 2007), and the so-called primary processing, which was thought to be independent of ping-pong (Malone et al. 2009). Recent studies suggested that in germ cells 'primary' biogenesis is triggered by initial ping-pong processing (Han et al. 2015; Mohn et al. 2015; Senti et al. 2015; Wang et al. 2015). These results suggested that 'primary' piRNA biogenesis is actually downstream of ping-pong, or 'secondary' biogenesis, making these titles obsolete. Our data shows that piRNA biogenesis that is independent of the ping-pong machinery can be achieved – at least by the artificial tethering approach – in germ cells. Therefore, both germ cells and somatic follicular cells (which lack the ping-pong machinery (Malone et al. 2009)) possess the same machinery, which allows them to process piRNAs in a ping-pong independent fashion. We propose that Aub and Ago3 might promote Zuc-dependent processing of piRNA precursors through sequestration of these RNA to a specific cellular compartment rather than through its slicing activity (see below).

The sequestration to nuage might explain selection of substrates for piRNA processing

How piRNA precursors are recognized by the piRNA processing machinery is still unresolved. The core piRNA processing machinery – composed of Zuc and Armi – is similar in both cell types, suggesting that a common principle for precursor selection should exist. piRNAs could be generated from artificial sequences inserted into piRNA clusters (Muerdter et al. 2012). piRNA processing from genic and tRNA transcripts was also observed – correlating with transcript abundance – in Kc167 cells. Both of these observations argue against the need for specific sequence/structure motifs in precursor transcripts (Vrettos et al. 2017). We propose that sequestration of RNA into a distinct cellular compartment might be the central principle for precursor selection that is shared by germline and follicular cells. Based on our proposed model any RNA that is localized

to the processing compartment will be processed to piRNAs in a sequence-independent fashion (Fig. 5B).

Several lines of evidence support this hypothesis. First, efficient piRNA biogenesis can be triggered by recruitment of several different proteins to RNA, which have little in common except for their localization to nuage. Some of these factors – such as Aub – localize to nuage in steady state, while others localize transiently – like Piwi on its way to the nucleus (Lim and Kai 2007; Le Thomas et al. 2013; Huang et al. 2014; Webster et al. 2015). The only compartment where both Aub and Piwi are present together is nuage (Malone et al. 2009). Recruitment of Zuc – the central player in the processing according to both previous results and our data – does not lead to efficient processing into piRNAs (Fig. 2). Zuc associates with mitochondria and therefore is localized throughout the cytoplasm and not exclusively to nuage (Handler et al. 2013). Second, we found that even though Zuc-dependent piRNA biogenesis does not require ping-pong, biogenesis is reduced upon knockdown of Vasa and Krimp, proteins that localize to nuage and previously were implicated in ping-pong (Fig. 4) (Xiol et al. 2014; Webster et al. 2015). Knockdown of Vasa disrupts the overall structure of nuage (Liang et al. 1994); this might be true for knockdown of Krimp as well, making biogenesis that relies on physical integrity of this compartment less efficient. Finally, we observe that reporter RNA – normally distributed throughout the cytoplasm – is localized to nuage upon tethering (Fig. 5A). A study published during the revision of this manuscript showed that artificial reporter tethering to Armitage in follicular cells, where Armi localizes to Yb bodies, leads to reporter processing into piRNAs. Mutating either the helicase or the ATP binding domain of Armi delocalizes it from Yb bodies and, consistent with our results, reporter tethering to these mutant (and thus delocialized) versions of Armi greatly reduced reporter processing into piRNAs (Pandey et al. 2017). Thus, it seems that processing directly correlates with localization of the substrate to nuage.

How are pre-piRNAs recruited to nuage in germ cells and to Yb bodies in follicular cells? In our experiments, artificial tethering of the heterologous RNA to Piwi was sufficient to sequester it to nuage. Natural piRNA substrates might be recruited to Yb bodies through RNA-binding proteins that recognize specific motifs in RNA sequences,

such as the sequences identified in *traffic-jam* or *Flamenco* (Homolka et al. 2015; Ishizu et al. 2015). In germ cells, pre-piRNAs might be recruited to nuage through interaction with Aub and Ago3 loaded with complementary piRNAs. Piwi-loaded piRNAs are substantially reduced in the Aub/Ago3 double mutant, which was interpreted as a requirement for slicer cleavage to trigger Zuc-dependent processing (Han et al. 2015; Mohn et al. 2015; Senti et al. 2015; Wang et al. 2015). We propose that these results can be explained by failure to sequester pre-piRNAs to nuage in the Aub/Ago3 double mutant. Expression of catalytically-impaired Ago3 and Aub, which cannot cleave precursor, but are likely able to bind and recruit it to nuage, can partially rescue Zuc-dependent processing of Piwi-associated piRNAs (Wang et al. 2015).

Compartmentalization is important for many RNA processing pathways such as splicing, rRNA maturation, etc. (Zhang et al. 1994; Henras et al. 2015). However, our hypothesized model of piRNA precursor selection goes beyond stating that compartmentalization is important for piRNA processing: while in all other pathways RNA sequence motifs are still required for processing (Zeng et al. 2005; Park et al. 2011; Barrangou and Marraffini 2014; Li and Patel 2016; Tsai and Joung 2016), indicating that localization of substrates to the processing compartments is necessary but not sufficient to trigger processing, we propose that for the piRNA pathway localization of RNA into nuage/Yb granules is both necessary and sufficient to initiate piRNA processing.

Our model of piRNA biogenesis allows unification of two seemingly different mechanisms for selection of piRNA precursors operating in germ and follicular cells. Processing induced by artificial recruitment enables separation of proteins involved in upstream steps of piRNA biogenesis from factors involved in Zuc-mediated processing. Finally, selection of piRNA precursors through recruitment to the processing machinery enables the design of artificial sequences that can be efficiently processed into piRNAs, and thus opens possibilities to use the piRNA pathway as a tool for both transcriptional and posttranscriptional gene regulation.

Methods and Materials

Fly stocks:

The sh-White, sh-AGO3, sh-Armitage, sh-Aubergine, sh-Cutoff, sh-Krimper, sh-Vasa, and sh-Zucchini lines were obtained from the Bloomington stocks center (BDSC #33623, BDSC #35232, BDSC #34789, BDSC #33728, BDSC #35182, BDSC #35230, BDSC #32434, and BDSC #35227). The GFP-Piwi line was a generous gift from the Hannon lab. The sh-SetDB1, sh-Panoramix, λN-FLAG-Piwi, λN-FLAG, and Tubulin-GFP-BoxB lines were generous gifts from the Brennecke lab. The TJ-Gal4 line was obtained from the Kyoto DGGR stocks center (DGRC # 104055). To obtain the sh-Thoc5 line, the short hairpin sequences were ligated into the pValium20 vector (Ni et al. 2011) using T4 DNA ligase from NEB (M0202), according to the manual, and then integrated into the attP2 landing site (BDSC #25710). Hairpin sequences are listed in Supplementary Table S2. UASp-λN-eGFP-Piwi, UASp-λN-eGFP-Aubergine, UASp-Armitage-HA-λN, UASp-Zucchini-HA-λN, UASp-λN-eGFP-GFP, UASp-λN-HA-Yb, UASp-λN-eGFP-Thoc5, UASp-λN-eGFP-Cutoff, UASp-λN-eGFP-Asterix, UASp-λN-eGFP-Panoramix, UASp-λN-eGFP-Egg, and UASp-mKate2-4xBoxB-K10polyA were generated in lab as previously described (Chen et al. 2016). These constructs were generated by P-element integration and driven by maternal alpha-tubulin67C-Gal4 (BDSC #7063).

Flies carrying the λN-transgenes and the alpha-tubulin67C-Gal4 driver were crossed to flies carrying the UASp-mKate2-4xBoxB-K10polyA reporter, and where applicable the respective hairpin, to obtain flies that express all components of the system in the female germline. For expression in follicular cells flies carrying the λN-transgenes and the TJ-gal4 driver were crossed to flies carrying the UASp-mKate2-4xBoxB-K10polyA reporter. Flies were put on yeast for 3 to 4 days prior to dissection and were 5-8 days old at the time of dissection.

RT-qPCR:

Total RNA was isolated from 20 ovaries with Ribozol (Amresco, N580), followed by DNase treatment with Amplification grade DNaseI (Invitrogen, 18068-015) according to the manual. Reverse transcription was carried out using Superscript III (Invitrogen) with oligo $d(T)_{15}$. qPCR was performed on a Mastercycler® ep Realplex PCR thermal

cycler machine (Eppendorf). RT-qPCR data target expression was normalized to the *rp49* locus expression. PCR primers are listed in Table S2.

Small RNA-seq:

Ovaries were dissected and total RNA was isolated with Ribozol (Amresco, N580). Small RNAs within a 19-nt to 29-nt window were isolated from 15% polyacrylamide gels from 4 μg of ovarian total RNA. For samples that were $NaIO_4$ treated, 5x Borate Buffer pH 8.6 (150mM borax, 150mM boric acid) and 200mM sodium periodate were added to the size selected small RNA, then the samples were incubated at 25C for 30 minutes. The $NaIO_4$-treated small RNA was then ethanol precipitated before proceeding to library construction. The small RNA libraries were constructed using the NEBNext Small RNA Library Prep Set for Illumina (#E7330S), according to the protocol using NEBNext Multiplex Oligos for Illumina (#E7335S). Libraries were sequenced on the Illumina HiSeq 2500 (SE 50bp reads) platform. The resulting reads were mapped against the reporter sequence and against the DM3 genome using Bowtie 0.12.7 (Langmead et al. 2009) with the following settings: " –v 0 –m 10000 --best –strata", retaining reads that map with zero mismatches. The reads were also mapped to RepBase using Bowtie 0.12.7 (Langmead et al. 2009) with the following settings: "-v 3 --best --strata". Reads that specifically mapped to the reporter sequence and not to the DM3 genome with zero mismatches were extracted from the libraries. Where applicable, 22-23nt reads or 25-28nt reads were extracted from the reads mapping exclusively to the reporter. All four sets of reads (all reads, all reads specific, 22-23nt reads specific, and 25-28nt reads specific) were plotted along the reporter sequence normalized to the number of total reads or 25-28nt reads mapping to the DM3 genome respectively using IGV 2.3.68 (Robinson et al. 2011). Reads that specifically mapped to the 42AB locus (chr2R: 2144349-2386719), and not to the rest of the DM3 genome, were extracted from the libraries. Weblogos were generated using an R script that utilizes the seqLogo library (Bembom 2016) to display the nucleotide frequency occurring at each position of the inputted reads calculated within a homemade bash script. Sequencing data is summarized in Table S1. High-throughput sequencing data for small RNA-seq experiments are available through Gene Expression Omnibus: GSE102961.

IP Small RNA-Seq:

Ovaries (~100/IP) from flies expressing UASp-λN-eGFP-Piwi, UASp-λN-eGFP-Aubergine, or UASp-Zucchini-HA-λN with UASp-GFP-Piwi, and the UASp-mKate2-4xBoxB-K10polyA reporter line driven by maternal alpha-tubulin67C-Gal4 were dissected and lysed on ice in 250µl lysis buffer (30mM Hepes-KOH pH7.4, 100mM KOAc, 2mM $Mg(OAc)_2$, 5mM DTT, 0.5% [v/v] NP40, proteinase inhibitor (Roche, 11836170001), RNasin Plus (Promega, N2611). Lysate was dounced and clarified by centrifugation at 4C at max speed. The supernatant was incubated with protein A/G agarose beads (Thermo fisher, 20421) pre-conjugated with rabbit polyclonal anti-GFP (Covance, affinity-purified in our laboratory) or anti-Piwi (Brennecke et al. 2007) at 4C for 2 hours. The immunoprecipitation and RNA isolation was carried out as previously described (Vagin et al. 2006). A fifth of the RNA was CIP (NEB, M0290S) treated in NEB Buffer #3 (NEB, B7003S) for 30' at 37C and then ethanol precipitated after phenol:chloroform and chloroform extraction. The CIP-treated RNA was then PNK treated with 1µl 10x T4 Polynucleotide Kinase Buffer (NEB, B0201S), 2µl γ-P^{32} ATP (PerkinElmer, BLU502A250UC), and 1µl T4 Polynucleotide Kinase (NEB, M0201S) for 45' at 37C. The CIP-and-PNK-treated RNA was added back to the remainder of the RNA isolated from the IP. Size selection, library preparation, and analysis were performed as described above, except that fragments were gel extracted based on labeled IP material.

RNA-FISH and Immunofluorescence:

Probes specific to the mKate2 transcript were designed and ordered from Stellaris (Biosearch Technologies) and resuspended in 200µl DEPC-treated TE buffer (10mM Tris-HCl, 1mM EDTA, pH 8.0) and stored at -20C as stock. FISH was performed from 30-50 pairs of ovaries that were dissected and washed with ice cold DEPC-treated PBS. Prior to fixation, the ovaries were exposed to blue light for 15 minutes to quench GFP fluorescence from the tethering construct (UASp-λN-eGFP-Piwi and UASp-λN-eGFP-GFP). The ovaries were fixed with 300µl fixation solution (4% paraformaldehyde, 0.15% Triton X-100 in DEPC-treated PBS) at room temp (RT), shaking at 300 rpm and washed 3x with PBX (DEPC-treated PBS, 0.3% Triton X-100) at RT for 10 minutes each. Samples were dehydrated in 500µl 70% Ethanol (in DEPC-treated water) and permeabilized overnight at 4C on a nutator. Stock probes were diluted 100x in

hybridization buffer (DEPC-treated 2x SSC, with 10% [w/v] Dextran sulfate and 10% [v/v] formamide). Following rehydration of ovaries in 500µl wash buffer (DEPC-treated 2x SSC, 10% [v/v] Formamide) for 5 minutes the diluted probe was added to the samples for overnight incubation at 37C in a hybridization chamber. Samples were washed twice with wash buffer and incubated in wash buffer twice at 37C for 30 minutes. Samples were blocked with SBX buffer (DEPC-treated 2x SSC, 1% [w/v] BSA, 0.3% Triton X-100) at RT for 2 hours and incubated with 1:100 diluted anti-Vasa (DSHB, Spradling, A.C. & Williams, D.) overnight at 4C followed by five washes with SBX buffer for 10 minutes at RT and incubation with 1:400 diluted Chicken anti-Rat IgG secondary antibody, Alexa Fluor 647 (Thermo Fisher, A21472) in SBX buffer for 4 hours at RT in the dark without agitation. Ovaries were washed 5x for 10 minutes with SSCT buffer (DEPC-treated 2xSSC, 0.1% Triton X-100) at RT and incubated with 1:5000 diluted DAPI in 2xSSC buffer for 10 minutes at RT on a nutator. Ovaries were washed 2x with 2xSSC buffer and then mounted on a glass slide using Prolong Gold antifade reagent (Thermo Fisher, P36934).

Acknowledgements

We are very grateful to Alexei Aravin for discussions and critical reading and editing of the manuscript. We thank members of the Fejes Toth and Aravin labs for discussions and suggestions and Maria Ninova for assistance with data analysis. We thank Howard Lipshitz for critical reading and comments on the manuscript. We are grateful to the Brennecke lab, the Bloomington stock center and the Kyoto stock center for fly stocks. Stocks obtained from the Bloomington Drosophila Stock Center (NIH P400D018537) and from the Drosophila Genomics Resource Center (NIH 2P40OD010949) were used in this study. We are grateful to Igor Antoschechkin at the Millard and Muriel Jacobs Genetics and Genomics Laboratory at Caltech for sequencing and the Caltech Biological Imaging Facility for support with confocal imaging. The Vasa antibody developed by Spradling, A.C. & Williams, D. was obtained from the Developmental Studies Hybridoma Bank, created by the NICHD of the NIH and maintained at The University of Iowa, Department of Biology, Iowa City, IA 52242. This work was supported by a grant from the National Institute of Health (R01GM110217) and the Ellison Medical Foundation award to KFT.

AKR is an NSF GRFP fellow. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program (DGE-1144469). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

Barrangou R, Marraffini LA. 2014. CRISPR-Cas Systems: Prokaryotes Upgrade to Adaptive Immunity. *Molecular cell* **54**: 234-244.

Bembom O. 2016. seqLogo: Sequence logos for DNA sequence alignments.

Brennecke J, Aravin AA, Stark A, Dus M, Kellis M, Sachidanandam R, Hannon GJ. 2007. Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. *Cell* **128**: 1089-1103.

Chen YC, Stuwe E, Luo Y, Ninova M, Le Thomas A, Rozhavskaya E, Li S, Vempati S, Laver JD, Patel DJ et al. 2016. Cutoff Suppresses RNA Polymerase II Termination to Ensure Expression of piRNA Precursors. *Molecular cell* **63**: 97-109.

Cox DN, Chao A, Lin H. 2000. piwi encodes a nucleoplasmic factor whose activity modulates the number and division rate of germline stem cells. *Development* **127**: 503-514.

Donertas D, Sienski G, Brennecke J. 2013. Drosophila Gtsf1 is an essential component of the Piwi-mediated transcriptional silencing complex. *Genes & development* **27**: 1693-1705.

Gunawardane LS, Saito K, Nishida KM, Miyoshi K, Kawamura Y, Nagami T, Siomi H, Siomi MC. 2007. A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in Drosophila. *Science* **315**: 1587-1590.

Han BW, Wang W, Li CJ, Weng ZP, Zamore PD. 2015. piRNA-guided transposon cleavage initiates Zucchini-dependent, phased piRNA production. *Science* **348**: 817-821.

Handler D, Meixner K, Pizka M, Lauss K, Schmied C, Gruber FS, Brennecke J. 2013. The genetic makeup of the Drosophila piRNA pathway. *Molecular cell* **50**: 762-777.

