Doctoral Thesis

Tejas Functions in piRNA Biogenesis Via Nuage Assembly in *Drosophila*

Lin Yuxuan

Graduate School of Frontier Biosciences

Osaka University

March 2022

ABSTRACT

In *Drosophila* germline cells, the piRNA amplification, called the ping-pong cycle, occurs at perinuclear structures named '*nuage*'. Two PIWI-clade Argonaute proteins, Aubergine (Aub) and Argonaute3 (Ago3), alternately bind and slice the piRNA precursors and TE transcripts into piRNAs assisted by other components localized to *nuage*, such as Tejas (Tej) and Spindle-E (Spn-E), which are indispensable for piRNA ping-pong amplification, but the molecular basis of how they cooperate in this processing pathway remains elusive.

In the absence of Tej, high-resolution imaging of ovaries containing endogenously fluorescent-tagged SpnE, Vas, Ago3, and Aub revealed an apparent defect in the localization of several piRNA factors. Core helicase Vas lost its granulation, and the function unknown helicase SpnE displayed a unique nuclear accumulation in *tej* mutant ovaries. piRNA precursors processing and the proper PIWI proteins piRNA loading is fractured.

I suggested that Tej regulates the dynamics of Spn-E and Vas as an organizer for nuage and potentially sustains the limited contact between the Vas and Spn-E. I newly identified a functional NLS in Spn-E and an Spn-E recruiting motif in Tej that cooperatively controlled the subcellular localization of Spn-E. Further research revealed that distinct parts of Tej engaged in the proper formation of the nuage granules by recruiting different nuage components, which ensure the processing of the piRNA precursors. Tej also regulations the molecular kinetic of the nuage components *in vitro* by utilizing its disordered region, indicating the contribution of Tej in the dynamic of nuage granules. Taken together, Tej maintains the female *Drosophila* piRNA biogenesis machinery in the germline cells as a multifunctional organizer of nuage, which takes a vital role in the piRNA precursor processing of germline cells.

ACKNOWLEDGEMENTS

Being a Ph.D. student at Osaka University and a member of the Integrated Biology Laboratories led by Dr. Toshie Kai was a memorable period of my life. All the help from my supervisors, families, colleagues, and friends supported me to finish my project successfully.

I want to express my sincere gratitude to my supervisor Professor Dr. Toshie Kai, for allowing me to be a Ph.D. student under her invaluable guidance. I am deeply grateful to her for leading me into the Drosophila germline biology and piRNA field with incredible patience to discuss and design the precessing of my project. I want to especially thank her for giving me a free investigation atmosphere during my Ph.D. studies and shaping me into a qualified researcher. I also appreciate her kindly effort and generous support in the fellowship and financial issues, which allow me to study Intently.

My gratitude extends to all the staff of Kai lab. To associate professor Taichiro Iki and assistant professor Ritsuko Suyama, thank you for your kindly help on my project. All the discussions with you vastly broadened my horizons on scientific research and life. To assistant professor Shinichi Kawaguchi, thank you for providing generous help on bioinformatics data analysis and various experiments. To secretaries and technicians, thanks for your help on affairs and experiments. To Ph.D. students of Kai-lab, Alisha Chakrabarti, thank you for providing the HetA antibody, and to Isshiki Wakana, many thanks for making the Spn-E knock-In fly line. Finally, thanks for all the cooperation and help to all of my colleagues. I will remember all the cherished time spent together in the lab.

I also appreciate all my friends. To Xue Chunxu and Ma Ziyu, thank you for concerning me and supporting me in my difficult time, grateful that our friendship did not weaken over time and distance. To my old friend Huang Chuankun, as a fellow traveler in the Ph.D. journey, thanks for sharing your time with me. I hope fortune smiles on us when we try to push the boundaries of science. Thanks for sharing happiness with me, to all the friends I met in China, Japan, and the internet.

Finally, I am extraordinarily grateful to my mother Tang Xinjuan and father Lin Zuwei, who shaped my body and mind. Without your full support with unrepayable love at any stage of my life, I will never get so far.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

1. INTRODUCTION

1.1 Transposons

Transposons, also named Transposable elements, are selfish elements that exist on all eukaryotic organisms. TE content in each animal correlates strongly with their genome size, and in some species, it constitutes nearly 85% of the genome, involves gene expression regulation, and promotes genetic innovation (Wells and Feschotte 2020). Transposons are classified into two types depending on the transposition mechanisms. Class I transposon, also called retrotransposon, is a category in transcripts from DNA to an RNA intermediator for proliferation. Then the RNA reverse transcript to DNA and insert the other genome locus (Dombroski et al. 1994). The class I transposons are further divided into three subgroups: long terminal repeats (LTRs), long interspersed nuclear elements (LINEs), and short interspersed nuclear elements (SINEs). LTRs, like retroviruses, contain the sequence encode reverse transcriptase, LINEs also encode reverse transcriptase but have no LTR, SINEs contain none of the two factors (Brown et al. 1987; Luan et al. 1993). Class II transposon named DNA transposon has the function of direct transposition into the genome via a cut-and-paste mechanism. Inverted tandem repeats flank the transposon and have been recognized by two transposases, respectively. They will joint and promote DNA double-strand cleavage. The cargo DNA is inserted into the specific motif in the genome (Lambowitz and Zimmerly 2011).

Although the transposons contribute to genome evolution, they highly threaten genome completeness due to their ability to move and insert into the genome (Orgel and Crick 1980; Moon et al. 2018). Their transposase activity mediated transportation directly interrupt the gene by causing double-strand breaks, ectopic recombination, altered gene expression (Hedges and Deininger 2007; Hedges and Belancio 2011).

1.2. Transposons in *Drosophila*

Global genome screening has already revealed that the transposons occupy 23% of the *D. melanogaster* genome. Ninety-six families of transposable elements in *D. melanogaster* had been defined and localized, including 49 LTR families, 27 LINE-like families, and 19 TIR families (Kaminker et al. 2002).