Handler D, Olivieri D, Novatchkova M, Gruber FS, Meixner K, Mechtler K, Stark A, Sachidanandam R, Brennecke J. 2011. A systematic analysis of Drosophila TUDOR domain-containing proteins identifies Vreteno and the Tdrd12 family as essential primary piRNA pathway factors. *The EMBO journal* **30**: 3977-3993.

Hay B, Jan LY, Jan YN. 1988. A protein component of Drosophila polar granules is encoded by vasa and has extensive sequence similarity to ATP-dependent helicases. *Cell* **55**: 577-587.

Henras AK, Plisson-Chastang C, O'Donohue MF, Chakraborty A, Gleizes PE. 2015. An overview of pre-ribosomal RNA processing in eukaryotes. *Wiley interdisciplinary reviews RNA* **6**: 225-242.

Homolka D, Pandey RR, Goriaux C, Brasset E, Vaury C, Sachidanandam R, Fauvarque MO, Pillai RS. 2015. PIWI Slicing and RNA Elements in Precursors Instruct Directional Primary piRNA Biogenesis. *Cell Rep* **12**: 418-428.

Huang H, Li Y, Szulwach KE, Zhang G, Jin P, Chen D. 2014. AGO3 Slicer activity regulates mitochondria-nuage localization of Armitage and piRNA amplification. *The Journal of cell biology* **206**: 217-230.

Hur JK, Chung YD. 2016. A novel model of THO/TREX loading onto target RNAs in metazoan gene expression. *BMB Rep* **49**: 355-356.

Hur JK, Luo Y, Moon S, Ninova M, Marinov GK, Chung YD, Aravin AA. 2016. Splicing-independent loading of TREX on nascent RNA is required for efficient expression of dual-strand piRNA clusters in Drosophila. *Genes & development* **30**: 840-855.

Ipsaro JJ, Haase AD, Knott SR, Joshua-Tor L, Hannon GJ. 2012. The structural biochemistry of Zucchini implicates it as a nuclease in piRNA biogenesis. *Nature* **491**: 279-283.

Ishizu H, Iwasaki YW, Hirakata S, Ozaki H, Iwasaki W, Siomi H, Siomi MC. 2015. Somatic Primary piRNA Biogenesis Driven by cis-Acting RNA Elements and trans-Acting Yb. *Cell Rep* **12**: 429-440.

Kirino Y, Kim N, de Planell-Saguer M, Khandros E, Chiorean S, Klein PS, Rigoutsos I, Jongens TA, Mourelatos Z. 2009. Arginine methylation of Piwi proteins catalysed by dPRMT5 is required for Ago3 and Aub stability. *Nat Cell Biol* **11**: 652-658.

Klattenhoff C, Xi H, Li C, Lee S, Xu J, Khurana JS, Zhang F, Schultz N, Koppetsch BS, Nowosielska A et al. 2009. The Drosophila HP1 homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters. *Cell* **138**: 1137-1149.

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**: R25.

Lasko PF, Ashburner M. 1990. Posterior localization of vasa protein correlates with, but is not sufficient for, pole cell development. *Genes & development* **4**: 905-921.

Le Thomas A, Rogers AK, Webster A, Marinov GK, Liao SE, Perkins EM, Hur JK, Aravin AA, Toth KF. 2013. Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. *Genes & development* **27**: 390-399.

Le Thomas A, Stuwe E, Li S, Du J, Marinov G, Rozhkov N, Chen YC, Luo Y, Sachidanandam R, Toth KF et al. 2014. Transgenerationally inherited piRNAs trigger piRNA biogenesis by changing the chromatin of piRNA clusters and inducing precursor processing. *Genes & development* **28**: 1667-1680.

Li SS, Patel DJ. 2016. Drosha and Dicer: Slicers cut from the same cloth. *Cell Res* **26**: 511-512.

Liang L, Diehl-Jones W, Lasko P. 1994. Localization of vasa protein to the Drosophila pole plasm is independent of its RNA-binding and helicase activities. *Development* **120**: 1201-1211.

Lim AK, Kai T. 2007. Unique germ-line organelle, nuage, functions to repress selfish genetic elements in Drosophila melanogaster. *Proceedings of the National Academy of Sciences of the United States of America* **104**: 6714-6719.

Malone CD, Brennecke J, Dus M, Stark A, McCombie WR, Sachidanandam R, Hannon GJ. 2009. Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary. *Cell* **137**: 522-535.

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

Mohn F, Sienski G, Handler D, Brennecke J. 2014. The rhino-deadlock-cutoff complex licenses noncanonical transcription of dual-strand piRNA clusters in Drosophila. *Cell* **157**: 1364-1379.

Muerdter F, Guzzardo PM, Gillis J, Luo Y, Yu Y, Chen C, Fekete R, Hannon GJ. 2013. A genome-wide RNAi screen draws a genetic framework for transposon control and primary piRNA biogenesis in Drosophila. *Molecular cell* **50**: 736-748.

Muerdter F, Olovnikov I, Molaro A, Rozhkov NV, Czech B, Gordon A, Hannon GJ, Aravin AA. 2012. Production of artificial piRNAs in flies and mice. *Rna* **18**: 42-52.

Ni JQ, Zhou R, Czech B, Liu LP, Holderbaum L, Yang-Zhou D, Shim HS, Tao R, Handler D, Karpowicz P et al. 2011. A genome-scale shRNA resource for transgenic RNAi in Drosophila. *Nat Methods* **8**: 405-407.

Nishimasu H, Ishizu H, Saito K, Fukuhara S, Kamatani MK, 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 YW, Shibuya A, Siomi H, Siomi MC, Saito K. 2013. DmGTSF1 is necessary for Piwi-piRISC-mediated transcriptional transposon silencing in the Drosophila ovary. *Genes & development* **27**: 1656-1661.

Pandey RR, Homolka D, Chen KM, Sachidanandam R, Fauvarque MO, Pillai RS. 2017. Recruitment of Armitage and Yb to a transcript triggers its phased processing into primary piRNAs in Drosophila ovaries. *PLoS genetics* **13**: e1006956.

Pane A, Jiang P, Zhao DY, Singh M, Schupbach T. 2011. The Cutoff protein regulates piRNA cluster expression and piRNA production in the Drosophila germline. *The EMBO journal* **30**: 4601-4615.

Park JE, Heo I, Tian Y, Simanshu DK, Chang H, Jee D, Patel DJ, Kim VN. 2011. Dicer recognizes the 5 ' end of RNA for efficient and accurate processing. *Nature* **475**: 201-U107.

Qi H, Watanabe T, Ku HY, Liu N, Zhong M, 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. *The Journal of biological chemistry* **286**: 3789-3797.

Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011. Integrative genomics viewer. *Nat Biotechnol* **29**: 24-26.

Senti KA, Jurczak D, Sachidanandam R, Brennecke J. 2015. piRNA-guided slicing of transposon transcripts enforces their transcriptional silencing via specifying the nuclear piRNA repertoire. *Genes & development* **29**: 1747-1762.

Sienski G, Batki J, Senti KA, Donertas D, Tirian L, Meixner K, Brennecke J. 2015. Silencio/CG9754 connects the Piwi-piRNA complex to the cellular heterochromatin machinery. *Genes & development* **29**: 2258-2271.

Todeschini AL, Teysset L, Delmarre V, Ronsseray S. 2010. The epigenetic trans-silencing effect in Drosophila involves maternally-transmitted small RNAs whose production depends on the piRNA pathway and HP1. *PLoS One* **5**: e11032.

Tsai SQ, Joung JK. 2016. Defining and improving the genome-wide specificities of CRISPR-Cas9 nucleases. *Nat Rev Genet* **17**: 300-312.

Vagin VV, Sigova A, Li C, Seitz H, Gvozdev V, Zamore PD. 2006. A distinct small RNA pathway silences selfish genetic elements in the germline. *Science* **313**: 320-324.

Vrettos N, Maragkakis M, Alexiou P, Mourelatos Z. 2017. Kc167, a widely used Drosophila cell line, contains an active primary piRNA pathway. *Rna* **23**: 108-118.

Wang W, Han BW, Tipping C, Ge DT, Zhang Z, Weng ZP, Zamore PD. 2015. Slicing and Binding by Ago3 or Aub Trigger Piwi-Bound piRNA Production by Distinct Mechanisms. *Molecular cell* **59**: 819-830.

Wang W, Yoshikawa M, Han BW, Izumi N, Tomari Y, Weng Z, Zamore PD. 2014. The initial uridine of primary piRNAs does not create the tenth adenine that Is the hallmark of secondary piRNAs. *Molecular cell* **56**: 708-716.

Webster A, Li S, Hur JK, Wachsmuth M, Bois JS, Perkins EM, Patel DJ, Aravin AA. 2015. Aub and Ago3 Are Recruited to Nuage through Two Mechanisms to Form a Ping-Pong Complex Assembled by Krimper. *Molecular cell* **59**: 564-575.

Xiol J, Spinelli P, Laussmann MA, Homolka D, Yang Z, Cora E, Coute 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.

Yu Y, Gu J, Jin Y, Luo Y, Preall JB, Ma J, Czech B, Hannon GJ. 2015. Panoramix enforces piRNA-dependent cotranscriptional silencing. *Science* **350**: 339-342.

Zeng Y, Yi R, Cullen BR. 2005. Recognition and cleavage of primary microRNA precursors by the nuclear processing enzyme Drosha. *Embo Journal* **24**: 138-148.

Zhang F, Wang J, Xu J, Zhang Z, Koppetsch BS, Schultz N, Vreven T, Meignin C, Davis I, Zamore PD et al. 2012. UAP56 couples piRNA clusters to the perinuclear transposon silencing machinery. *Cell* **151**: 871-884.

Zhang G, Taneja KL, Singer RH, Green MR. 1994. Localization of pre-mRNA splicing in mammalian nuclei. *Nature* **372**: 809-812.

Zhang Z, Wang J, Schultz N, Zhang F, Parhad SS, Tu S, Vreven T, Zamore PD, Weng Z, Theurkauf WE. 2014. The HP1 homolog rhino anchors a nuclear complex that suppresses piRNA precursor splicing. *Cell* **157**: 1353-1363.

Figures and Figure Legends

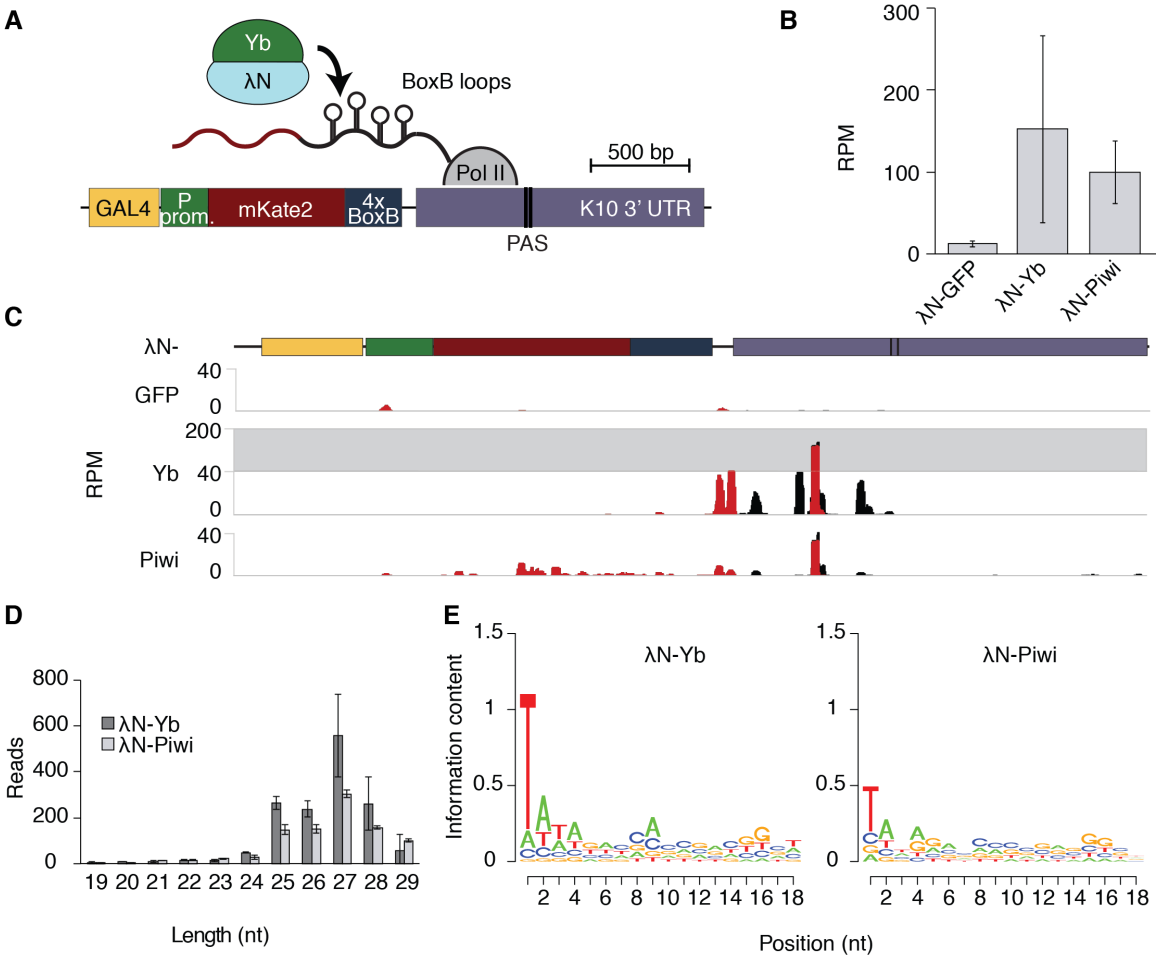

Figure 1. Yb and Piwi recruitment to RNA in follicular cells triggers piRNA production from reporter mRNA.

A) Schematic diagram of the reporter construct used to study the effect of piRNA pathway component recruitment to RNA. (PAS) PolyA signal. **B)** Recruitment of Yb and Piwi, but not GFP, to a reporter results in small RNA reads mapping to the reporter sequence. Shown are small RNA-seq reads from 19-30nt total ovarian RNA. RPMs were calculated for reads mapping exclusively to the reporter normalized to total reads mapping to the DM3 genome. Error bars indicate standard deviation between replicates (n=2). **C)** The small RNA reads produced upon tethering Yb or Piwi map to the reporter. Shown are profiles for all reads (black) and reporter-specific reads (red) mapping to the reporter normalized to all reads mapping to the DM3. **D)** Tethering of Yb or Piwi results in production of piRNA-length reads from the reporter. Shown are size profiles of all

reads mapping to the reporter. Error bars indicate standard deviation between replicates (n=2). **E)** Reporter-derived sequences upon Yb or Piwi tethering exhibit a 1U bias. The weblogos were generated from all reads mapping to the reporter.

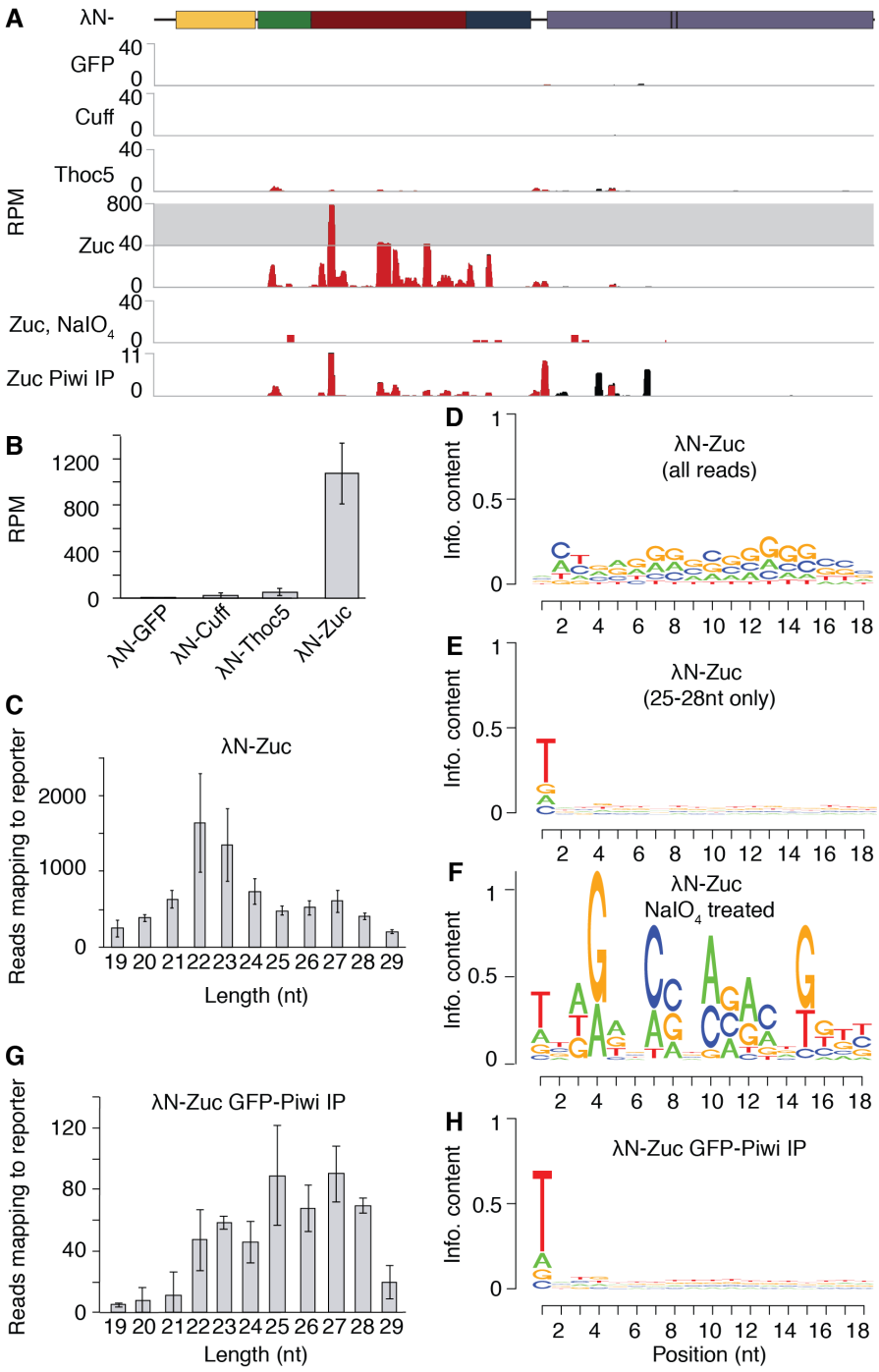

Figure 2. Recruitment of the nuclear RDC and TREX complexes, or Zucchini is not sufficient to trigger piRNA processing in germ cells.

A) Recruitment of Zuc, but not Cuff, Thoc5, or GFP to the reporter results in small RNA production. However, upon $NaIO_4$-treatment reporter-derived small RNAs generated upon Zuc tethering are lost. Shown are profiles for all reads (black) and reporter-specific

reads (red) mapping to the reporter, normalized to all reads mapping to the DM3. **B)** Normalized small RNA-seq reads (cumulative RPMs) upon germline tethering of GFP, Cuff, Thoc5, or Zuc. RPMs were calculated for 19-30nt total ovarian RNA reads mapping exclusively to the reporter normalized to total reads mapping to the DM3 genome. Error bars indicate standard deviation between replicates (n=2). **C)** Zuc tethering leads to production of two classes of small RNAs. Shown is the size profile of all reads mapping to the reporter. Error bars indicate standard deviation between replicates (n=2). **D)** Reporter-derived sequences upon Zuc tethering do not show any nucleotide bias. The web logo was generated from all reads mapping to the reporter. **E)** 25-28nt (piRNA-sized) reporter sequences produced upon Zuc tethering exhibit a slight 1U bias. The web logo was generated from all 25-28nt reads mapping to the reporter. **F)** Web logo of $NaIO_4$-treated reporter-derived small RNAs upon Zuc tethering indicates moderate 1U bias but very few reads. **G)** Piwi-loaded small RNAs upon Zuc tethering are piRNA-length. Shown is the size profile of all reads mapping to the reporter. Error bars indicate standard deviation between replicates (n=2). **H)** Reporter-derived sequences upon Zuc tethering that are loaded into Piwi show a 1U bias, indicating they are *bona-fide* piRNAs. The web logo was generated from all reads mapping to the reporter.

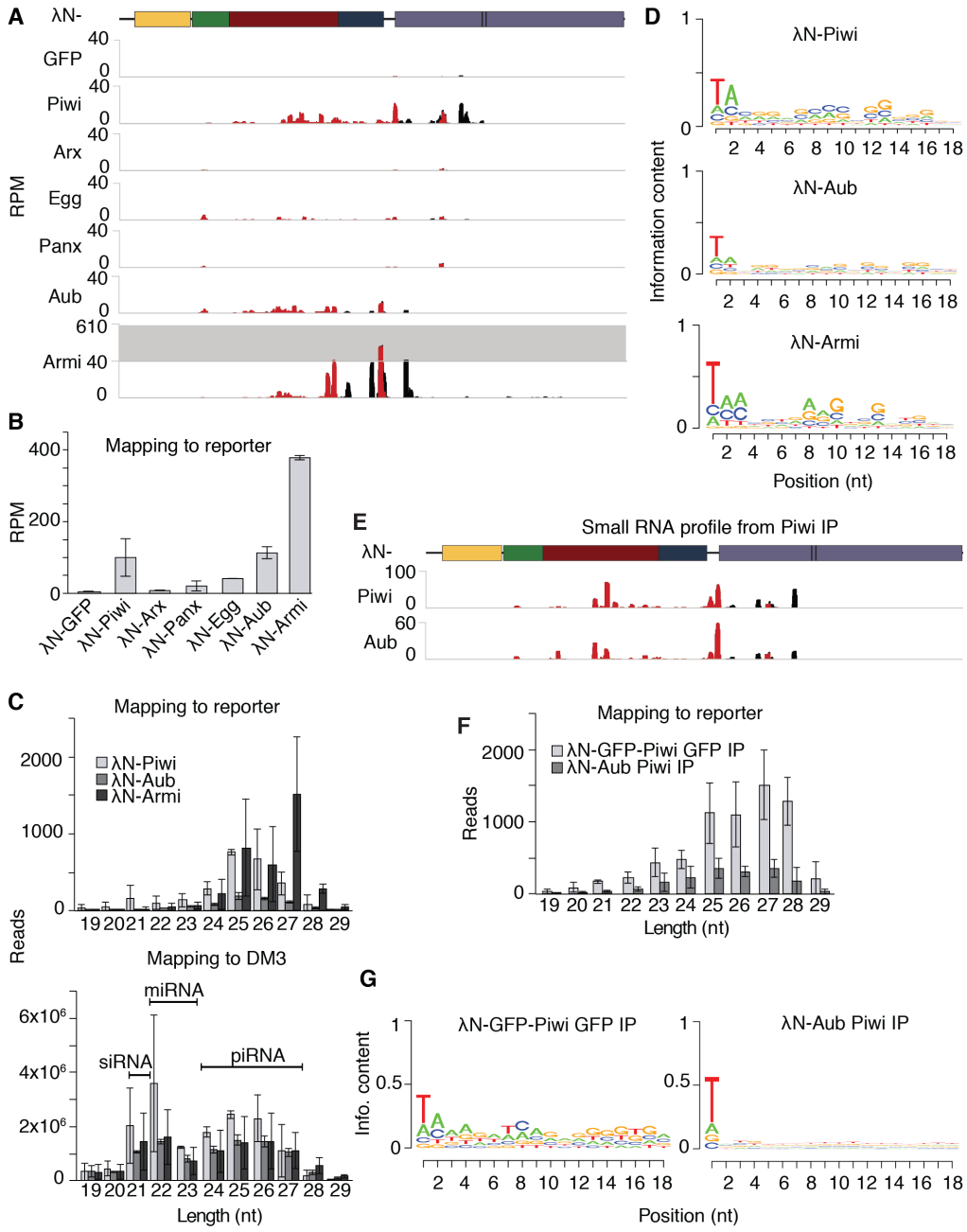

Figure 3. In germ cells, recruitment of Piwi and cytoplasmic piRNA pathway components triggers piRNA production from the reporter.

A) Tethering of Piwi, Aub, or Armi in the germline produces piRNA from the reporter, whereas tethering of the chromatin factors, Arx, Egg, and Panx does not. Shown are profiles for all reads (black) and reporter-specific reads (red) mapping to the reporter normalized to all reads mapping to the DM3. **B)** Shown are cumulative RPMs from 19-30nt total ovarian RNA. RPMs were calculated for reads mapping exclusively to the

reporter normalized to total reads mapping to the DM3 genome. Error bars indicate standard deviation between replicates (n=2). **C)** Tethering of Piwi, Aub, or Armi leads to piRNA-length reads from the reporter. Shown are size profiles of all reads mapping to the reporter (top) and – as a size selection control – to the DM3 genome (bottom) upon tethering Piwi, Aub, or Armi. Error bars indicate standard deviation between replicates (n=2). **D)** Upon recruitment of Piwi, Aub, or Armi, reporter-derived sequences show a 1U bias. The web logos were generated from all reads mapping to the reporter. **E)** Upon tethering of Piwi or Aub, small RNA sequences derived from the reporter are loaded into Piwi. Shown are profiles for all reads (black) and reporter-specific reads (red) mapping to the reporter, normalized to all reads mapping to the DM3. **F)** Piwi-loaded reporter small RNA sequences obtained upon Piwi or Aub tethering are piRNA-length. Shown are size profiles of all reads mapping to the reporter. Error bars indicate standard deviation between replicates (n=2). **G)** Piwi-loaded reporter small RNA sequences show a 1U bias. The web logos were generated from all reads mapping to the reporter.

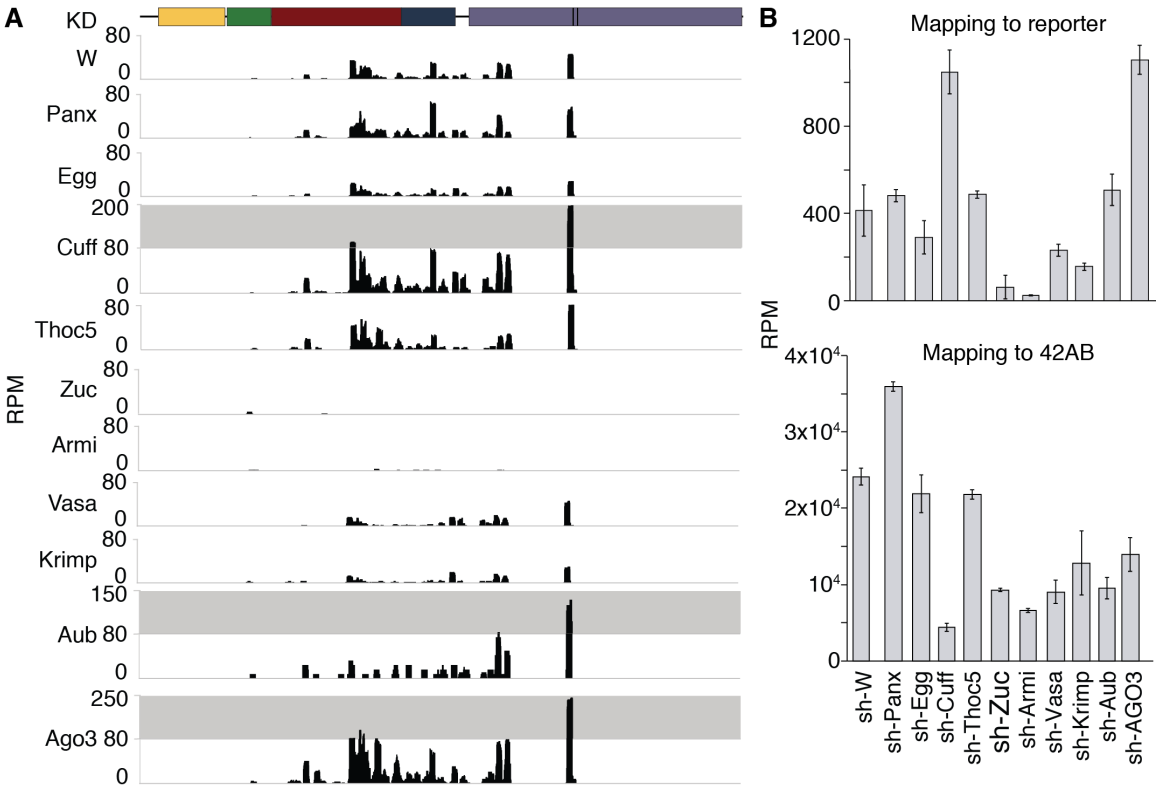

Figure 4. Chromatin factors are dispensable, but primary biogenesis factors are required for piRNA generation upon Piwi tethering.

A) Knockdown of piRNA factors concomitantly to Piwi tethering leads to loss of piRNA, only if Zuc-mediated cytoplasmic biogenesis factors are affected. Shown are small RNA profiles along the reporter for 25-28nt reads mapping exclusively to the reporter normalized to 25-28nt reads mapping to the DM3. **B)** Shown are cumulative RPMs of small RNA from 25-28nt total ovarian RNA calculated for all reads mapping to the reporter (top) and total reads uniquely mapping to cluster 42AB (bottom) normalized to total reads mapping to the DM3 genome. Knockdown of White is used as a control.

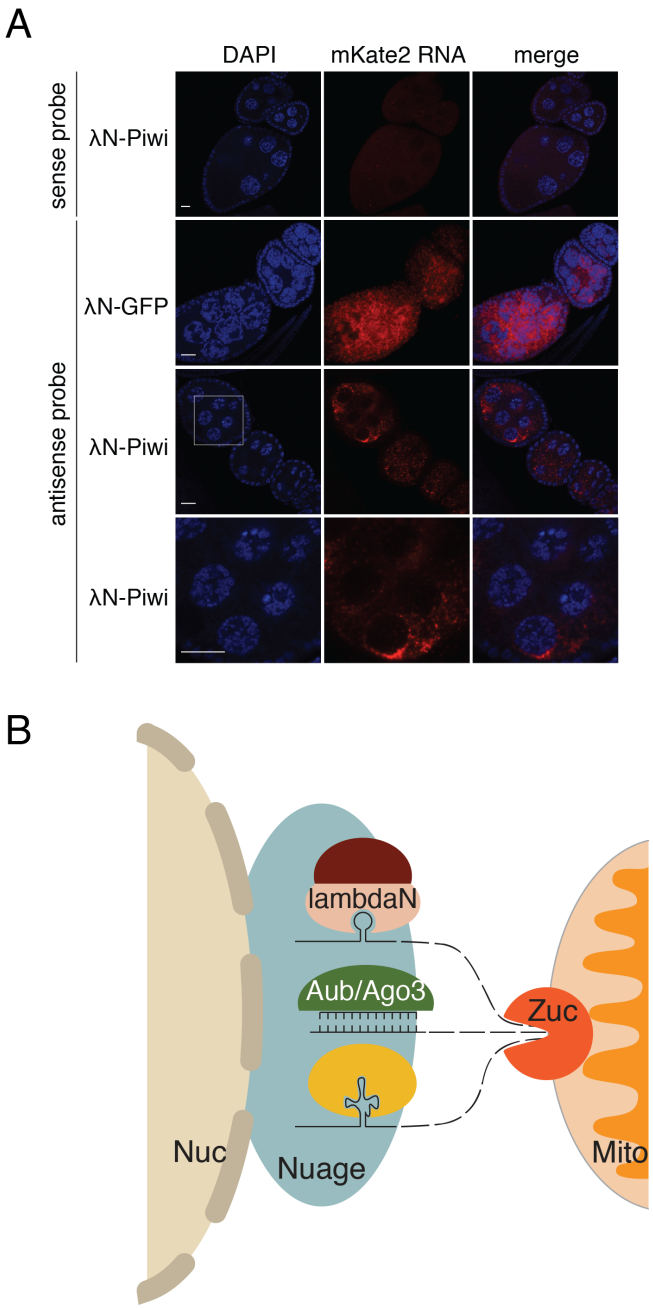

Figure 5. Tethering to Piwi leads to transcript sequestration in nuage.

A) Artificial tethering of RNA to Piwi leads to its accumulation in perinuclear granules. RNA-FISH experiments detect dispersed mKate-4BoxB reporter mRNA upon tethering to GFP (control) and perinuclear colocalization of the RNA with Vasa upon Piwi tethering. Higher magnification of the boxed-in section is shown in the 3rd row. Scale bars indicate 10μm. **B)** An overarching model of precursor selection for primary piRNA biogenesis.

Chapter V

IDENTIFYING DIRECT PROTEIN INTERACTIONS WITH PIWI USING A HETEROLOGOUS CELL CULTURE ASSAY

This work will be submitted for publication as:

Rogers, A. K. et al. "Mago and Arp6 identified as novel *Drosophila* Piwi interactors using a heterologous two-hybrid system". (Submitted)

Abstract

The piRNA pathway acts as a safeguard against transposable elements for the germline genomes of all animals. In *Drosophila melanogaster*, the piRNA pathway is composed of three piwi clade Argonuate proteins – Aubergine (Aub), Argonaute3 (Ago3), and Piwi – and their associated piwi-interacting RNAs (piRNAs). Piwi differs from Aub and Ago3 in subcellular localization and its role in the piRNA pathway. However, the exact method by which Piwi accomplishes its function remains unresolved. An essential step to understanding how Piwi achieves its function is establishing an interactome map. While a complete list of all proteins that interact with Piwi has not been established, there are many factors that have been identified genetically to act within the piRNA pathway. We developed a novel strategy to distinguish between direct and indirect interactions with Piwi and employed our strategy on a candidate list of factors. Our approach identifies direct protein interactions with Piwi using a microscopy-based assay in a heterologous cell culture system in which Piwi is stably tethered to DNA. By performing our assay in a heterologous cell culture system, we limit the possibility of identifying false positives. Our assay allowed us to discern from previously identified Piwi interactors those that directly interact with Piwi – thus further elucidating their roles within the piRNA pathway. We also utilized our approach to identify novel direct protein interactions with Piwi from a list of factors genetically identified to participate in the piRNA pathway.

Introduction

Members of the Argonaute protein family associate with distinct classes of small RNAs to transcriptionally or post-transcriptionally silence targets. The small RNA loaded provides specificity for a target, while the Argonaute protein provides the effector function to silence the target. In the germline of metazoans, the piwi clade of Argonuate proteins associates with piwi interacting RNAs (piRNAs) to form effector complexes that recognize and silence selfish mobile genetic elements called transposable elements (TEs). In *Drosophila melanogaster*, the genome encodes three piwi clade Argonaute proteins: Piwi, Aubergine (Aub), and Argonaute3 (Ago3).

All three piwi clade proteins are expressed in germline cells of *D. melanogaster* ovaries, while in the somatic follicular cells only Piwi is present (Cox et al. 2000; Harris and Macdonald 2001; Saito et al. 2006; Brennecke et al. 2007; Gunawardane et al. 2007). Piwi is loaded with piRNA in cytoplasmic, perinuclear granules termed nuage, and then the Piwi-piRNA complex is translocated to the nucleus where Piwi functions. In the nucleus, Piwi scans transcripts for a complementary target to its piRNA. Once a target has been identified, Piwi recruits the silencing machinery necessary to establish the H3K9me3 silencing mark at the site of the target (Saito et al. 2009; Olivieri et al. 2010; Handler et al. 2011; Ishizu et al. 2011; Klenov et al. 2011; Shpiz et al. 2011; Sienski et al. 2012; Gu and Elgin 2013; Le Thomas et al. 2013; Rozhkov et al. 2013). To accomplish its function, Piwi must engage in many interactions with a distinct set of, mostly unidentified, proteins within the cytoplasmic and nuclear compartments.