The transposons in *Drosophila* deliver an essential role in maintaining chromosome ends. The *Drosophila* have no telomerase and simple repeats sequences generated by telomerase. Instead, *Drosophila* telomeres contain three non-LTR retrotransposons, TAHRE, HeT-A, and TART. They are transposed into the heterochromatic region in the genome, which is localized at the end of chromosomes, as construct head-to-tail joined telomere arrays. (Blackburn 1992; Pardue and DeBaryshe 2003; Abad et al. 2004; Pardue and Debaryshe 2011).

The transposons in male and female *Drosophila* gonads spread through interbreeding to propagate in a population. The DNA transposons P-elements that spread through wild *Drosophila melanogaster* do not exist in the previously isolated laboratory strains in the early 20th century. Crosses between P-elements carrying P strain males and P-elements deficient lab stocks, M strain females, thus lead to P-M hybrid dysgenesis, causing sterility due to the invading of P-element (Kidwell, Kidwell, and Sved 1977; Rubin, Kidwell, and Bingham 1982; Khurana et al. 2011; Kofler et al. 2015).

The active transposons threaten the genome integrity of germline cells by the products from both transposition and translation. *gypsy,* which belongs to the LTR family, is one of the well-studied examples in *D. melanogaster*. This endogenous retrovirus encodes a Gag-like protein, a protease-polymerase fusion protein, and an envelope. These translation products are generated in follicle cells and been delivered to the oocyte. (Chalvet et al. 1999; Mejlumian et al. 2002). Another member of the *gypsy* family is *ZAM*, a retrotransposon that expresses similar Gag and Env polypeptides in follicle cells and passes to the germline cell via

the vitellogenin secretion pathway (Leblanc et al. 1997; Leblanc et al. 2000; Brasset et al. 2006).

Transposons can invade the genome in novel populations and species, and uncontrolled transposon invasions will potentially lead to a crisis to the host (Kofler et al. 2015). Therefore, the organism needs to suppress the spread of TEs. Piwi-interacting RNAs, an immune system-like mechanism against transposon invasions, have appeared due to the 'evolutional' pressure to protect the genome integrity in the host organism.

1.3. piRNAs

RNA silencing is a defense weapon against invading poly nucleic acid molecules. By producing short non-coding RNA molecules and utilizing them as guiders, the RNA interference (RNAi) pathways achieve the ability of sequence-specific silencing. Two well-studied pathways are small RNAs (siRNAs) and microRNA (miRNA) RNA interference pathways. Nevertheless, the transposon resistibility in the animal germline is preserved by a class of Dicer-independent small RNA, distinct from miRNAs and siRNAs (Vagin et al. 2006). These germline-specific small RNAs are presented to defend the germline genome against transposon mobilization in most animals. They have been found in mouse, rat, and human germline cells (Aravin et al. 2006; Girard et al. 2006; Lau et al. 2006; Grivna et al. 2006).

Distinct from miRNAs and siRNAs, these small RNAs bear 2'-O-methyl-modified 3' termini and have their role in binding and guiding PIWI-clade Argonaute proteins rather than the AGO-clade, which function in the miRNA and siRNA pathways (Brennecke et al. 2007; Cenik and Zamore 2011; Czech and Hannon 2011; Ozata et al. 2019). These sets of small RNAs longer than siRNA, 24-29nt, can be identified as repeat-associated small interfering RNAs (rasiRNAs) also recognized in the *Drosophila* germ cells. This type of sRNAs is firstly identified in testis, with the ability to repress the testis-expressed Stellate genes, following in ovaries

and embryos (Aravin et al. 2001; Aravin et al. 2004; Vagin et al. 2006). Their biogenesis in processing, highly associated with PIWI-family proteins Piwi, Aubergine (Aub), and Argonaute 3 (AGO3), requires neither Dicer-1 nor Dicer-2, thus named as piwi-interacting RNAs, piRNAs (Vagin et al. 2006; Gunawardane et al. 2007; Brennecke et al. 2007).

1.4. piRNA clusters

By mapping the piRNA sequence to the genome of *Drosophila*, researchers reported that the major sources of piRNAs are discrete heterochromatic loci. These loci contain inactivated sense and antisense transposon truncations and can be the sources of piRNAs, identified as piRNA clusters (Brennecke et al. 2007; Brennecke et al. 2008). Thus, the piRNA clusters provide a template library for the piRNA biogenesis mechanism to generate the piRNAs homologous to transposons.

The piRNA clusters form a genetic record of transposon invading as a transposon immune system. The newly invaded transposons will trigger the production of *de novo* piRNA that is homologous to the invaders. When the invading transposons are integrated into the paternally inherited cluster loci, they will be 'trapped' and provide the temples for generating corresponding piRNAs that fight against the transposon themselves (Malone and Hannon 2009; Khurana et al. 2011; Rozhkov et al. 2013; Ozata et al. 2019). In *Drosophila,* piRNA clusters are activated in both germline and ovarian somatic cells and can be sorted into two groups depending on the molecular mechanism of transcription. Most of the somatic cell clusters are uni-strand clusters that are transcribed from one orientation. By contrast, the dual-strand clusters activated in germline cells start their transcription in two directions (Gleason et al. 2018; Ozata et al. 2019).

1.4.1 Dual-strand clusters

Transcription of a dual-strand cluster depends on RNA polymerase II, same as canonical transcription. However, their transcripts are remarkably different (Chen et al. 2016). Most of the dual-strand cluster transcripts are uncapped in 5'-end and lack the polyadenylated tail (Klattenhoff et al. 2009; Malone and Hannon 2009; Zhang et al. 2014; Mohn et al. 2014). They also lack the hallmarks of canonical transcription, such as the active promoter mark histone H3 lysine4 trimethylation (H3K4me3) and the unique sequence for RNA splicing or transcriptional termination (Ozata et al. 2019; Li et al. 2009). As the promoters are not located in piRNA cluster loci, one possibility of transcription of dual-strand clusters relies on the read-through from the flanking protein-coding genes. However, depletion of the promoter region on the *Pld* gene upstream of *42AB* piRNA precursors did not abolish the transcription of downstream precursors (Chen et al. 2016). This evidence shows that the piRNA dual-strand clusters utilize a distinct initiation mechanism from the canonical Pol II transcription.