In nuage Piwi must interact with the piRNA biogenesis machinery to be loaded, and once in the nucleus Piwi must interact with the silencing effector complex. Many additional factors have been genetically identified as part of the piRNA pathway. To date, only a few of the piRNA pathway factors have been biochemically dissected, with the majority of interactions demonstrated predominantly through the use of coimmunoprecipitation followed by western blotting. Further, the ability to dissect the interactome is complicated by the complexity of the interactions. While studies have been carried out to determine whether the factors that interact with Piwi are involved in loading Piwi with piRNA in nuage or recruiting the silencing machinery to target loci, it is still unclear exactly whether Piwi engages in direct or indirect biochemical interactions with these factors. The ability to

discern direct biochemical interactions with Piwi would provide important insights into understanding the sequence of events that must occur for Piwi to be loaded in the cytoplasm, translocated to the nucleus, and accomplish its function in the nucleus.

To detect protein-protein interactions, two commonly used techniques are yeast two-hybrid screening and affinity purification of complexes *in vitro* followed by mass spectrometry. Both techniques have weaknesses that must be considered before utilizing them to identify protein interactions. Previously, we performed immunoprecipitation followed by mass spectrometry of tagged-Piwi in the presence and absence of RNAse A treatment, and found that Piwi predominantly interacted with RNA-binding proteins in an RNA-dependent manner (Le Thomas et al. 2013). However, due to the bias of affinity purification (AP) or coimmunoprecipitation (co-IP) towards high affinity or highly abundant proteins, we suspect that the list of interactors we obtained may be missing less abundant factors as well as any weak or transient interactions. Since we were unable to fix our samples prior to mass spectrometry, we also suspect that the list of interactors we obtained contains unspecific hits. Furthermore, we cannot discriminate between direct and indirect protein interactions based off the mass spectrometry data. Based on FLIP data, we expect that Piwi interacts very transiently with chromatin at the site of targets (Le Thomas et al. 2013). Piwi's transient interaction with chromatin suggests that factors downstream of Piwi recognition of a target could be interacting transiently as well, resulting in the inability to capture these interactions with co-IP followed by mass spectrometry.

Yeast two-hybrid (Y2H) systems, which identify defined, direct binary interactions, are relatively inexpensive and do not rely on specialized or large equipment. When looking for direct interactions with *D. melanogaster* Piwi, both AP/MS using *D. melanogaster* ovary lysate and traditional Y2H systems are plagued with false positives – indirect interactions – generated by interactions mediated by other proteins in the piRNA pathway. We developed a fluorescence based optimized two-hybrid (2H) approach to identify biochemical interactions of Piwi. Our method is capable of detecting transient or weak protein-protein interactions *in vivo* and allows for discrimination between direct and indirect interactions with *D. melanogaster* Piwi. Our work establishes a technique that allowed us to take a candidate list of factors and discern between factors that most likely directly interact with Piwi and those that indirectly interact with Piwi (possibly mediated by

RNA or other *D. melanogaster* proteins). Furthermore, our method allowed us to identify novel direct Piwi interactors amongst factors implicated in the piRNA pathway, which we confirmed by co-IP experiments in S2 cells.

Results

A heterologous cell culture assay utilizing lacO tethering to detect direct protein interactions

To determine if factors previously identified to interact with Piwi were directly or indirectly interacting with Piwi, we employed a new method capable of detecting direct protein-protein interactions *in vivo*. Our approach utilizes fusion of a bait protein to the DNA-binding domain of lacI and a fluorescent tag (mKate2). The lacI-mKate2 fused bait protein is stably expressed in cells that contain 3 genomic loci with arrays of the lacO sequence (Jegou et al. 2009), resulting in the recruitment of the bait protein to the lacO loci which can be visualized using fluorescent microscopy as three distinct nuclear dots. Cells stably expressing the tethered bait protein are transiently transfected with prey proteins fused with a different fluorescent tag (CFP). If the bait and prey proteins interact, the prey is recruited to the nuclear lacO loci, resulting in co-localization of the mKate2 and CFP signals (Fig 1A). Tethering of the bait protein to lacO foci allows for weak or transient interactions, which may be missed by other techniques, to be identified. We utilized our system with Piwi as bait in human U2OS cells which lack other *Drosophila* germline proteins – thus favoring detection of direct interactions between Piwi and the prey proteins. When we stably express lacI-mKate2-Piwi in U2OS cells containing lacO arrays, Piwi localizes to nuclear dots, whereas transiently transfected CFP (control) is evenly distributed throughout the cell and is not enriched at lacO arrays (p-value=0.234; std=0.193; n=5) (Fig 1B). We composed a list of candidate interaction partners from factors previously identified to interact with Piwi (Fig 1C) and the top hits identified from at least two out of three RNAi screens aimed to identify proteins that are involved in transposon repression (Czech et al. 2013; Handler et al. 2013; Muerdter et al. 2013). We used our heterologous cell culture system to determine whether factors on our candidate list engaged in direct interactions with Piwi. We quantified our images using a homemade script that employs the Costes method on the nuclei, deemed regions of interest (ROI), of cells expressing both the lacI-

mKate2-Piwi dots and CFP-tagged candidates where a p-value greater than or equal to 0.95 indicates colocalization (Table S1) (Costes et al. 2004; Schindelin et al. 2012).

Piwi directly interacts with some nuage-localized factors, but does not directly interact with the core components of the piRNA processing machinery

Prior to being translocated to the nucleus, Piwi is loaded with piRNA in nuage. Using our heterologous cell culture system, we tested for biochemical interactions between Piwi and factors known to localize to nuage. CFP-tagged Hen1 (p-value=0.096; std=0.124; n=5), Gasz (p-value=0.154; std=0.344; n=5), Nibbler (Nbr) (p-value=0.00; std=0.00; n=5), Minotaur (Mino) (p-value= 0.018; std=0.040; n=5), Vreteno (Vret) (p-value=0.212; std=0.172; n=5), and BoYb (p-value=0.062; std=0.139; n=5) were not recruited to lacI-mKate2-Piwi dots (Fig S1). Furthermore, the core cytoplasmic components of the primary processing machinery, including the helicase Armitage (Armi) (p-value=0.296; std=0.362; n=5) and the endonuclease Zucchini (Zuc) (p-value=0.196; std = 0.210; n=5) were not recruited by Piwi to the lacO foci (Fig 2A, S1). In steady state, Zuc is localized to the mitochondrial membrane (Malone et al. 2009; Haase et al. 2010). We generated Zuc with a mutated predicted mitochondrial localization signal (ΔMLS-Zuc) to test whether the subcellular localization of Zuc affected the result of our assay. We found that ΔMLS-Zuc was also not recruited to the lacI-mKate2-Piwi foci (p-value=0.098; std=0.197; n=5) (Fig 2A). Previously, other labs concluded that Armi interacts with Piwi in an RNA-independent manner (Olivieri et al. 2010; Ilyin et al. 2017). We performed co-IP of Armi-HA and FLAG-Piwi from S2 cells in the presence and absence of RNAse A treatment. While Armi and Piwi are pulled down together, the majority of the interaction between Piwi and Armi is lost upon RNase A treatment (Fig 2B). Based on this result, we believe the majority of stable interactions between Armi and Piwi are mediated by RNA.

Previously identified as Piwi interactors, CFP-tagged Vasa (p-value= 0.996; std=0.009; n=5), Shutdown (Shu) (p-value=0.994; std=0.013; n=5) and Papi (p-value=1.00; std=0.00; n=5) were recruited by Piwi to lacO foci, indicating a direct biochemical interaction (Fig 2C). Squash (Squ) (p-value=0.968; std=0.066;n=5), a RNase HII-related protein involved in the effector step of the piRNA pathway (Pane et al. 2007; Haase et al. 2010), and Qin (p-value=0.986; std=0.031; n=5), which is required for heterotypic ping-

pong amplification (Zhang et al. 2011; Zhang et al. 2014a) were also found to directly interact with Piwi (Fig 2C).

We identified the RNA helicase Spindle-E (Spn-E) (p-value=0.996; std=0.005; n=5), but not the Vasa paralog Belle (Bel) (p-value=0.028; std=0.063; n=5), to be a novel interactor with Piwi (Fig 2C, S1). We used co-immunoprecipitation of tagged Piwi and Spn-E proteins in S2 cells to confirm this interaction (Fig 2D). Immunoprecipitation of FLAG-tagged Piwi readily pulled-down HA-tagged Spn-E in an RNA-independent manner.

We tested for recruitment of components of the ping-pong piRNA processing (4P) complex to lacO foci and identified Aubergine (Aub) (p-value=0.990; std=0.010; n=5), but not Argonaute3 (Ago3) (p-value=0.010; std=0.017; n=5) or Krimper (p-value=0.00; std=0.00; n=5), as a novel interactor with Piwi (Fig 2C, S1). We also tested whether CFP-tagged Piwi could be recruited to the lacI-mKate2-Piwi foci and found that Piwi (p-value= 0.162; std=0.362) could not form dimers with itself in our assay (Fig S1). We used co-immunoprecipitation of tagged Piwi, Aub, and Ago3 proteins in S2 cells to confirm their interaction (Fig 2E). Immunoprecipitation of FLAG-tagged Piwi readily pulled-down HA-tagged Aub but not HA-tagged Ago3; as a control, immunoprecipitated FLAG-tagged Ago3 pulled-down HA-tagged Aub in a predominantly RNA-dependent manner. Importantly, the interaction between Piwi and Aub is not mediated by RNA as it was also observed upon treatment of the lysate with RNase A. Overall, the heterologous cell culture assay, with novel interactions validated by co-IP, suggests that Piwi directly interacts with nuage components – Vasa, Papi, Shu, Squash, Qin, Spn-E, and Aub.

Piwi directly interacts with the exon junction complex (EJC) component Mago

The mechanism by which RNA transcripts are translocated from the nucleus to the piRNA processing machinery remains elusive. Within the list of potential direct Piwi interacting partners were several core components of the exon junction complex (EJC) and EJC accessory proteins. Nxt1 (p-value=0.024; std=0.054; n=5), RnpS1 (p-value=0.268; std=0.343; n=5), and Acinus (Acn) (p-value=0.122; std=0.267; n=5) have previously been implicated in the piRNA pathway, but did not colocalize with Piwi (Fig 3A). Mago (p-value=1.00; std=0.00; n=5) was identified as a novel, direct Piwi interacting factor (Fig 3A). We confirmed the interaction of Mago and Piwi by co-IP

followed by western blot using overexpression of FLAG-tagged Piwi and HA-tagged Mago in S2 cells. We found that the interaction is not affected upon RNAse A treatment, indicating it is not mediated by RNA (Fig 3B). Interestingly, we were also able to co-IP FLAG-tagged Piwi and HA-tagged Nxt1 in the presence and absence of RNase A treatment (Fig 3C). These results suggest that Piwi interacts with multiple components of the EJC, but, of the factors tested, only directly interacts with Mago.

We tested other factors that affect the translation, surveillance, and localization of mRNAs. Of these factors, we found that Fmr1 (p-value=0.00; std=0.00; n=5), Karyβ3 (p-value=0.154; std=0.238; n=5) and Zn72D (p-value=0.00; std=0.00; n=5) were not recruited to lacI-mKate2-Piwi foci (Fig S2). Furthermore, Piwi failed to recruit the components of the transcription export (TREX) complex, UAP56 (p-value=0.002; std=0.004; n=5) and THOC5 (p-value=0.030; std=0.051; n=5) (Fig S2).

Piwi interacts with the transcriptional silencing machinery in the nucleus through Arx and HP1a

Since in steady-state Piwi is nuclear in localization, it is not surprising that many factors previously identified as Piwi interactors are also nuclear in localization. Several of these nuclear factors are known to be involved in chromatin remodeling and establishment of repressive chromatin modifications, thus potentially functioning in Piwi-mediated transcriptional silencing. The histone methyltransferases Su(Var)3-9 (p-value=0.00; std=0.00; n=5) and Eggless (Egg) (p-value=0.060; std=0.129; n=5), – the proposed "writer" of Piwi-guided H3K9me3 at sites of transposons (Sienski et al. 2015) – were not recruited to lacO foci by Piwi (Fig S3). Su(Var)2-10 (p-value= 0.006; std=0.005; n=5) was also not recruited by lacI-mKate2-Piwi (Fig S3). Maelstrom (Mael) (p-value=0.056; std=0.114; n=5) and Panoramix (Panx) (p-value=0.222; std=0.361; n=5) were also not recruited to lacO foci (Fig S3). We found that Piwi directly interacts with Asterix (Arx) (p-value=1.00; std=0.00; n=5), a nuclear protein of yet unknown function in piRNA-mediated transcriptional repression (Donertas et al. 2013; Ohtani et al. 2013) (Fig 4). Piwi further failed to recruit the linker histone H1 (p-value=0.092; std=0.200; n=5), His2AV (p-value=0.00; std=0.00; n=5), YL-1 (p-value=0.008; std=0.018; n=5), or chromatin assembly factor 1 p105 subunit (Caf1-105) (p-value=0.136; std=0.266; n=5) to lacO foci (Fig S3). Piwi did however recruit HP1a to lacO foci (p-value=1.00; std=0.00; n=5) (Fig 4). This

data suggests there are two identified direct biochemical interactions, Arx and HP1a, which could provide potential RNA-independent "bridges" between Piwi and the effector silencing complexes involved in H3K9me3 deposition at target loci.

Another major mechanism of histone modification is mediated by Polycomb group (PcG) protein complexes. We found that Piwi does not directly interact with E(z) (p-value=0.338; std=0.368; n=5) or the core Polycomb Repressive Complex 2 (PRC2) subunit Esc (p-value=0.00; std=0.00; n=5) (Fig S3). Corto (p-value=0.486; std=0.459; n=5), which physically interacts with PcG factors (Kodjabachian et al. 1998; Lopez et al. 2001; Salvaing et al. 2003) and the PcG gene Cramped (Crm) (p-value=0.608; std= 0.380; n=5) were also not recruited to lacO foci (Fig S3). We were unable to test the interaction between Piwi and Su(z)12, another component of PRC2, due to low expression levels in *Drosophila* ovaries impeding our ability to generate an entry vector. Interestingly, Cp190 – a component of many *Drosophila* chromatin insulator complexes – directly interacts with Piwi in our assay (p-value=0.964; std=0.050; n=5) (Fig 4). The complex interplay between chromatin insulator complexes and PcG proteins has yet to be fully elucidated, but our data suggests Piwi may mediate interactions between the PcG proteins and chromatin insulators by directly interacting with Cp190.

Exclusive recruitment of the Rhino-Deadlock-Cutoff (RDC) complex to piRNA clusters may be mediated by Piwi's direct interactions with Cutoff and a novel interactor – Arp6

We tested for direct interactions between Piwi and the factors that comprise the Rhino-Deadlock-Cutoff (RDC) complex, which exclusively binds H3K9me3 marks along piRNA clusters and is required for pre-piRNA transcription (Klattenhoff et al. 2009; Pane et al. 2011; Le Thomas et al. 2014; Mohn et al. 2014; Zhang et al. 2014b; Chen et al. 2016). Rhino was not recruited to lacO foci (p-value=0.018; std=0.040; n=5) (Fig 5A). However, another RDC component, Cutoff (Cuff) (p-value=1.00; std=0.00; n=5), was recruited by Piwi (Fig 5A). Not much is known about the actin-related protein 6 (Arp6), which colocalizes with HP1 on centric heterochromatin (Kato et al. 2001), however our assay identified Arp6 as a novel direct interactor with Piwi (p-value=1.00; std=0.00; n=5) (Fig 5A). We were able to confirm this interaction by co-IP of tagged Arp6 and Piwi expressed in S2 cells in the presence and absence of RNAse A treatment (Fig 5B). Furthermore, we were able to show in S2 cells Arp6 RNA-independently co-IPs with Rhino, but does not

co-IP with HP1a (Fig 5C). This data suggests that the exclusive recruitment of the RDC complex to piRNA clusters is mediated by direct interactions between Piwi and the factors Arp6 and Cuff.

Previously, our lab showed Cuff plays an important role in the protection of pre-piRNA transcripts from degradation by the exonuclease Rat1 (Chen et al. 2016). We found that Piwi directly interacts with Rat1 (p-value=0.998; std=0.004; n=5) but not with its binding partner, Rai1 (p-value=0.438; std=0.267; n=5) (Fig 5D). This data suggests that Piwi may play a role in determining whether piRNA cluster derived transcripts will be protected for transport to nuage or will be degraded by exonucleases.

Piwi does not directly interact with factors involved in its regulation and phosphorylation

From our list of candidates, one factor had been shown previously to antagonistically regulate Piwi at the genetic and molecular levels – Tudor-SN (TSN) (Ku et al. 2016). We found that TSN (p-value=0.044; std=0.062; n=5) wasn't recruited to lacI-mKate2-Piwi foci in our assay (Fig 6). Previously, it was shown that Piwi interacts with HOP and Hsp83 in the process of canalization (Gangaraju et al. 2011). While Hsp83 was shown to be involved in phosphorylation of Piwi (Gangaraju et al. 2011), it was not recruited to lacO foci in our assay (p-value=0.00; std=0.00; n=5) (Fig 6). Furthermore, HOP, which was shown to post-translationally regulate Piwi with Hsp83(Gangaraju et al. 2011) did not colocalize with lacI-mKate2-Piwi (p-value=0.006; std=0.013; n=5) (Fig 6). It seems surprising that neither the chaperone Hsp83, nor its cofactor HOP directly interacted with Piwi in our assay; however, this data suggests that Piwi may need to undergo a conformational change – by binding a piRNA or some other mechanism – before it can be phosphorylated.

Discussion

Due to its complexity, a biochemical dissection of the piRNA pathway's interactome has been limited. In this study, we focused on elucidating the interactome of Piwi – the nuclear piwi protein responsible for identifying targets and recruiting the silencing machinery necessary to establish a repressive chromatin state at target loci (Saito et al. 2009; Olivieri et al. 2010; Handler et al. 2011; Ishizu et al. 2011; Klenov et al. 2011; Shpiz et al. 2011; Sienski et al. 2012; Gu and Elgin 2013; Le Thomas et al. 2013; Rozhkov et al. 2013). We

employed an optimized two-hybrid (2H) system to discriminate between direct and indirect interactions with Piwi.

Experimental approach to discriminate between direct and indirect interactions with Piwi

Previously, many factors were identified genetically to be involved in the piRNA pathway. Some of these factors were determined to engage in biochemical interactions with Piwi via co-IP followed by western blot. However, such experiments do not distinguish between direct biochemical interactions and indirect interactions. In addition, traditional techniques such as affinity purification followed by mass spectrometry (AP/MS) can be plagued with unspecific hits and miss detection of transient or weak interactions. We developed a heterologous cell culture system capable of detecting transient or weak direct protein-protein interactions. The read-out of our assay is observed using colocalization of fluorescently labeled candidate proteins and lacI-mKate2-Piwi dots (Fig 1A-B). At the time of our study, there were 24 factors previously identified to interact with Piwi (Fig 1C). We generated a list of candidates from these factors, as well as factors previously identified as potential piRNA pathway components in genetic screens (Czech et al. 2013; Handler et al. 2013; Muerdter et al. 2013). Interestingly, Piwi did not recruit the majority of our candidate factors to the lacO foci, suggesting that at each step of Piwi function only a small number of proteins directly interact with Piwi, while other factors' interactions with Piwi are mediated by a complex or RNA (Fig 7, Table S1). It should be noted, that while our heterologous cell culture assay does not produce false positive results due to the low likelyhood of mammalian somatic proteins mediating interactions with Piwi, we acknowledge that our assay is susceptible to false negatives. False negative results could be due to (1) the necessity of a modification made on the prey protein in *D. melanogaster* to interact with Piwi that is not made in the mammalian cells, (2) a conformational change that is made when Piwi is loaded with piRNAs or (3) the prey protein is bound to a structure that cannot enter the nucleus.

Piwi does not colocalize with core cytoplasmic components of the primary processing machinery

There are three classifications of factors involved in piRNA biogenesis: (I) factors required for primary piRNA biogenesis, (II) factors specifically involved in the ping-pong amplification cycle, and (III) factors that are necessary for all piRNA biogenesis (Olivieri

et al. 2012). Despite previously being identified as Piwi interactors, we saw that Piwi does not colocalize with the core cytoplasmic components of the primary processing machinery Zucchini (Zuc) and Armitage (Armi) (Fig 2A, S1). It has been previously accepted that Armi interacts with Piwi in an RNA-independent manner (Olivieri et al. 2010; Ilyin et al. 2017); however, we find that the majority of the interaction between Piwi and Armi, detected by co-IP of tagged proteins from S2 cells followed by western blot, is lost upon RNase A treatment, indicating their interaction is mediated by RNA (Fig 2B). We acknowledge that Zuc is anchored to the mitochondrial membrane (Malone et al. 2009; Haase et al. 2010) and thus will not enter into the nucleus; however our ΔMLS-Zuc was also unable to be recruited to the lacO foci (Fig 2A). As Gasz is predicted to recruit piRNA precursors, as well as piRNA biogenesis factors, to the mitochondrial surface for cleavage by Zuc (Cook et al. 2004; Ma et al. 2009; Haase et al. 2010; Olivieri et al. 2010; Saito et al. 2010; Handler et al. 2013; Vrettos et al. 2017), it is possible that the inability of Gasz to be recruited to lacO foci (Fig S1) could be a false negative. To completely address whether factors that localize in subcellular regions other than the nucleus directly interact with Piwi, our assay could be utilized with mutants that allow these factors to enter the nucleus to reassess their recruitment to lacO foci by lacI-mKate2-Piwi. Overall, the lack of a direct interaction between Piwi and core components of the primary piRNA processing machinery suggests that these interactions are typically RNA-mediated.

Piwi directly interacts with core components of nuage structure

While the majority of Piwi exhibits nuclear localization in steady-state, it has been proposed that Piwi is loaded with piRNA within nuage, or Yb granules in follicular cells (Saito et al. 2009; Haase et al. 2010; Olivieri et al. 2010; Handler et al. 2011; Ishizu et al. 2011; Klenov et al. 2011; Qi et al. 2011). We found the core architectural component of nuage and piRNA biogenesis group II factor, Vasa (Lim and Kai 2007; Li et al. 2009; Malone et al. 2009; Patil and Kai 2010; Kibanov et al. 2011), directly interacts with Piwi (Fig 2C). Our assay also revealed direct interactions occur between Piwi and Qin (Fig 2C), which is required for nuage assembly (Anand and Kai 2012). Furthermore, the RNA helicase and piRNA biogenesis group II factor, Spindle-E (Spn-E), was identified as a novel interactor of Piwi (Fig 2C-D). Recently, it has been shown that in the absence of Spn-E, the nuage components Aub, Ago3, and Vasa are mislocalized (Lo et al. 2016). This

implies that Spn-E is essential for proper nuage structure. In addition, it has been shown that Spn-E and Aub or Ago3 mutually depend on each other to enter nuage, whereas Spn-E and Qin mutually depend on each other for their exit of nuage, resulting in cycling of Spn-E in and out of nuage (Andress et al. 2016). The direct Piwi/Qin and Piwi/Spn-E interactions could imply that Piwi's localization to nuage, like Aub and Ago3's localization, is dependent on the proper cycling of Spn-E.

Piwi directly interacts with Papi, Shutdown and Aubergine potentially to initiate loading of piRNAs into PIWI proteins

Interestingly, we identified the piwi clade argonaute Aub, which is essential for ping-pong (Brennecke et al. 2007; Gunawardane et al. 2007), as a novel interactor of Piwi (Fig 2C, 2E). The other essential components of the ping-pong piRNA processing (4P) complex, Krimper and Ago3, were not recruited to lacO foci by Piwi (Fig 2E, S1). Recent studies have shown that piRNA-dependent processing of precursors via ping-pong also leads to phased processing of the transcript downstream of the initial cleavage site (Han et al. 2015; Mohn et al. 2015; Wang et al. 2015), thus Piwi's direct interaction with Aub might be responsible for recruiting unloaded Piwi to the ping-pong machinery within nuage to initiate phased processing and loading of piRNAs into Piwi.

Shutdown (Shu) is a group III factor as is it is required for all piRNA biogenesis (Olivieri et al. 2012). We found that Shu directly interacts with Piwi (Fig 2C), supporting the hypothesis that unloaded Piwi recruits Shu to nuage, downstream of primary and secondary piRNA biogenesis, to participate in loading of piRNAs into Piwi proteins (Olivieri et al. 2012).

Once a piRNA is loaded into Piwi, it will be trimmed and modified to form a mature piRNA. Despite group III factor Vreteno (Vret) affecting Piwi protein stability and primary piRNA biogenesis in the soma and germline (Handler et al. 2011; Zamparini et al. 2011), it was not recruited to lacO foci in our assay (Fig S1). We also found the methyltransferase Hen1, which is responsible for methylating the terminal 2' hydroxyl group of piRNAs (Horwich et al. 2007; Saito et al. 2007) as well as other factors downstream of piRNA biogenesis did not directly interact with Piwi in our assay (Fig S1). However, one factor involved in trimming some pre-piRNA transcripts, Papi (Hayashi et al. 2016), directly interacts with Piwi (Fig 2C). Piwi's direct interaction with Papi might be

responsible for recruiting other factors necessary for the trimming, and modification, of Zuc-generated pre-piRNA loaded into Piwi.

Taken together, this data suggests that unloaded Piwi may interact directly with Aub to initiate phasing of piRNA precursors and the subsequent loading of Piwi, aided by direct interactions with Shu. Then, loaded Piwi directly interacts with Papi to trim and recruit other group III factors needed to modify the loaded pre-piRNA.

Mago, a component of the exon junction complex (EJC), provides the connection between Piwi and the EJC to facilitate the export of piRNA precursors to nuage

The mechanism by which RNA transcripts are translocated from the nucleus to nuage, the site of piRNA precursor selection (Pandey et al. 2017; Rogers et al. 2017), for processing into piRNAs is still not understood. Previously, it had been proposed that the transcript export (TREX) complex component UAP56, along with Rhino, may bind piRNA precursor transcripts and facilitate their release into nuage by the nuclear pore complex (Zhang et al. 2012; Zhang et al. 2014b). When we tested the THO/TREX complex components UAP56 and THOC5, we did not find any direct interactions with Piwi (Fig S2). Our candidate list contained several components of the exon junction complex (EJC) and EJC accessory proteins. Despite previously being shown to be important for proper localization of *flam* piRNA precursors outside of the nucleus (Dennis et al. 2016), Nxt1 did not directly interact with Piwi in our assay (Fig 3A). However, Nxt1 can be pulled down with Piwi in an RNA-independent manner (Fig 3C), suggesting that Piwi interacts with Nxt1 via interactions with other EJC components. While RnpS1 and Acinus (Acn) have been implicated in mRNA quality control, pre-mRNA splicing, and transcriptional regulation (Schwerk et al. 2003), they only affect *flam* transcript intranuclear localization (Dennis et al. 2016) and did not directly interact with Piwi (Fig 3A). We did, however, identify the EJC component Mago as a novel direct Piwi interactor (Fig 3A-B). Mago was also previously shown to affect *flam* transcript export from the nucleus, but not *flam* piRNA production (Dennis et al. 2016). Taken together, this data suggests that Mago directly interacts with Piwi, allowing for the EJC to associate with bound target transcripts to be channeled into nuage for piRNA processing.

The silencing machinery is directed to target loci by Piwi via biochemical interactions with Arx and HP1a

After Piwi is loaded with a piRNA, the Piwi/piRNA-complex is translocated to the nucleus and searches for a complimentary target. Once a target has been identified, Piwi must recruit the silencing machinery necessary to establish a repressive H3K9me3 mark at the site of the target (Saito et al. 2009; Olivieri et al. 2010; Handler et al. 2011; Ishizu et al. 2011; Klenov et al. 2011; Shpiz et al. 2011; Sienski et al. 2012; Gu and Elgin 2013; Le Thomas et al. 2013; Rozhkov et al. 2013). Many factors have been implicated in Piwi-mediated transcriptional silencing, but exactly how Piwi and these factors coordinate the establishment of a repressive mark at the sites of targets is unclear. Several nuclear factors we tested are involved in chromatin remodeling and establishment of repressive chromatin modifications, thus potentially functioning in Piwi-mediated transcriptional silencing.

The histone methyltransferases Su(Var)3-9 and Eggless (Egg) – the proposed "writers" of Piwi-guided H3K9me3 at sites of transposons (Sienski et al. 2015) – were not recruited to lacO foci by Piwi (Fig S3). Based on previous results (chapter II), we believe Su(Var)2-10 may be part of the missing link between the general silencing machinery and Piwi/piRNA complex in establishing H3K9me3 at target loci. In our 2H assay, Su(Var)2-10 failed to be recruited by lacI-mKate2-Piwi (Fig S3), indicating it is not a direct link. It has been proposed that Panoramix (Panx), a protein of unknown function, acts as a scaffold between the Piwi/piRNA effector complex and the general silencing machinery at transposon loci (Sienski et al. 2015; Yu et al. 2015). However, Panx did not directly interact with Piwi in our assay (Fig S3). Interestingly, we identified another protein of yet unknown function, Asterix (Arx), which interacts with Panx (Donertas et al. 2013; Ohtani et al. 2013), to be directly interacting with Piwi (Fig 4). Taken together with data from Chapter II, this suggests that in the piRNA-induced transcriptional silencing (piRITS) complex the Piwi/piRNA complex connects with the effector silencing complex via direct interactions with Arx.

Notably, the transcriptional repressor HP1a that was previously identified to interact with Piwi using several techniques – nuclear magnetic resonance (NMR), co-IP followed by western blot, and Y2H (Brower-Toland et al. 2007; Mendez et al. 2011; Mendez et al. 2013) – did interact with Piwi in our assay (Fig 4). The factors His2AV, required for H3K9 methylation and HP1a binding of chromatin (Verni and Cenci 2015),

and YL-1, proposed to interact with His2AV (Liang et al. 2016) did not directly interact with Piwi (Fig S3). Furthermore, chromatin assembly factor 1, p105 subunit (Caf1-105) – the medium subunit of the CAF-1 that interacts with HP1a (Roelens et al. 2017), was also not recruited to lacO foci by Piwi (Fig S3). Linker histone H1, whose localization at the site of transposon loci chromatin is potentially regulated by Piwi/piRNA effector complexes and plays an interdependent role in Piwi-mediated transcriptional silencing with HP1a (Iwasaki et al. 2016), also failed to be recruited to lacO foci by Piwi (Fig S3). While His2AV, YL-1, and Caf1-105 may be important for HP1a function, it seems Piwi directly interacts with HP1a and not with its co-factors.

Another factor essential to Piwi-mediated transcriptional repression of transposons (Sienski et al. 2012; Iwasaki et al. 2016), but poorly understood, Maelstrom (Mael) did not directly interact with Piwi (Fig S3). The lack of Mael localizing to the lacO foci supports the hypothesis that Mael functions downstream of Piwi-guided deposition of H3K9me3 to increase chromatin compaction as the penultimate Piwi-piRNA silencing mechanism (Sienski et al. 2012; Iwasaki et al. 2016).

Changes in histone modifications can also be mediated by Polcomb group (PcG) protein complexes. We probed for direct interactions between Piwi and several PcG factors. We did not find direct interactions between Piwi and E(z) (Fig S3), the catalytic component of the Polycomb Repressive Complex 2 (PCR2) methyltransferase that methylates H3K27. Nor did Piwi directly interact with the core PRC2 component Esc or PcG factor Cramped (Crm) (Fig S3). Piwi's interactions with PcG complexes were also not mediated by direct interactions with Corto, which physically interacts with PcG factors (Kodjabachian et al. 1998; Lopez et al. 2001; Salvaing et al. 2003) (Fig S3). However, we identified direct interactions between Piwi and Cp190 – a core component of many *Drosophila* chromatin insulator complexes that was previously shown to interact with Piwi in an RNA-independent manner (Moshkovich et al. 2011). Currently, there is a shroud of mystery around the complex interplay between chromatin insulators and PcG complexes. It has been proposed that Piwi plays a role in inhibiting H3K37 trimethylation by preventing PcG proteins from binding their genomic targets by sequestration of the PcG complexes in the nucleoplasm (Peng et al. 2016). *Peng et al.* proposes that this sequestration is accomplished by Piwi binding to the Su(z)12 and Esc subunits of the PRC2 in a piRNA-

independent manner. Unfortunately, we were unable to test the direct interaction between Piwi and Su(z)12; however, our data indicates that Piwi's role in regulating genomic target binding by PRC2 may be mediated by Piwi's direct interaction with Cp190.

Piwi interacts with the Rhino-Deadlock-Cutoff (RDC) complex component Cutoff and a novel interactor – Arp6

Discreet genomic regions that give rise to piRNA precursors, termed piRNA clusters, require the RDC complex to bind H3K9me3 marks to allow for the clusters' transcription (Klattenhoff et al. 2009; Pane et al. 2011; Le Thomas et al. 2014; Mohn et al. 2014; Zhang et al. 2014b; Chen et al. 2016). A major question in the field is how the RDC complex exclusively binds the H3K9me3 deposited on piRNA clusters. We found that Piwi directly interacts with the RDC component Cutoff (Cuff) but not Rhino (Fig 5A). In addition, we identified Arp6 as a novel direct interactor of Piwi in our system (Fig 5A-B). We showed that Arp6 interacts with Piwi in an RNA-independent manner and can be pulled down with Rhino in an RNA-independent manner (Fig 5B-C). It was previously shown that Arp6 colocalized with HP1a (Kato et al. 2001); however, we were able to pull down Arp6 specifically with Rhino (HP1d) and not with HP1a (Fig 5C). This data suggests a direct biochemical "bridge" between Piwi and the RDC complex at piRNA clusters that is mediated by interactions with Cuff and Arp6. We propose that when Piwi/piRNA complexes recognize nascent pre-piRNA transcripts, Piwi biochemically interacts with Cuff and Arp6 resulting in the exclusive recruitment and binding of RDC complexes to the H3K9me3 deposited on piRNA clusters. This would expand the RDC complex to the PARDC complex, where Arp6 and Rhino interact and Piwi guides the ARDC complex to target binding sites via its interactions with Cuff and Arp6.

Cuff has been shown to be essential in the protection of piRNA precursor transcripts from degradation before their transportation to nuage (Chen et al. 2016). *Chen et al.* showed that Cuff and the exonuclease Rat1, and its binding partner Rai1, play antagonistic roles in the processing of piRNA precursor transcripts. We found that Piwi directly interacts with Rat1, but not Rai1 (Fig 5D). We previously showed that Piwi can sequester a transcript in nuage, resulting in the processing of the transcript by the piRNA processing machinery (Rogers et al. 2017). It is tempting to suggest that Piwi's interaction with Rat1 implies that after directing the RDC complex to piRNA clusters, Piwi is

"handed" nascent precursor transcripts and ultimately determines the transcript's fate. Piwi's interactions with Rat1 and the EJC means it could channel the transcript for degradation by Rat1 or protect the transcript and channel it ultimately to be exported to nuage for processing into piRNAs.