Instead of using canonical promoters, dual-strand clusters initiate the transcription relying on germline-specific complexes H3K9me3, not H3K4me3 utilized in the canonical Pol II-dependent transcription. Missing the critical methyltransferase dSETDB1 (egg) will lead to a block of piRNA cluster transcription and collapse the piRNA biogenesis pathway (Rangan et al. 2011). A unique promoter-independent protein complex regulates the dual-strand cluster transcription in the germline cells, called the RDC protein complex. The members of the RDC complex are H3K9me3-binding protein Rhino (Rhi), a variant of heterochromatin protein 1a (HP1a), Cutoff (Cuff), and Deadlock (Del). Rhino recognizes H3K9me3 through the chromodomain, then the gathering of Deadlock, Cutoff, and Rhino licenses the transcription(Klattenhoff et al. 2009; Mohn et al. 2014; Zhang et al. 2014). Continuously, Moonshiner, a paralog of germline-specific transcription initiation factor II A subunit 1 (GTF2A1), is recruited to RDC. Then TATA box-binding protein-related factor TRF2 is then

sequentially recruited and initiates the transcription on the piRNA cluster (Andersen et al. 2017). Unlike the canonical transcription followed by RNA slicing processing, in piRNA cluster transcripts, Rhino together with Cuff and the DEAD-box helicase UAP56 suppresses the RNA slicing by occupying the splice sites composed by consensus sequences on piRNA cluster transcripts (Zhang et al. 2014). The importance of keeping unsliced intron-containing transcripts remains unknown in piRNA biogenesis processing.

Additionally, the Cuff proteins perform as guardians of piRNA precursors. They can suppress the premature termination of RNA Pol II and restrain the degradation of RNA caused by the nuclear exonuclease Rat1. Cuff proteins preserve the Poly (A) tail of nascent piRNA precursors by stopping the recruitment of the polyadenylation specificity factor (CPSF) complex, which can slice the Poly (A) tail (Chen et al. 2016). Transcription/export (TREX) complex, a well-studied protein complex that is critical for pre-mRNA processing and mRNA nuclear export in yeasts and mammals (Masuda et al. 2005; Katahira 2012). Tho5, the TREX in the THO subunit complex recruits Cuff and accumulates in the nascent piRNA precursors. It was revealed piRNA cluster transcripts were transported to the downstream piRNA-producing machinery with UAP56 (Strässer et al. 2002; Zhang et al. 2012; Hur et al. 2016) (Figure 1.1 A).

1.4.2 Uni-strand clusters

In somatic follicle cells surrounding the germline cells, the uni-strand piRNA clusters substituted the function for dual-strand piRNA clusters (Brennecke et al. 2007; Malone et al. 2009). Unlike the dual-strand clusters, the uni-strand clusters are transcribed by an RDC complex independent manner. The largest uni-strand cluster is the *flamenco locus,* which is localized in the X chromosome, produces a long noncoding RNA through the canonical transcriptional machinery by the transcription factor Cubitus interruptus (Ci). Transcripts

from the *flamenco* are alternatively spliced into multiple *flamenco* precursors that share the first exon at their 5' end (Goriaux et al. 2014). After the transcripts are produced from the cluster, the Yb protein could recognize the unique cis-acting RNA elements in the precursors. The binding of Yb proteins to the precursor transcript determines the further slicing of precursors (Ishizu et al. 2015). Transcripts derived from *Flamenco* cluster contain the somatic cell specific antisense transposons, like ZAM family, *Idefix*, and *gypsy* (Desset et al. 2003; Goriaux et al. 2014; Prud'homme et al. 1995). Not all piRNA in somatic cells of *Drosophila* ovaries are derived from the uni-strand cluster. It also reported that the 3'UTRs of traffic jam protein-encoding mRNAs are processed into piRNAs (Robine et al. 2009) (Figure 1.1 B).

1.5 PIWI proteins

Once long transcripts are derived from the piRNA clusters, they are delivered to the molecular processing machinery localized at the cytoplasm. As the piRNA precursors are driven from distinct clusters in germline cells and gonadal somatic cells, the machineries of piRNA biogenesis are also different between the two types of cells. However, they shared the same core components, PIWI-family proteins, an Argonaute family subclade (Cox et al. 1998; Carmell et al. 2002). By forming the RNA-induced silencing complex (RISC), the Argonaute protein family induces RNA interference (RNAi). Several groups of small non-coding RNAs, including miRNAs, siRNAs directly bind to Argonaute proteins and guide them to their sequence complementarity targets induce the mRNA cleavage or translation inhibition. In both male and female gonads, piRNAs are loaded to the PIWI family, a germline specific Argonaute protein family. The members of PIWI family are PIWI, Aubergine (Aub), and Argonaute3 (Ago3) that share three highly conserved domains with the AGO proteins, PAZ (Piwi/Argonaute/Zwille), MID and PIWI domains (Song et al. 2004; Parker, Roe, and Barford 2005; Elkayam et al. 2012; Schirle and MacRae 2012; Wei et al. 2012). Among different AGO

and PIWI proteins, the diversity of PAZ domains has a unique binding pocket by crystal structure and shows a distinct binding preference to the 3’ end modification of guide RNAs. The guide RNAs are terminal modified RNAs that bring the AGO and PIWI proteins to the slicing targets. To the members of the PIWI family, the matured piRNA functions as guide RNAs (Ma, Ye, and Patel 2004; Tian et al. 2011). The MID and PIWI domain constitute a hydrophilic interface and provide a binding pocket for the 5’ end phosphate of guide RNA (Parker, Roe, and Barford 2005; Boland et al. 2011). After the guide RNA is anchored to the PIWI proteins through the PAZ and MID domain, the PIWI protein is directed to the target RNA that is complementary to the guide RNA. PIWI domain proteins hydrolyses the phosphodiester bond linking of nucleotides, causing a cleavage in between the nucleotides of target RNA (Parker, Roe, and Barford 2004; Schwarz, Tomari, and Zamore 2004; Yuan et al. 2005). The disordered regions in N-terminal PIWI proteins have also been reported to provide an essential site for post-translational modification in piRNA biogenesis and silencing (Huang et al. 2021). These cleavage activities of PIWI proteins provide the molecular basis for piRNA biogenesis, supporting the other piRNA processing mechanism such as the Zucchini(Zuc)-dependent piRNA biogenesis pathway in somatic cells or the piRNA amplification pathway known as the ping-pong cycle.