Our optimized 2H system in mammalian cell culture allows for detection of weak or transient direct protein-protein interactions *in vivo*. This approach utilizes fluorescent detection of recruitment of the prey protein to the lacO tethered prey protein. Our approach can be employed in a modular fashion, switching the cell type containing the lacO arrays (as necessary) and the lacI-tagged bait protein, to identify direct protein-protein interactions for any factor. We utilized our technique in two ways: (1) determining whether previously identified interactions with Piwi are direct and (2) screening candidates for potential novel direct interactions with Piwi. Our approach has allowed for elaboration of the Piwi interactome by discerning direct interactions (Fig 7), proving it to be a useful tool for extending our understanding of the piRNA pathway. Furthermore, we believe this approach will prove useful for a wide audience screening for and establishing verified direct protein-protein interactions for factors in a variety of complex pathways.

Methods and Materials

Piwi-tethering assay in U2OS cells:

The lacO U2OS cells were a generous gift from the Rippe lab (Jegou et al. 2009), and were grown in complete DMEM media (DMEM high glucose GlutaMAX supplement pyruvate (ThermoFisher, 10569010), 10% heat inactivated FBS (Gemini Bio-Products, 100-106), 1% Penicillin-Streptomycin (ThermoFisher, 15140122)). The lacI-mKate2-Piwi vector was created using pcDNA6.2/N-EmGFP-DEST (Invitrogen, V32620) and lacI and mKate2 overlap PCR products. All primers are listed in Table S2. lacO U2OS cells expressing lacI-mKate2-Piwi were selected using 5μg/mL blasticidin (Invitrogen, A11139-03). A stable cell line was established by colony picking and cells positive for mKate2 expression were sorted by the Caltech Sorting Facility, using their recommended protocol based on mKate2 expression. The CFP-destination vector was constructed by ligation of the CFP sequence from the pPCW vector (DGRC#1085) into the XbaI and NotI restriction

sites of the backbone of pcDNA6.2/N-EmGFP-DEST (Invitrogen, V32620) using T4 Ligase (NEB, M0202) according to the manual. Entry vectors containing Asterix, RnpS1, Mago, Arp6, Nxt1, His2AV, Panoramix, Caf1-105, H1, Acn, and YL-1 were generated by PCR amplification of the respective cDNA and Topo cloning into the pENTR-D-TOPO vector according to the manufacturer's suggestion (Invitrogen K240020). Expression vectors for CFP-tagged candidates were generated from the CFP-destination vector and entry vectors using LR clonase (Invitrogen, 11791100) according to the manual. LacI-mKate2-Piwi LacO U2OS cells were transfected with 2.5µg of the CFP or CFP-tagged constructs using TransIT-LT1 transfection reagent (Mirus, MIR 2305) according to the manual. 48 hours after transfection, cells were fixed with 4% paraformaldehyde in PBS at room temperature (RT) for 20 minutes then washed 3x in PBS. Cells were incubated in 1% Triton X-100 for 10 minutes at RT then stained with 1:1000 diluted 5mg/ml DAPI at RT for 5 minutes. Then the cells were washed 3x in PBS and mounted with a coverslip using Prolong Gold antifade reagent before imaging on a confocal microscope. Quantification of colocalization was done with in-house script to determine the Costes p-value (Costes et al. 2004). A region of interest (ROI) for each cell was defined by where DAPI signal overlapped with CFP and/or mKate2 signals. The Costes p-value is based off 10,000 iterations of scrambling the ROI for the channels, where a Costes p-value ≥ 0.95 indicates positive colocalization (Costes et al. 2004). For each factor, an N of 5 images was used and the average p-value and standard deviation was calculated (Table S1).

Co-immunoprecipitation with Piwi from S2 cells:

Schneider S2 cells were cultured in complete Schneider's *Drosophilia* medium (Schneider's *Drosophilia* medium (ThermoFisher, 21720024), 10% heat inactivated FBS (Gemini Bio-Products, 100-106), 1% Penicillin-Streptomycin (ThermoFisher, 15140122)). Plasmids were generated using LR clonase (Invitrogen, 11791100) according to the manual using entry vectors created in the lab (see above) and Drosophila Gateway Vector Collection (DGVC) destination vectors – pAGW (DGRC#1071), pAFW (DGRC#1111), pAHW (DGRC#1095), pAWH (DGRC#1096) – expressed by the Actin5c promoter. Cells were transfected using TransIT-LT1 transfection reagent (Mirus, MIR 2305) according to the manual with 3µg of total plasmid. For co-immunoprecipitation of proteins expressed in

S2 cells, a 3.5mm culture plate of transfected S2 cells was mechanically lysed and incubated for 20min on ice in 150µL S2 lysis buffer (20mM Tris pH7.4, 150mM KCL, 0.1% Tween-20, 0.1% NP-40 Igepal, EDTA-free Complete Protease Inhibitor Cocktail (Roche, 11836170001), 100µg/mL RNAse A). Supernatant was cleared by centrifugation at 4,000g for 20 minutes at 4°C. 10% Input sample was collected from the supernatant at concentrations of 1-3µg/µL. Anti-FLAG M2 beads (Sigma Aldrich, M8823) were blocked in 5mg/ml BSA for 10 minutes at 4°C, followed by washing in S2 lysis buffer. Blocked beads were added to the supernatant and rotated at 4°C overnight. Beads were washed three times in PBS + 0.05% Tween-20 and eluted by boiling in reducing SDS loading buffer. Antibodies used for western blots were mouse monoclonal anti-FLAG M2 (Sigma Aldrich, F3165) and mouse monoclonal anti-HA (Sigma Aldrich, H3663) at 1:10000 in 5% milk in PBST. Homemade Anti-GFP rabbit polyclonal antibody was used at 1:2000 in 5% milk in PBST.

Acknowledgements

We thank members of the Fejes Toth and Aravin labs for discussions and suggestions. We are grateful to the Rippe lab for the LacO U2OS cell line. We are grateful to Andres Collazo and Steven Wilbert at the Caltech Biological Imaging Facility for support with confocal imaging. We would also like to thank the Caltech Cell Sorting Facility. We would like to thank Justin Bois with help designing the script used for quantitative analysis. We would like to acknowledge Yelim Lee for help generating constructs during her time as a participant in the Caltech Freshman Summer Research Program. This work was supported by a grant from the National Institute of Health (R01GM110217) and the Ellison Medical Foundation award to KFT. AKR is an NSF GRFP fellow. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program (DGE-1144469). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

Anand A, Kai T. 2012. The tudor domain protein kumo is required to assemble the nuage and to generate germline piRNAs in Drosophila. *The EMBO journal* **31**: 870-882.

Andress A, Bei Y, Fonslow BR, Giri R, Wu Y, Yates JR, 3rd, Carthew RW. 2016. Spindle-E cycling between nuage and cytoplasm is controlled by Qin and PIWI proteins. *The Journal of cell biology* **213**: 201-211.

Brennecke J, Aravin AA, Stark A, Dus M, Kellis M, Sachidanandam R, Hannon GJ. 2007. Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. *Cell* **128**: 1089-1103.

Brower-Toland B, Findley SD, Jiang L, Liu L, Yin H, Dus M, Zhou P, Elgin SC, Lin H. 2007. Drosophila PIWI associates with chromatin and interacts directly with HP1a. *Genes & development* **21**: 2300-2311.

Chen YC, Stuwe E, Luo Y, Ninova M, Le Thomas A, Rozhavskaya E, Li S, Vempati S, Laver JD, Patel DJ et al. 2016. Cutoff Suppresses RNA Polymerase II Termination to Ensure Expression of piRNA Precursors. *Molecular cell* **63**: 97-109.

Cook HA, Koppetsch BS, Wu J, Theurkauf WE. 2004. The Drosophila SDE3 homolog armitage is required for oskar mRNA silencing and embryonic axis specification. *Cell* **116**: 817-829.

Costes SV, Daelemans D, Cho EH, Dobbin Z, Pavlakis G, Lockett S. 2004. Automatic and quantitative measurement of protein-protein colocalization in live cells. *Biophysical journal* **86**: 3993-4003.

Cox DN, Chao A, Lin H. 2000. piwi encodes a nucleoplasmic factor whose activity modulates the number and division rate of germline stem cells. *Development* **127**: 503-514.

Czech B, Preall JB, McGinn J, Hannon GJ. 2013. A transcriptome-wide RNAi screen in the Drosophila ovary reveals factors of the germline piRNA pathway. *Molecular cell* **50**: 749-761.

Dennis C, Brasset E, Sarkar A, Vaury C. 2016. Export of piRNA precursors by EJC triggers assembly of cytoplasmic Yb-body in Drosophila. *Nature communications* **7**: 13739.

Donertas D, Sienski G, Brennecke J. 2013. Drosophila Gtsf1 is an essential component of the Piwi-mediated transcriptional silencing complex. *Genes & development* **27**: 1693-1705.

Gangaraju VK, Yin H, Weiner MM, Wang J, Huang XA, Lin H. 2011. Drosophila Piwi functions in Hsp90-mediated suppression of phenotypic variation. *Nature genetics* **43**: 153-158.

Gu T, Elgin SC. 2013. Maternal depletion of Piwi, a component of the RNAi system, impacts heterochromatin formation in Drosophila. *PLoS genetics* **9**: e1003780.

Gunawardane LS, Saito K, Nishida KM, Miyoshi K, Kawamura Y, Nagami T, Siomi H, Siomi MC. 2007. A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in Drosophila. *Science* **315**: 1587-1590.

Haase AD, Fenoglio S, Muerdter F, Guzzardo PM, Czech B, Pappin DJ, Chen C, Gordon A, Hannon GJ. 2010. Probing the initiation and effector phases of the somatic piRNA pathway in Drosophila. *Genes & development* **24**: 2499-2504.

Han BW, Wang W, Li CJ, Weng ZP, Zamore PD. 2015. piRNA-guided transposon cleavage initiates Zucchini-dependent, phased piRNA production. *Science* **348**: 817-821.

Handler D, Meixner K, Pizka M, Lauss K, Schmied C, Gruber FS, Brennecke J. 2013. The genetic makeup of the Drosophila piRNA pathway. *Molecular cell* **50**: 762-777.

Handler D, Olivieri D, Novatchkova M, Gruber FS, Meixner K, Mechtler K, Stark A, Sachidanandam R, Brennecke J. 2011. A systematic analysis of Drosophila TUDOR domain-containing proteins identifies Vreteno and the Tdrd12 family as essential primary piRNA pathway factors. *The EMBO journal* **30**: 3977-3993.

Harris AN, Macdonald PM. 2001. Aubergine encodes a Drosophila polar granule component required for pole cell formation and related to eIF2C. *Development* **128**: 2823-2832.

Hayashi R, Schnabl J, Handler D, Mohn F, Ameres SL, Brennecke J. 2016. Genetic and mechanistic diversity of piRNA 3'-end formation. *Nature* **539**: 588-592.

Horwich MD, Li C, Matranga C, Vagin V, Farley G, Wang P, Zamore PD. 2007. The Drosophila RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC. *Current biology : CB* **17**: 1265-1272.

Ilyin AA, Ryazansky SS, Doronin SA, Olenkina OM, Mikhaleva EA, Yakushev EY, Abramov YA, Belyakin SN, Ivankin AV, Pindyurin AV et al. 2017. Piwi interacts with chromatin at nuclear pores and promiscuously binds nuclear transcripts in Drosophila ovarian somatic cells. *Nucleic acids research* **45**: 7666-7680.

Ishizu H, Nagao A, Siomi H. 2011. Gatekeepers for Piwi-piRNA complexes to enter the nucleus. *Current opinion in genetics & development* **21**: 484-490.

Iwasaki YW, Murano K, Ishizu H, Shibuya A, Iyoda Y, Siomi MC, Siomi H, Saito K. 2016. Piwi Modulates Chromatin Accessibility by Regulating Multiple Factors Including Histone H1 to Repress Transposons. *Molecular cell* **63**: 408-419.

Jegou T, Chung I, Heuvelman G, Wachsmuth M, Gorisch SM, Greulich-Bode KM, Boukamp P, Lichter P, Rippe K. 2009. Dynamics of telomeres and promyelocytic leukemia nuclear bodies in a telomerase-negative human cell line. *Molecular biology of the cell* **20**: 2070-2082.

Kato M, Sasaki M, Mizuno S, Harata M. 2001. Novel actin-related proteins in vertebrates: similarities of structure and expression pattern to Arp6 localized on Drosophila heterochromatin. *Gene* **268**: 133-140.

Kibanov MV, Egorova KS, Ryazansky SS, Sokolova OA, Kotov AA, Olenkina OM, Stolyarenko AD, Gvozdev VA, Olenina LV. 2011. A novel organelle, the piNG-body, in the nuage of Drosophila male germ cells is associated with piRNA-mediated gene silencing. *Molecular biology of the cell* **22**: 3410-3419.

Klattenhoff C, Xi H, Li C, Lee S, Xu J, Khurana JS, Zhang F, Schultz N, Koppetsch BS, Nowosielska A et al. 2009. The Drosophila HP1 homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters. *Cell* **138**: 1137-1149.

Klenov MS, Sokolova OA, Yakushev EY, Stolyarenko AD, Mikhaleva EA, Lavrov SA, Gvozdev VA. 2011. Separation of stem cell maintenance and transposon silencing functions of Piwi protein. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 18760-18765.

Kodjabachian L, Delaage M, Maurel C, Miassod R, Jacq B, Rosset R. 1998. Mutations in ccf, a novel Drosophila gene encoding a chromosomal factor, affect progression through mitosis and interact with Pc-G mutations. *The EMBO journal* **17**: 1063-1075.

Ku HY, Gangaraju VK, Qi H, Liu N, Lin H. 2016. Tudor-SN Interacts with Piwi Antagonistically in Regulating Spermatogenesis but Synergistically in Silencing Transposons in Drosophila. *PLoS genetics* **12**: e1005813.

Le Thomas A, Rogers AK, Webster A, Marinov GK, Liao SE, Perkins EM, Hur JK, Aravin AA, Toth KF. 2013. Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. *Genes & development* **27**: 390-399.

Le Thomas A, Stuwe E, Li S, Du J, Marinov G, Rozhkov N, Chen YC, Luo Y, Sachidanandam R, Toth KF et al. 2014. Transgenerationally inherited piRNAs trigger piRNA biogenesis by changing the chromatin of piRNA clusters and inducing precursor processing. *Genes & development* **28**: 1667-1680.

Li C, Vagin VV, Lee S, Xu J, Ma S, Xi H, Seitz H, Horwich MD, Syrzycka M, Honda BM et al. 2009. Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies. *Cell* **137**: 509-521.

Liang X, Shan S, Pan L, Zhao J, Ranjan A, Wang F, Zhang Z, Huang Y, Feng H, Wei D et al. 2016. Structural basis of H2A.Z recognition by SRCAP chromatin-remodeling subunit YL1. *Nature structural & molecular biology* **23**: 317-323.

Lim AK, Kai T. 2007. Unique germ-line organelle, nuage, functions to repress selfish genetic elements in Drosophila melanogaster. *Proceedings of the National Academy of Sciences of the United States of America* **104**: 6714-6719.

Lo PK, Huang YC, Poulton JS, Leake N, Palmer WH, Vera D, Xie G, Klusza S, Deng WM. 2016. RNA helicase Belle/DDX3 regulates transgene expression in Drosophila. *Developmental biology* **412**: 57-70.

Lopez A, Higuet D, Rosset R, Deutsch J, Peronnet F. 2001. corto genetically interacts with Pc-G and trx-G genes and maintains the anterior boundary of Ultrabithorax expression in Drosophila larvae. *Molecular genetics and genomics : MGG* **266**: 572-583.

Ma L, Buchold GM, Greenbaum MP, Roy A, Burns KH, Zhu H, Han DY, Harris RA, Coarfa C, Gunaratne PH et al. 2009. GASZ is essential for male meiosis and suppression of retrotransposon expression in the male germline. *PLoS genetics* **5**: e1000635.

Malone CD, Brennecke J, Dus M, Stark A, McCombie WR, Sachidanandam R, Hannon GJ. 2009. Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary. *Cell* **137**: 522-535.

Mendez DL, Kim D, Chruszcz M, Stephens GE, Minor W, Khorasanizadeh S, Elgin SC. 2011. The HP1a disordered C terminus and chromo shadow domain cooperate to select target peptide partners. *Chembiochem : a European journal of chemical biology* **12**: 1084-1096.

Mendez DL, Mandt RE, Elgin SC. 2013. Heterochromatin Protein 1a (HP1a) partner specificity is determined by critical amino acids in the chromo shadow domain and C-terminal extension. *The Journal of biological chemistry* **288**: 22315-22323.

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

Mohn F, Sienski G, Handler D, Brennecke J. 2014. The rhino-deadlock-cutoff complex licenses noncanonical transcription of dual-strand piRNA clusters in Drosophila. *Cell* **157**: 1364-1379.

Moshkovich N, Nisha P, Boyle PJ, Thompson BA, Dale RK, Lei EP. 2011. RNAi-independent role for Argonaute2 in CTCF/CP190 chromatin insulator function. *Genes & development* **25**: 1686-1701.

Muerdter F, Guzzardo PM, Gillis J, Luo Y, Yu Y, Chen C, Fekete R, Hannon GJ. 2013. A genome-wide RNAi screen draws a genetic framework for transposon control and primary piRNA biogenesis in Drosophila. *Molecular cell* **50**: 736-748.

Ohtani H, Iwasaki YW, Shibuya A, Siomi H, Siomi MC, Saito K. 2013. DmGTSF1 is necessary for Piwi-piRISC-mediated transcriptional transposon silencing in the Drosophila ovary. *Genes & development* **27**: 1656-1661.