1.6 piRNA precursor export in the ovary

1.6.1 piRNA precursor exportation in germline cells

The transport of precursors in the germline cells is largely different from the somatic cells. Once the precursors are transcribed from heterochromatic clusters, UAP56, the DEAD-box containing nuclear helicase, is recruited to the nascent transcripts by the Rhino protein. UAP56 then carries the precursors to the nuclear periphery and interacts with the nuclear pore components, triggering RNA release. The precursors are delivered to nuage located on

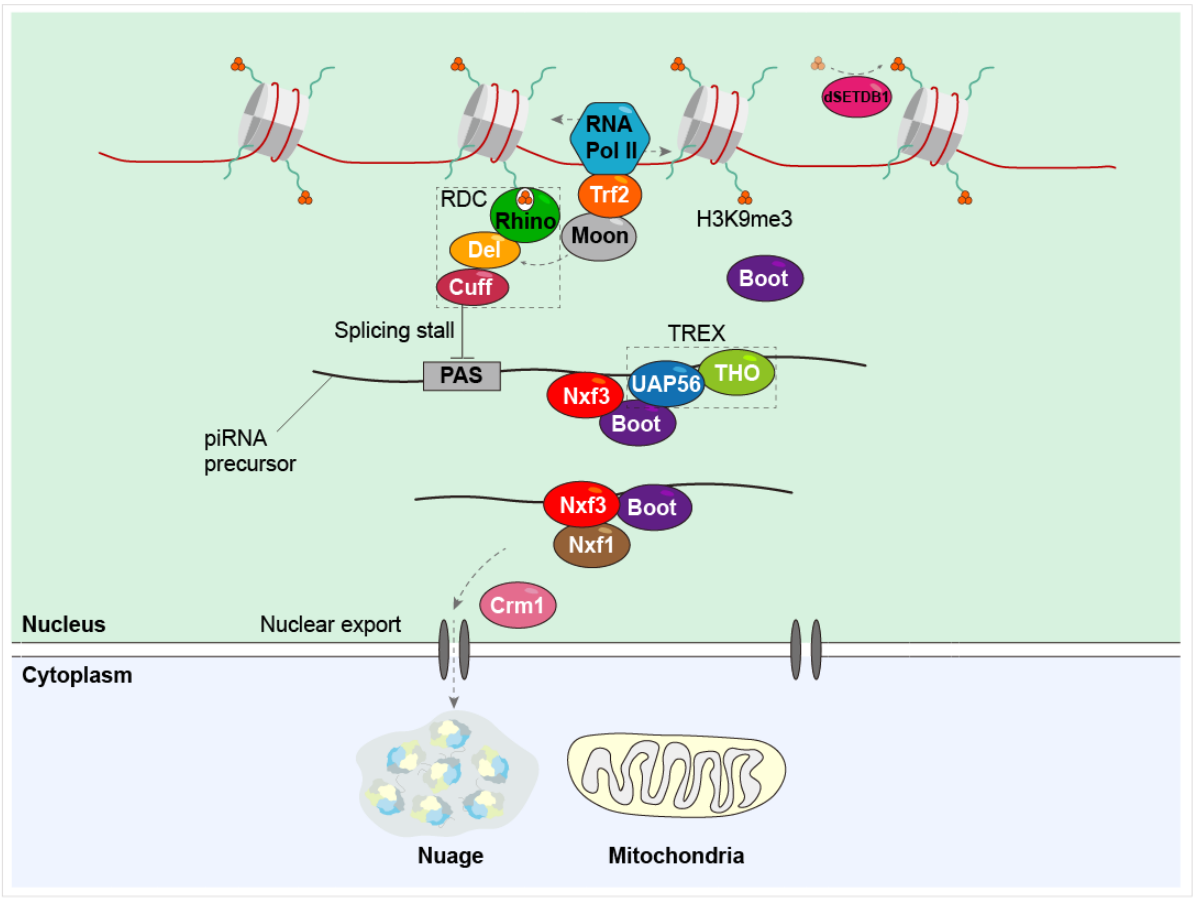
the perinuclear region and are captured by Vas, a DEAD-box containing helicase (Zhang et al. 2012). A germline cell-specific piRNA precursor transport to the cytoplasm has been recently identified in coordination with the UAP56. Nxf3, a germline-specific paralog of Nxf1 (Czech et al. 2013), and its cofactor Bootlegger (Boot) have been visualized to be localized at Rhi-dependent dual-strand cluster transcripts foci inside the germline nucleus. RNA immunoprecipitation sequencing (RIP) shows that the Nxf3 proteins bind to piRNA precursors derived from Rhi-dependently transcribed *42AB* and *38C1*, but not the Rhi-independent *20A* cluster (Kneuss et al. 2019). Bootlegger leads the localization of Nxf3 and UAP56 to the RDC complex that recognizes dual strand piRNA clusters. Loss of Nxf3 vastly reduced the cytoplasmic perinuclear localization of piRNA precursors derived from *42AB* and *80EF clusters*, while the nuclear foci are unaffected. These results indicate that the export machinery of Rhino-dependent piRNA precursors in germline cells is impaired in the absence of Nxf3 (Kneuss et al. 2019; ElMaghraby et al. 2019; Mendel and Pillai 2019). Distinct from the Nxf1-Nxt1 transport in the somatic cells, the transport of precursors derived from the dual-strand cluster in germline cells requires the interaction of Nxf3 and Nxt1 (Mohn et al. 2014; Goriaux et al. 2014; ElMaghraby et al. 2019). The piRNA precursors carried by Nxf3 are recognized by Crm1, the cellular exportin protein, and were translocated to the nuclear pore complexes (NPCs) for the pronuclear deposition(Yang et al. 2001) (Figure 1.1 B).

1.6.2 piRNA precursor export in ovarian somatic cells

As canonical RNA polymerase II transcripts, the somatic piRNA precursors, like flamenco, are capped on the 5’ and polyadenylated on the 3’ end. Same with the mature mRNA, their 5’ cap will be recognized by the cap-binding complex (CBC). Together with the Transcript/export complex (TREX), the nuclear export factor 1 (Nxf1) coordinating with the partner NTF2-related export protein 1 (Nxt1) transport the somatic piRNA precursors into the cytoplasm

(Viphakone et al. 2012; Katahira 2012; Handler et al. 2013; Czech et al. 2013; Dennis et al. 2016). Once the precursors have been transported to the cytoplasm, as an example, the *flam* cluster-derived transcripts are accumulated into the perinuclear structure named flam body. The flam body, neighboring but segregated from the Yb-bodies, may act as a storage site for the precursors, but the details remain unknown (Murota et al. 2014). Then the precursors are delivered to the unique piRNA biogenesis pathway, the Zucchini-dependent piRNA biogenesis pathway, in the ovarian somatic cells (Figure 1.1 A).

A


dSETDB1
RNA Pol II
Trf2
RDC
Rhino
Moon
Del
Cuff
H3K9me3
Boot
Splicing stall
TREX
PAS
Nxf3
UAP56
THO
Boot
piRNA precursor
Nxf3
Boot
Nxf1
Crm1
Nucleus
Nuclear export
Cytoplasm
Nuage
Mitochondria

B

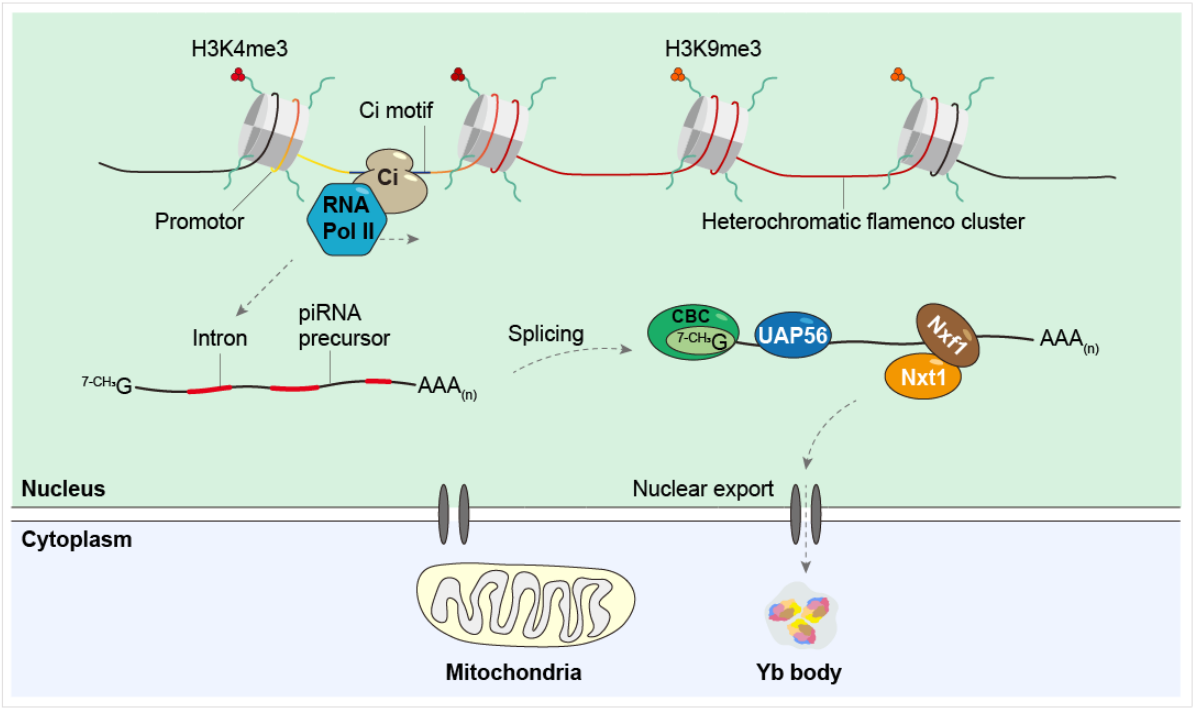

H3K4me3
Ci motif
H3K9me3
Ci
RNA Pol II
Promotor
Heterochromatic flamenco cluster
Intron
piRNA precursor
7-CH3G
AAA(n)
Splicing
CBC
7-CH3G
UAP56
Nxf1
Nxt1
AAA(n)
Nucleus
Nuclear export
Cytoplasm
Mitochondria
Yb body

Figure 1.1. Transcription and Transportation of piRNA precursors derived from Clusters.