Olivieri D, Senti KA, Subramanian S, Sachidanandam R, Brennecke J. 2012. The cochaperone shutdown defines a group of biogenesis factors essential for all piRNA populations in Drosophila. *Molecular cell* **47**: 954-969.

Olivieri D, Sykora MM, Sachidanandam R, Mechtler K, Brennecke J. 2010. An in vivo RNAi assay identifies major genetic and cellular requirements for primary piRNA biogenesis in Drosophila. *The EMBO journal* **29**: 3301-3317.

Pandey RR, Homolka D, Chen KM, Sachidanandam R, Fauvarque MO, Pillai RS. 2017. Recruitment of Armitage and Yb to a transcript triggers its phased processing into primary piRNAs in Drosophila ovaries. *PLoS genetics* **13**: e1006956.

Pane A, Jiang P, Zhao DY, Singh M, Schupbach T. 2011. The Cutoff protein regulates piRNA cluster expression and piRNA production in the Drosophila germline. *The EMBO journal* **30**: 4601-4615.

Pane A, Wehr K, Schupbach T. 2007. zucchini and squash encode two putative nucleases required for rasiRNA production in the Drosophila germline. *Developmental cell* **12**: 851-862.

Patil VS, Kai T. 2010. Repression of retroelements in Drosophila germline via piRNA pathway by the Tudor domain protein Tejas. *Current biology : CB* **20**: 724-730.

Peng JC, Valouev A, Liu N, Lin H. 2016. Piwi maintains germline stem cells and oogenesis in Drosophila through negative regulation of Polycomb group proteins. *Nature genetics* **48**: 283-291.

Qi H, Watanabe T, Ku HY, Liu N, Zhong M, 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. *The Journal of biological chemistry* **286**: 3789-3797.

Roelens B, Clemot M, Leroux-Coyau M, Klapholz B, Dostatni N. 2017. Maintenance of Heterochromatin by the Large Subunit of the CAF-1 Replication-Coupled Histone Chaperone Requires Its Interaction with HP1a Through a Conserved Motif. *Genetics* **205**: 125-137.

Rogers AK, Situ K, Perkins EM, Toth KF. 2017. Zucchini-dependent piRNA processing is triggered by recruitment to the cytoplasmic processing machinery. *Genes & development* **31**: 1858-1869.

Rozhkov NV, Hammell M, Hannon GJ. 2013. Multiple roles for Piwi in silencing Drosophila transposons. *Genes & development* **27**: 400-412.

Saito K, Inagaki S, Mituyama T, Kawamura Y, Ono Y, Sakota E, Kotani H, Asai K, Siomi H, Siomi MC. 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 KM, Siomi H, Siomi MC. 2010. Roles for the Yb body components Armitage and Yb in primary piRNA biogenesis in Drosophila. *Genes & development* **24**: 2493-2498.

Saito K, Nishida KM, Mori T, Kawamura Y, Miyoshi K, Nagami T, Siomi H, Siomi MC. 2006. Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the Drosophila genome. *Genes & development* **20**: 2214-2222.

Saito K, Sakaguchi Y, Suzuki T, Suzuki T, Siomi H, Siomi MC. 2007. Pimet, the Drosophila homolog of HEN1, mediates 2'-O-methylation of Piwi- interacting RNAs at their 3' ends. *Genes & development* **21**: 1603-1608.

Salvaing J, Lopez A, Boivin A, Deutsch JS, Peronnet F. 2003. The Drosophila Corto protein interacts with Polycomb-group proteins and the GAGA factor. *Nucleic acids research* **31**: 2873-2882.

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B et al. 2012. Fiji: an open-source platform for biological-image analysis. *Nature methods* **9**: 676-682.

Schwerk C, Prasad J, Degenhardt K, Erdjument-Bromage H, White E, Tempst P, Kidd VJ, Manley JL, Lahti JM, Reinberg D. 2003. ASAP, a novel protein complex involved in RNA processing and apoptosis. *Molecular and cellular biology* **23**: 2981-2990.

Shpiz S, Olovnikov I, Sergeeva A, Lavrov S, Abramov Y, Savitsky M, Kalmykova A. 2011. Mechanism of the piRNA-mediated silencing of Drosophila telomeric retrotransposons. *Nucleic acids research* **39**: 8703-8711.

Sienski G, Batki J, Senti KA, Donertas D, Tirian L, Meixner K, Brennecke J. 2015. Silencio/CG9754 connects the Piwi-piRNA complex to the cellular heterochromatin machinery. *Genes & development* **29**: 2258-2271.

Sienski G, Donertas D, Brennecke J. 2012. Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression. *Cell* **151**: 964-980.

Verni F, Cenci G. 2015. The Drosophila histone variant H2A.V works in concert with HP1 to promote kinetochore-driven microtubule formation. *Cell cycle* **14**: 577-588.

Vrettos N, Maragkakis M, Alexiou P, Mourelatos Z. 2017. Kc167, a widely used Drosophila cell line, contains an active primary piRNA pathway. *Rna* **23**: 108-118.

Wang W, Han BW, Tipping C, Ge DT, Zhang Z, Weng ZP, Zamore PD. 2015. Slicing and Binding by Ago3 or Aub Trigger Piwi-Bound piRNA Production by Distinct Mechanisms. *Molecular cell* **59**: 819-830.

Yu Y, Gu J, Jin Y, Luo Y, Preall JB, Ma J, Czech B, Hannon GJ. 2015. Panoramix enforces piRNA-dependent cotranscriptional silencing. *Science* **350**: 339-342.

Zamparini AL, Davis MY, Malone CD, Vieira E, Zavadil J, Sachidanandam R, Hannon GJ, Lehmann R. 2011. Vreteno, a gonad-specific protein, is essential for germline development and primary piRNA biogenesis in Drosophila. *Development* **138**: 4039-4050.

Zhang F, Wang J, Xu J, Zhang Z, Koppetsch BS, Schultz N, Vreven T, Meignin C, Davis I, Zamore PD et al. 2012. UAP56 couples piRNA clusters to the perinuclear transposon silencing machinery. *Cell* **151**: 871-884.

Zhang Z, Koppetsch BS, Wang J, Tipping C, Weng Z, Theurkauf WE, Zamore PD. 2014a. Antisense piRNA amplification, but not piRNA production or nuage assembly, requires the Tudor-domain protein Qin. *The EMBO journal* **33**: 536-539.

Zhang Z, Wang J, Schultz N, Zhang F, Parhad SS, Tu S, Vreven T, Zamore PD, Weng Z, Theurkauf WE. 2014b. The HP1 homolog rhino anchors a nuclear complex that suppresses piRNA precursor splicing. *Cell* **157**: 1353-1363.

Zhang Z, Xu J, Koppetsch BS, Wang J, Tipping C, Ma S, Weng Z, Theurkauf WE, Zamore PD. 2011. Heterotypic piRNA Ping-Pong requires qin, a protein with both E3 ligase and Tudor domains. *Molecular cell* **44**: 572-584.

Figures and Figure Legends

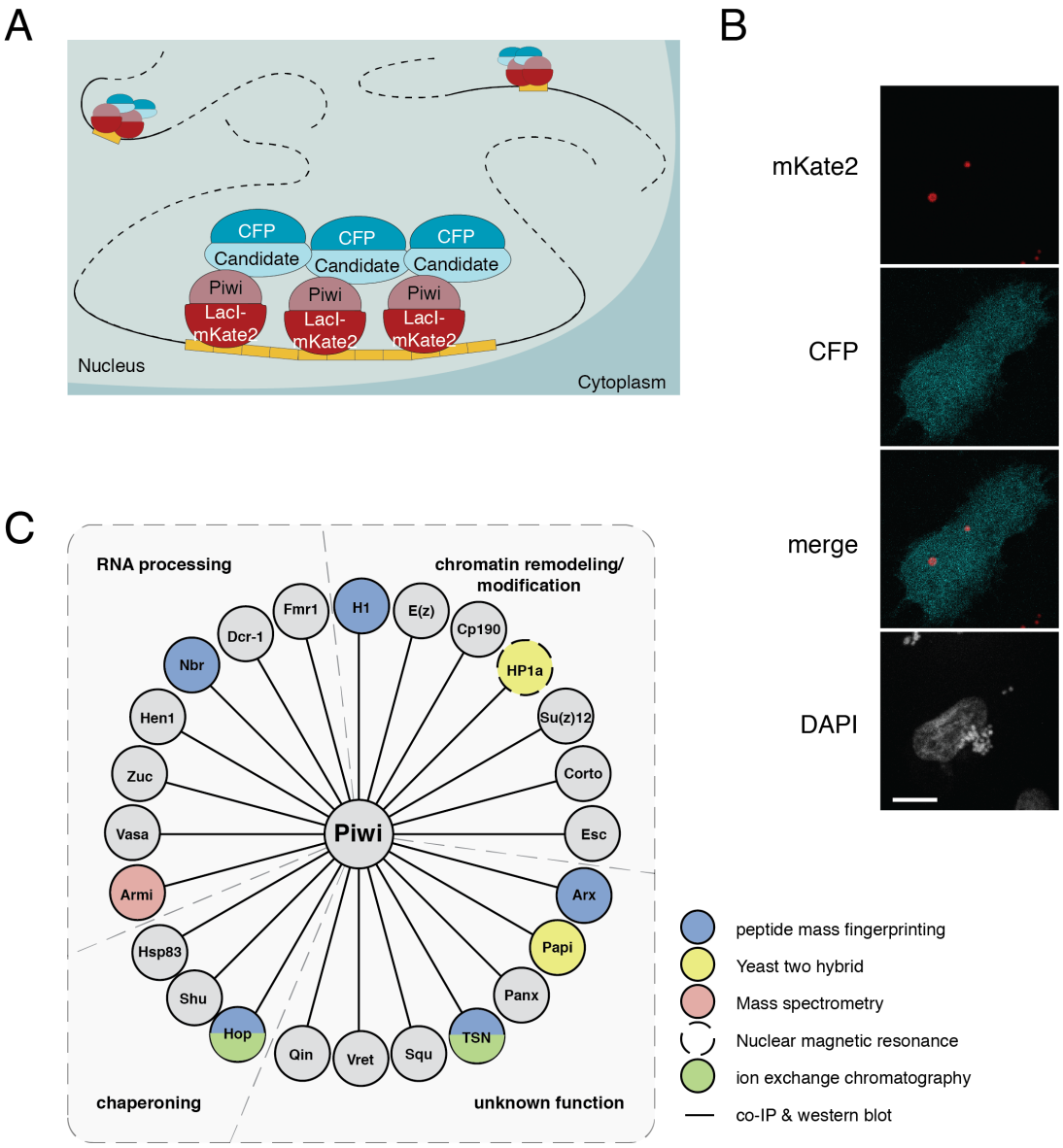

Figure 1. A heterologous cell culture assay to distinguish between direct and indirect protein-protein interactions with *Drosophila melanogaster* Piwi.

A) Schema of a heterologous cell culture assay to identify weak or transient interactions with Piwi. Human U2OS cells have three arrays composed of lacO repeats inserted in their genome. The bait protein, a fusion of the DNA binding domain of lacI with a fluorescent protein (mKate2) and Piwi, is tethered to the lacO arrays, resulting in distinct nuclear dots. The cells are transfected with CFP-tagged prey proteins. Colocalization of the CFP and mKate2 signals indicates a direct interaction between the bait and prey proteins. **B)** CFP alone does not colocalize with lacI-mKate2-Piwi. Representative confocal images of lacI-

mKate2-Piwi U2OS cells transfected with CFP (p-value=0.234; std=0.193; n=5). Scale bar represents 10μm. **C)** A map of factors previously identified to interact with *D. melanogaster* Piwi in the piRNA pathway. It is not known which of these interactions with Piwi are direct or indirect (mediated by RNA or other proteins). The techniques by which these interactions have been determined are marked according to the legend.

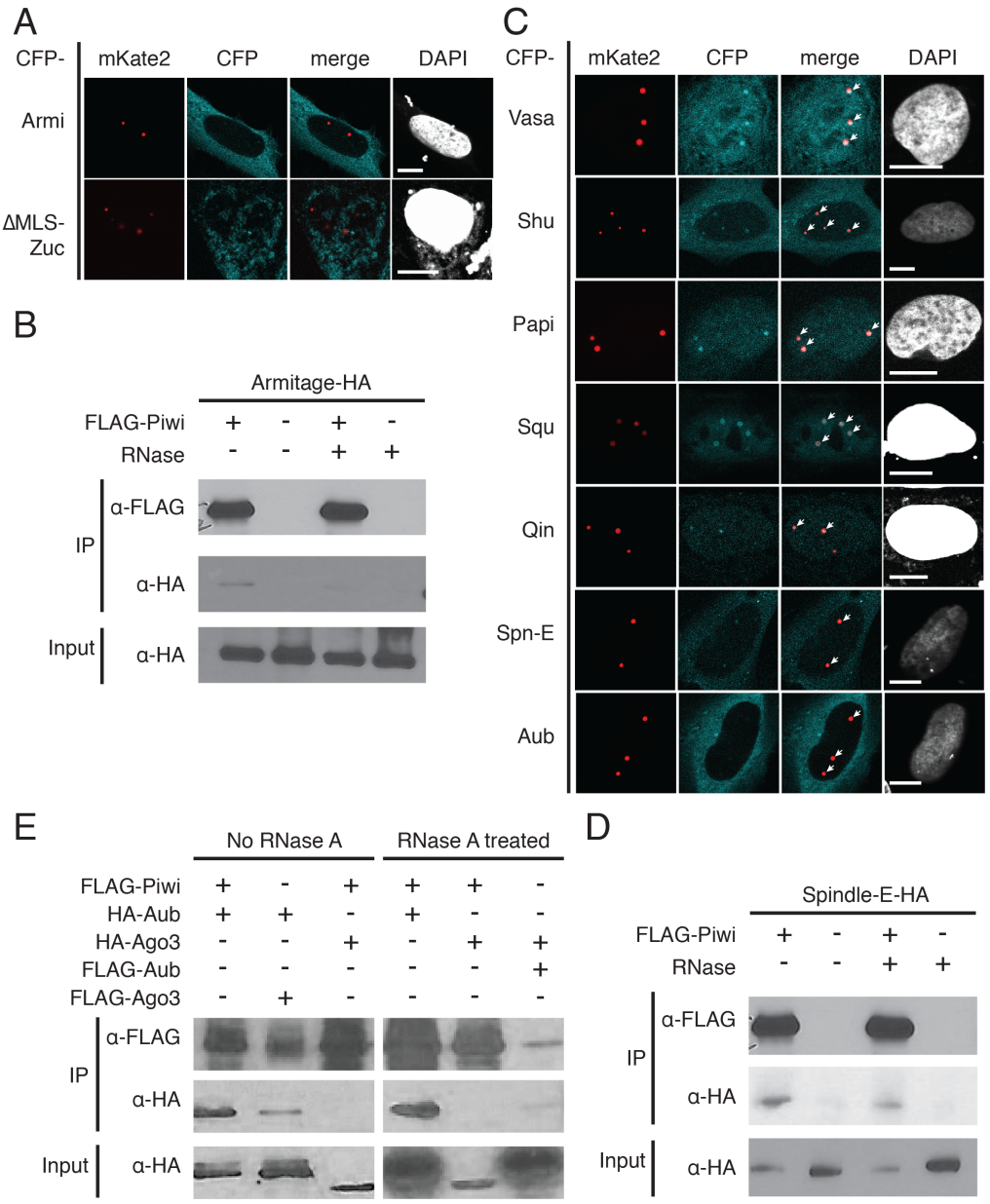

Figure 2. Piwi directly interacts with nuage components Vasa, Shutdown, Papi, Squash, Qin, Spindle-E, and Aubergine.

A) Piwi does not directly interact with core components of the primary piRNA processing machinery. Armitage (Armi) is not recruited by lacI-mKate2-Piwi (p-value=0.296; std=0.362; n=5). In steady state, Zucchini (Zuc) is localized to the mitochondrial membrane. Zuc with a mutated predicted mitochondrial localization signal (ΔMLS-Zuc) is not recruited to lacO foci by Piwi (p-value=0.098; std=0.197; n=5). Representative images are shown. Scale bar represents 10μm. **B)** Piwi and Armi interactions are RNA-dependent.

Co-immunoprecipitation **(**Co-IP) of HA-tagged Armi and FLAG-tagged Piwi in S2 cells indicates that the Piwi/Armi interaction is depleted upon RNAse A treatment. **C)** Piwi does interact with some nuage localized factors. Vasa (p-value= 0.996; std=0.009; n=5), Shutdown (Shu) (p-value=0.994; std=0.013; n=5), Papi (p-value=1.00; std=0.00; n=5), Squash (Squ) (p-value=0.968; std=0.066; n=5), Qin (p-value=0.986; std=0.031; n=5), Spindle-E (Spn-E) (p-value=0.996; std=0.005; n=5), and Aubergine (Aub) (p-value=0.990; std=0.01; n=5) were recruited by lacI-mKate2-Piwi. Representative images are shown. White arrows indicate sites of recruitment of CFP-tagged candidates to lacI-mKate2-Piwi foci. Scale bars represent 10μm. **D)** Piwi and Spn-E interactions are RNA-independent. Co-IP of HA-tagged Spn-E and FLAG-tagged Piwi in S2 cells indicates that the Piwi/Spn-E interaction is unaffected by RNAse A treatment. **E)** Co-IP of Aub with Piwi from S2 cells is RNA-independent. The third *D.melanogaster* piwi clade Argonaute protein, Argonuate3 (Ago3), does not co-IP with Piwi, but does co-IP with Aub in an RNA-dependent manner.