(A) Transcription and Transportation of Dual-strand clusters. Dual-strand clusters transcript rely on canonical RNA polymerase II, as an H3K9me3 depending way. H3K4me3 sustained by the critical methyltransferase dSETDB1. The members of the RDC complex are H3K9me3-binding protein Rhino (Rhi), a variant of heterochromatin protein 1a (HP1a), Cutoff (Cuff), and Deadlock (Del). Rhino recognizes H3K9me3 then recruits Deadlock, Cutoff to license the transcription. After that, Moon will be recruited sequentially to RDC, cooperate with TRF2 to initiate the transcription on piRNA cluster products the uncapped in 5'-end and lack the polyadenylated tail piRNA precursors. Meanwhile, together with the Cuff and UAP56, the Rhino suppresses the RNA slicing. Responding to the recruitment of Cuff, the THO accumulated in the nascent piRNA precursors together with UAP56 and assembles the Transcription/export (TREX) complex to help the piRNA cluster transcripts transport to the downstream piRNA-producing machinery. Coordinates with the UAP56, Nxt1, and cofactor Boot, Nxf3 specifically binds to piRNA precursors and be recognized by the cellular exportin Crm1 and finish nuclear pore complexes (NPCs) translocation.
(B) Transcription and Transportation of Uni-strand clusters in somatic follicular cells. The uni-strand clusters transcription initiates through an RDC complex independent con manner. A long noncoding RNA through the RNA polymerase II transcript mechanism activated by the transcription factor Cubitus interruptus (Ci) recognizing cis-acting RNA elements. The Uni-strand *flamenco* precursor will be capped and polyadenylated after alternatively spliced into multiple precursors. Further exportation relies on the cooperation of the TREX complex and the Nxf1-Nxt1 system. Adopted and modified from (Ozata et al. 2019) (Mendel and Pillai 2019)

1.7 piRNA biogenesis

As discussed above, the piRNA precursors, including the piRNA cluster transcripts, mRNA 3' UTR of genes, and the transposon transcripts, will be delivered to the further processing at the cytoplasmic machinery. Several protein complexes stepwise cleavage and modify the 5' and 3' end of the piRNA precursors and intermediate pre- piRNAs that shape them into 23-29 nucleotides mature piRNAs.

1.7.1 The 5' end formation of the piRNAs

The monophosphorylated 5' end is required to bind with PIWI proteins (Parker, Roe, and Barford 2005; Boland et al. 2011), which licensed only long single-stranded nascent piRNA precursors are loaded to the Piwi (Preall et al. 2012; Olivieri et al. 2012). At present, these 5' end-nascent piRNA precursors are made in two manners in the *Drosophila* piRNA biogenesis pathway (Figure 1.2 A).

One of the 5' end formations is a Zucchini(Zuc)-dependent. In both *Drosophila* somatic and germline cells, endonuclease Zuc plays an essential function in piRNA 5' end formation. The piRNAs disappeared in the absence of the Zuc n the ovaries (Ipsaro et al. 2012; Nishimasu et al. 2012). Zuc is located on the outer surface of the mitochondria through the mitochondrial targeting signal (MTS) (Pane, Wehr, and Schüpbach 2007; Olivieri et al. 2010; Saito et al. 2010). As a mitochondrial phospholipase D (PLD) superfamily protein plays crucial role in piRNA maturation, Zuc can further process the precursors (Munafò et al. 2019; Yamashiro et al. 2020). *in vitro* experiment suggested Zuc works as a single-strand specific endoribonuclease and forms a monophosphate 5' end in single-stranded products recognized by a narrow catalytic groove on their homodimerization interface. The cleavage products have no significant bias in the first nucleotide of the 5' end, which is inconsistent with the signature of the Piwi-bound piRNA, suggested that 1-U preference of 5' end of mature piRNAs was

determined by the preferentially binding of Piwi proteins during the loading step (Pane, Wehr, and Schüpbach 2007; Ipsaro et al. 2012; Nishimasu et al. 2012).

The other is distinct from the Zucchini-dependent in 5’ end formation. A feed-forward amplification cycle has been identified explicitly in *Drosophila* germline cells and called ‘ping-pong cycle’, to bear the 5’ end formation. The ping-pong pathway possesses high conservation among the different organisms, and a similar mechanism has already been unearthed from the Hydra, fruit fly, mouse, and human being (Aravin et al. 2007; Houwing et al. 2007; Grimson et al. 2008; Lim et al. 2014).

In this pathway, the maternally inherited piRNAs loaded PIWI proteins are directed to cleave the precursors in the complement site as an initial step. The PIWI proteins mono-phosphorylate the 5’ end of long RNA fragments and produce pre-pre-piRNA. Another empty PIWI protein further captures the pre-pre-piRNAs by its nascent 5’ end. New PIWI proteins continually bind and clip the pre-pre-piRNAs with endonuclease cleavage activity, thus digesting the long pre-pre-piRNA into pieces. After trimming and methylation, these matured so-called responder piRNAs become new initiator piRNAs that guide the newly formed piRNA-induced silencing complex (piRISC) to target the complementary precursors and shape the 5’ end of them. In the ‘ping-pong cycle’, the continuous cleavage of the transposon and piRNA cluster transcripts is enabled to amplify the population of 5’ phosphorylated pre-piRNA fragments effectively.

1.7.2 The 3’ end formation of the piRNAs

The mechanisms for establishing the 3’ end of the mature piRNA show more variety. In addition to the generation of 5’ end monophosphorylating of the precursors by Zuc, they are cleaved by Zuc and polished into the pre-piRNA with their 3’ end and 5’ end of trailing simultaneously. The evidence was revealed from the head to tail strain mapping of the piRNAs

to the precursor, indicating that the Zuc cleave and produce the similar e length of mature piRNAs, instead, with little or no extra 3’ nucleotides (Mohn, Handler, and Brennecke 2015; Czech and Hannon 2016; Hayashi et al. 2016). This ping-pong cycle independent piRNA biogenesis pathway has been defined as the ‘Phasing’ pathway, which produces trailing pre-piRNAs and then been further polished to trailing piRNA and mature piRNAs with proper size (Mohn, Handler, and Brennecke 2015; Han et al. 2015; Ozata et al. 2019).

Part of the pre-piRNAs remain some extra 3’ nucleotides even after the 5’ end formation and cleavage on 3’ end by Zuc, and these intermediate products need to be further polished on the 3’ end for obtaining a genius length and appropriately modified tail to improve its stability as mature piRNAs (Kamminga et al. 2010; Lim et al. 2015). Following the stepwise binding and endonuclease slicing of Piwi proteins, 3’ end of intermediate pre-piRNA fragments in Piwi-RISCs is polished by an exonucleolytic enzyme Nibbler (Nib) (Han et al. 2011; Liu et al. 2011; Hayashi et al. 2016), and then further 2’-O-methylated by 2’-O-methyltransferase Hen1 (Horwich et al. 2007; Saito et al. 2007; Kirino and Mourelatos 2007). In many animals, the pre-piRNAs require the exonuclease for this final trimming, like PNLDC1 in mice (Ding et al. 2017; Nishimura et al. 2018) and the PARN-1 in *C. elegans* (Tang et al. 2016). Distinct from mice and nematode, the piRNA is trimmed by the Nibbler (Nib)and also utilized in the miRNA biogenesis pathway (Liu et al. 2011; Han et al. 2011; Feltzin et al. 2015; Hayashi et al. 2016). As the mouse piRNA methylation depends on the HEN methyltransferase (HENMT1) (Kirino and Mourelatos 2007), the 2'-O-methylation on the 3' end of piRNA *in Drosophila* was revealed to be happened by the DmHen1, *Drosophila* homolog of HEN1 (Horwich et al. 2007).

In addition, it has been implied that the 3’ end is modified by another way in the ping-pong pathway. When both Zuc and Nib are eliminated in the ovaries, 3’ ends of piRNAs can

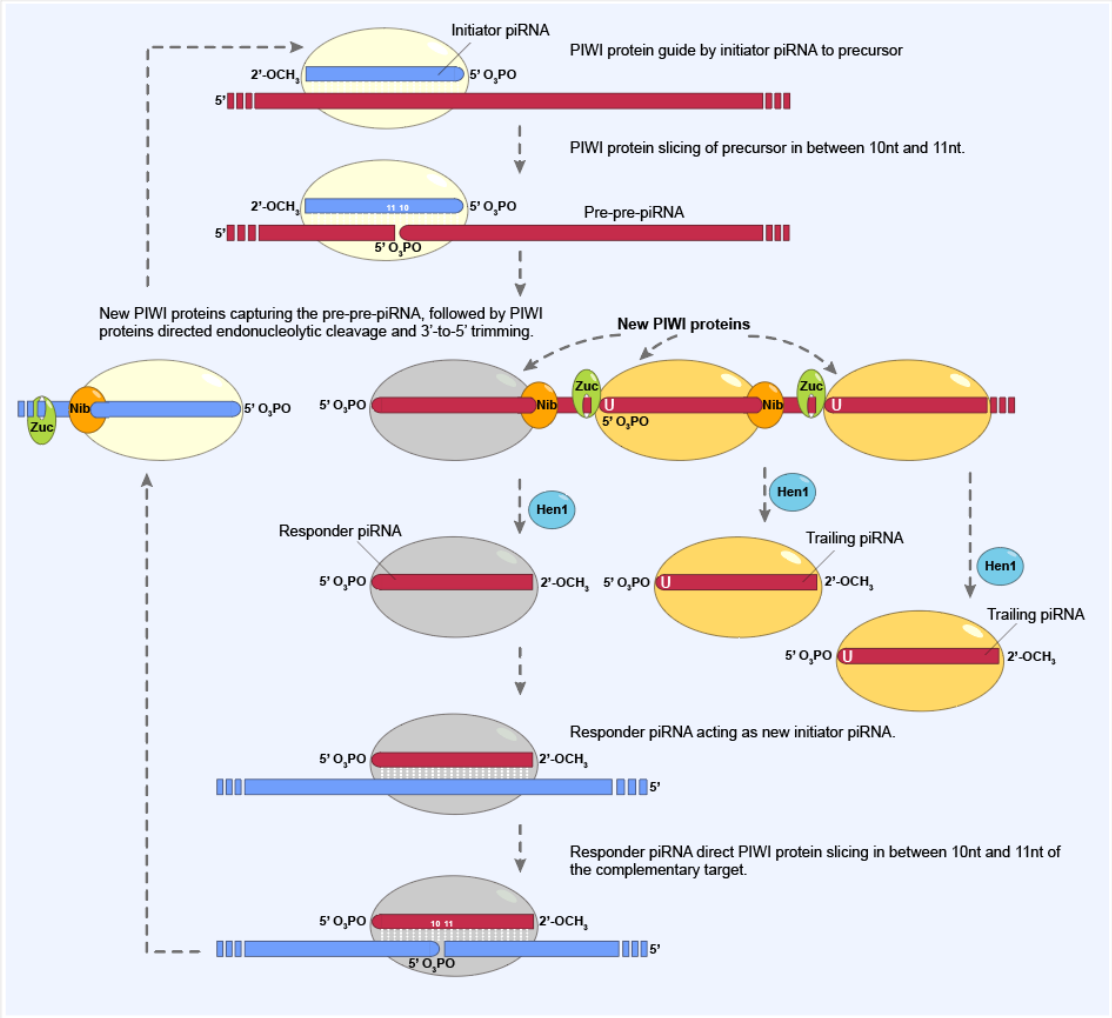
bind and be cleaved in piRNA-guided piRISC to become pre-piRNA, suggesting the existence of a presence (Hayashi et al. 2016). The Aub- and Ago3-dependent piRNA amplification cycle is only present at the germline cells, whereas the Piwi-Zuc dependent pathway functions at only somatic cells. Since the binding of initiator piRNA to PIWI proteins triggered the ping-pong cycle, it is speculated that these initiator piRNAs are produced by the Zuc dependent pathway and let the ping-pong pathway launch in the germline. Thus, piRNA generation in *Drosophila* is divided into the primary and secondary pathways, taking place respectively in the somatic cells and germline cells with distinct mechanisms (Brennecke et al. 2007; Li et al. 2009).