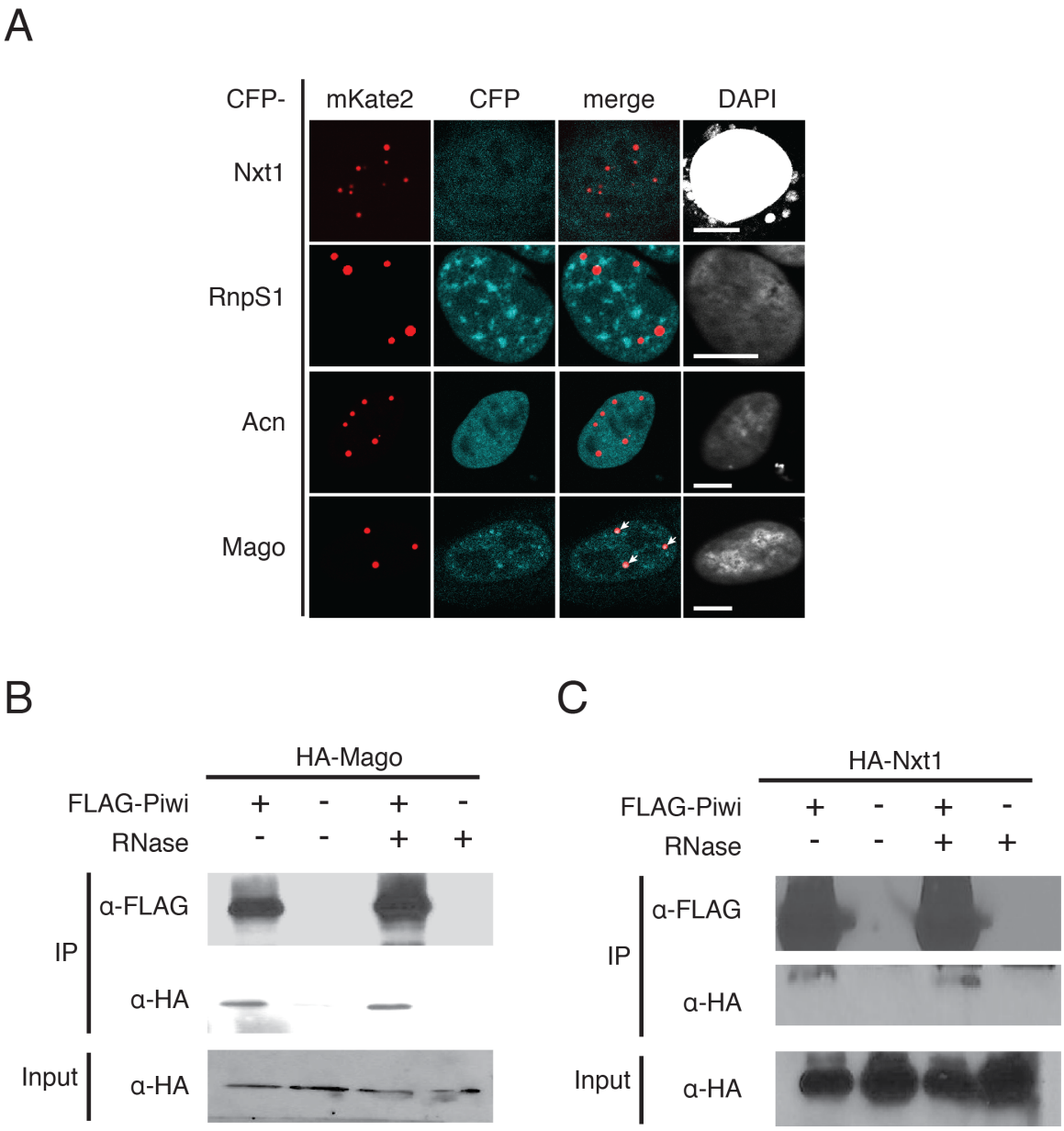

Figure 3. Piwi directly interacts with the exon junction complex (EJC) component Mago.

A) Exportin Nxt1 (p-value=0.024 ; std=0.054 ; n=5) and EJC components RnpS1 (p-value=0.268 ; std=0.343 ; n=5) and Acinus (Acn) (p-value=0.122 ; std=0.267 ; n=5) were not recruited to lacI-mKate2-Piwi foci. Piwi interacts with the EJC component Mago (p-value=1.00; std=0.00 ; n=5). Representative confocal images are shown. White arrows indicate colocalization. Scale bars represent 10μm. **B)** Piwi and Mago interactions are RNA-independent. Co-IP of HA-tagged Mago and FLAG-tagged Piwi in S2 cells indicates that the Piwi/Mago interaction is unaffected by RNAse A treatment. **C)** Piwi's interaction

with Nxt1 is likely mediated by a protein complex. Piwi and Nxt1 interactions are RNA-independent. Co-IP of HA-tagged Nxt1 and FLAG-tagged Piwi in S2 cells indicates that the Piwi/Nxt1 interaction is unaffected by RNAse A treatment.

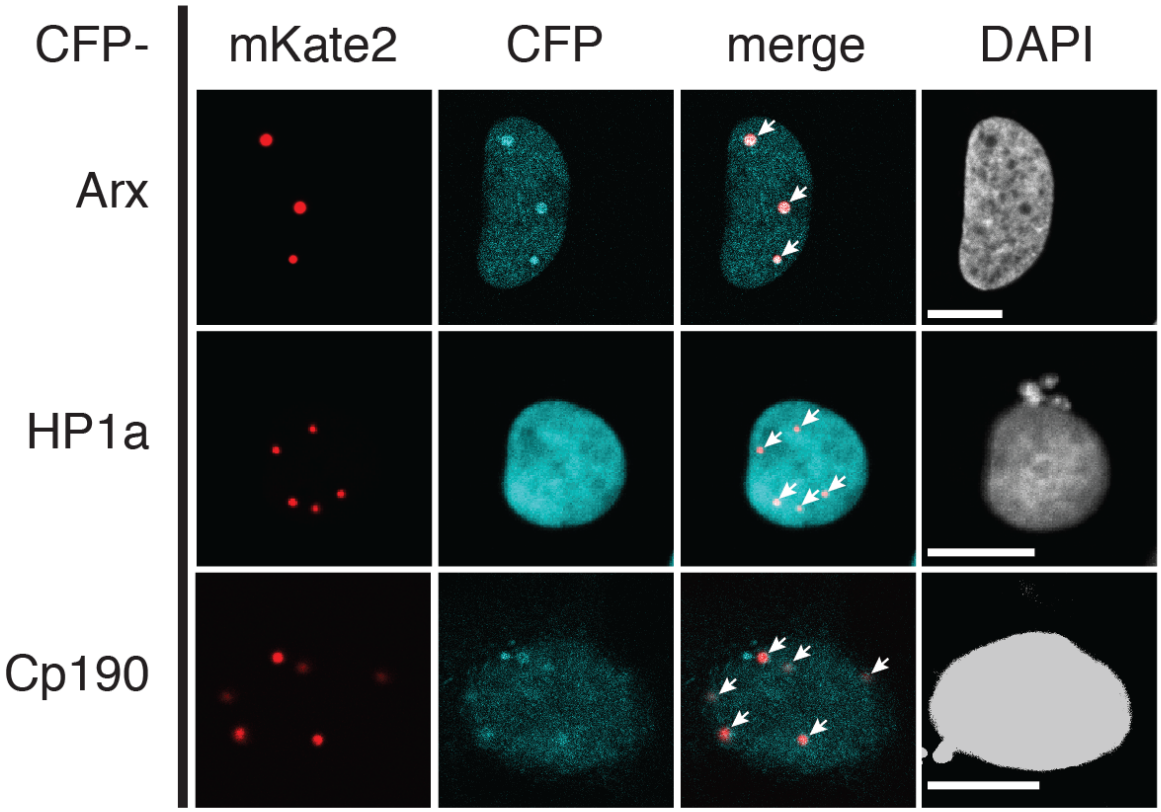

Figure 4. Piwi directly interacts with Asterix, HP1a, and Cp190 in the nucleus.

Piwi directly interacts with the protein of unknown function, Asterix (Arx) (p-value=1.00; std=0.00; n=5) and HP1a (p-value=1.00; std=0.00; n=5). CFP-tagged Cp190 (p-value=0.964 ; std=0.048 ; n=5) was also recruited to lacO foci by lacI-mKate2-Piwi. Representative images are shown. White arrows indicate sites of colocalization. Scale bars represent 10μm.

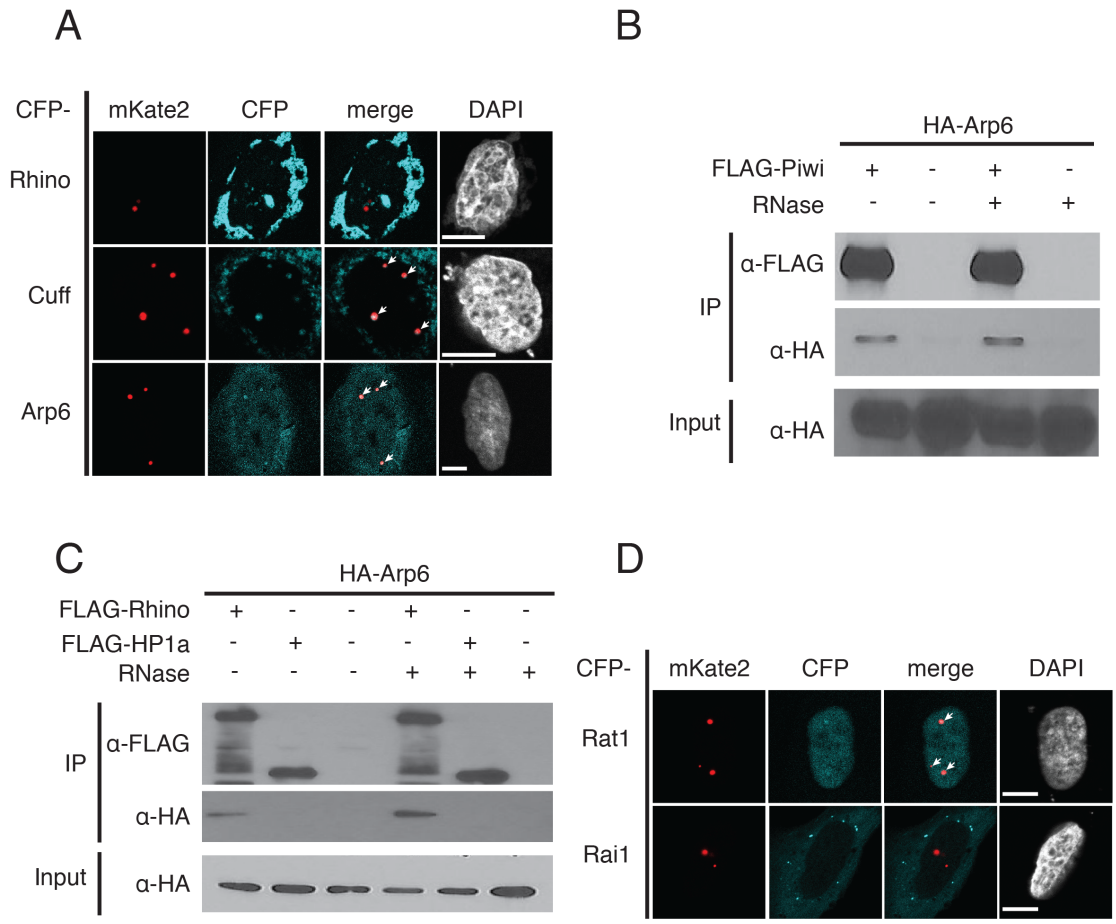

Figure 5. Piwi directly interacts with the Rhino-Deadlock-Cutoff (RDC) complex component Cutoff and the novel interactor Arp6.

A) Piwi directly interacts with the RDC component Cutoff (Cuff) (p-value=1.00; std=0.00; n=5), but not with Rhino (p-value=0.018; std=0.040; n=5). Arp6 (p-value=1.00; std=0.00; n=5) was identified as a novel interactor of Piwi. Representative images are shown. White arrows indicate sites of colocalization. Scale bars represent 10μm. **B)** Piwi and Arp6 interactions are RNA-independent. Co-IP of HA-tagged Arp6 and FLAG-tagged Piwi in S2 cells indicates that the Piwi/Arp6 interaction is unaffected by RNAse A treatment. **C)** Arp6 can be pulled down with Rhino, but not with HP1a. Co-IP of HA-tagged Arp6 and FLAG-tagged Rhino in S2 cells indicates that the Arp6/Rhino interaction is unaffected by RNAse A treatment. **D)** Piwi directly interacts with the exonuclease Rat1 (p-value=0.998; std=0.004; n=5), but not with its binding partner Rai1 (p-value=0.438; std=0.267; n=5).

Representative images are shown. White arrows indicate sites of colocalization. Scale bars represent 10μm.

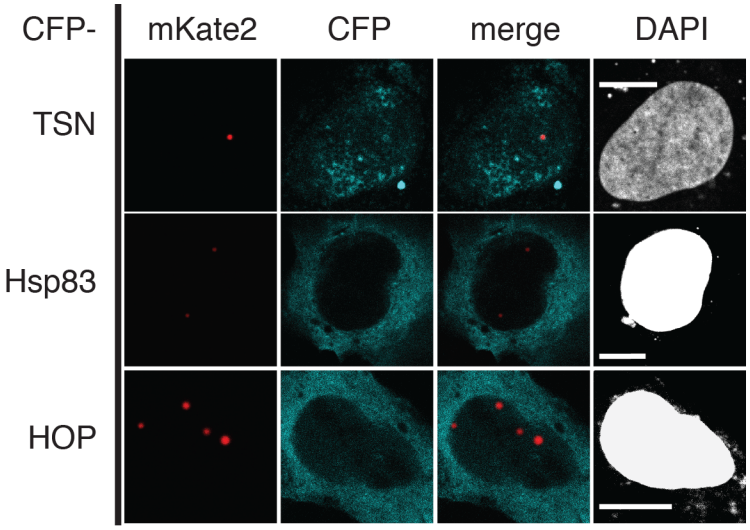

Figure 6. Interactions between Piwi and factors that modify Piwi are indirect.

LacI-mKate2-Piwi did not recruit Tudor-SN (TSN) (p-value=0.044; std=0.062; n=5), Hsp83 (p-value=0.00; std=0.00; n=5), or HOP (p-value=0.006; std=0.013; n=5). Representative images are shown. Scale bars represent 10μm.

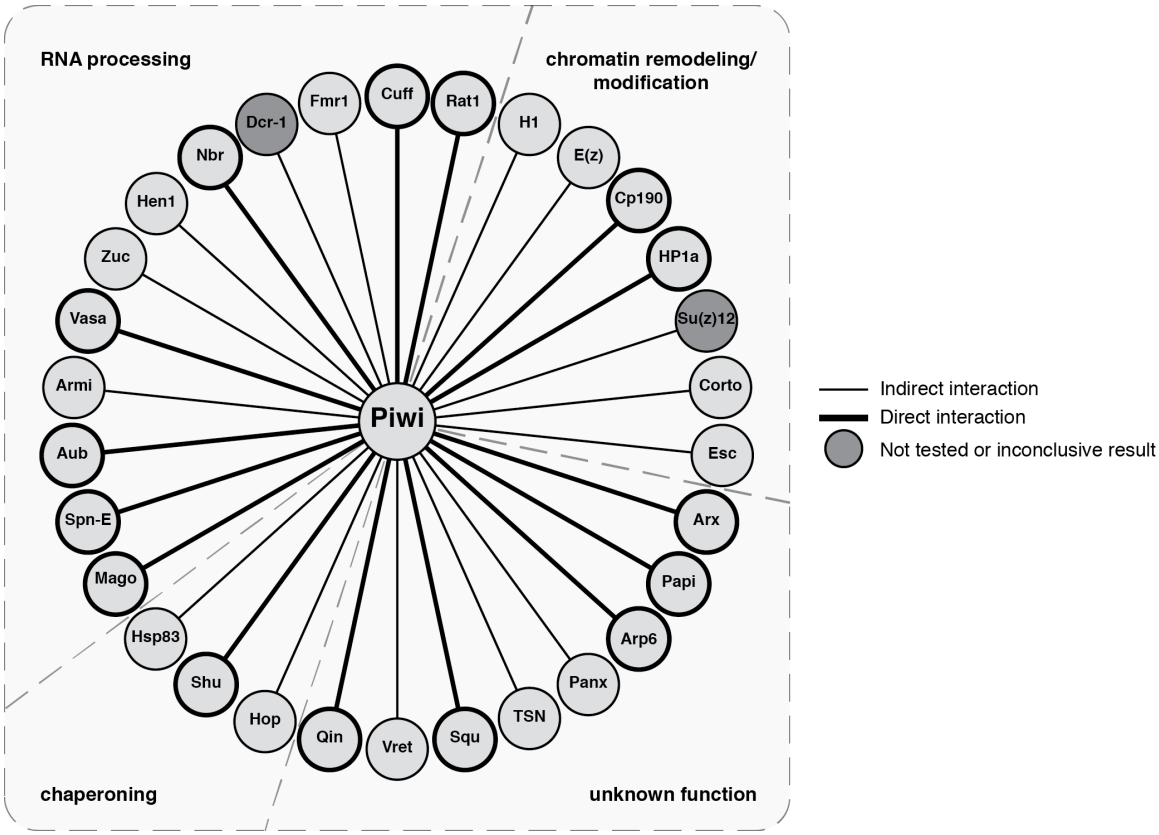

Figure 7. A map of the Piwi interactome with direct and indirect interactions indicated.

An updated map of factors identified to interact with *D. melanogaster* Piwi in the piRNA pathway. Direct protein-protein interactions are indicated (bolded lines). Newly identified factors have been added. Factors Dicer-1 (Dcr-1) and Su(z)12 were previously identified as interacting with Piwi, but we were unable to test whether these interactions were direct or indirect (dark gray circles).