However, the current research indicates that the primary and secondary piRNA biogenesis pathways are more intrinsically intercrossed systems. By utilizing BmN4 cells developed from silkworm ovary, which has similar ping-pong pathways organized by the two cytosolic PIWI proteins ortholog, Siwi and BmAgo3, unexpectedly, defined 'primary' pathway signature was observed by the BmAgo3 slicing (Homolka et al. 2015). In addition, the ovaries of Aub/Ago3 double mutant also eliminated the piRNA from the Piwi protein, which is thought to be produced only from the 'primary' pathway. Loss of Ago3 decreased Piwi-bound piRNAs more than in Aub mutants. As shown, most Piwi-bound piRNAs could be mapped to the sequence following the cleavage products of Ago3 and Aub. Given these facts, the cleavage of piRNAs by Ago3 or Aub triggers Piwi-bound piRNA production by the Piwi-Zuc dependent 'phasing' manner described above, the primary function of Ago3 is to generate Piwi-bound piRNAs during the ping-pong cleavage looping (Han et al. 2015; Mohn, Handler, and Brennecke 2015; Wang et al. 2015).

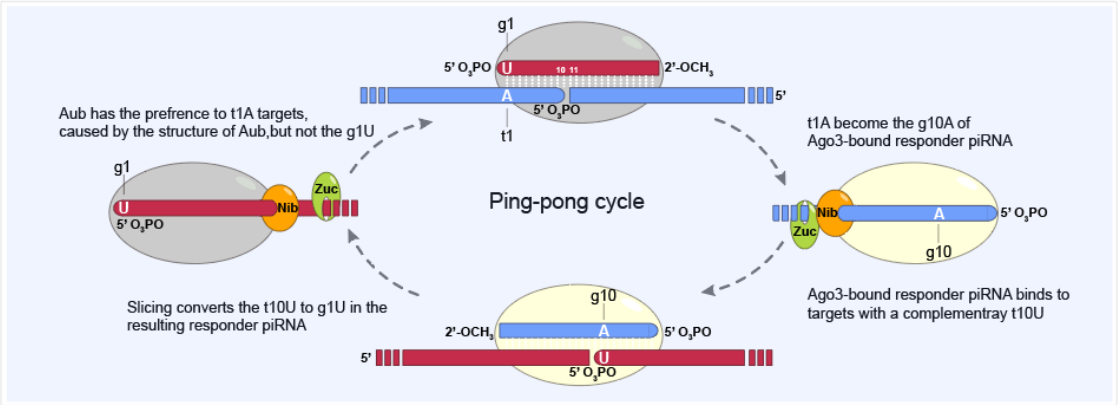
Thus, the current model proposed the piRNA biogenesis using the Zucchini-dependent and Ping-pong/Slicer-dependent piRNA biogenesis pathway, produced the phased Trailing

piRNAs and Initiator/Responder piRNAs, respectively (Huang, Fejes Tóth, and Aravin 2017; Ozata et al. 2019) (Figure 1.2 B).

A


Initiator piRNA
PIWI protein guide by initiator piRNA to precursor
PIWI protein slicing of precursor in between 10nt and 11nt.
Pre-pre-piRNA
New PIWI proteins capturing the pre-pre-piRNA, followed by PIWI proteins directed endonucleolytic cleavage and 3'-to-5' trimming.
New PIWI proteins
Responder piRNA
Trailing piRNA
Trailing piRNA
Responder piRNA acting as new initiator piRNA.
Responder piRNA direct PIWI protein slicing in between 10nt and 11nt of the complementary target.

B


Aub has the prefrence to t1A targets, caused by the structure of Aub,but not the g1U
t1A become the g10A of Ago3-bound responder piRNA
Ping-pong cycle
Slicing converts the t10U to g1U in the resulting responder piRNA
Ago3-bound responder piRNA binds to targets with a complementray t10U

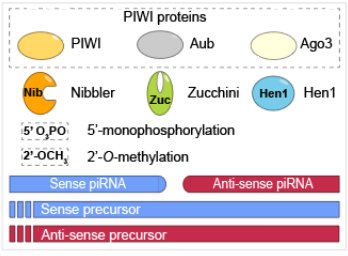

PIWI proteins
PIWI
Aub
Ago3
Nibbler
Zucchini
Hen1
5'-monophosphorylation
2'-O-methylation
Sense piRNA
Anti-sense piRNA
Sense precursor
Anti-sense precursor

Figure 1.2. Terminal and Nucleotide Preference Formation of piRNA.

(A) The terminal processing of the piRNAs. The monophosphorylated 5’ end could be formed through two pathways, the endonuclease slicing of the piRNA guided Aub and Ago3, or the cleavage by endonuclease Zuc. The 3’ end of piRNA can be shaped by the suitable endonuclease Zuc cleavage while generating the trailing 5’ end of piRNA or by 3’-to-5’ trimming of longer precursors by the exonuclease Nibbler. Then become mature piRNAs fit for PIWI-proteins loading with polished by 2’-O-methyltransferase Hen1.
(B) The preference of PIWI proteins for the t1A target is one of the sources of the g1U bias of piRNAs. The structure of Aub precludes pairing between the first nucleotide of the guide (g1) and the opposing nucleotide of the target (t1) of the Ago3-bound responder piRNA. However, t1 Adenine bias is caused by the intrinsic structure of Aub protein, but not g1U complement. The Aub slicing converts t1A to g10A of the responder piRNA further loaded into Ago3. Targets of Ago3 often bear a t10U because of complementary pairing with g10A. Slicing by Ago3 converts t10U to g1U in the novel generated responder piRNA. In flies, Heterotypic Aub–Ago3 ping- pong ensures the antisense bias of Aub- bound ping- pong piRNAs and Piwi- bound trailing piRNAs. Adopted and modified from (Ozata et al. 2019)

1.8 piRNA biogenesis machinery in *Drosophila* germline

In both somatic and germline cells, the further processing of piRNA precursors requires distinctive perinuclear structures, the somatic specific Yb-bodies (Saito et al. 2010; Olivieri et al. 2010) in cells and nuage structure in germline cells (Lim and Kai 2007; Pek, Patil, and Kai 2012). These non-membranous organelles aggregated RNAs and proteins, providing a robust molecular basis for piRNA cleavages.

1.8.1 Nuage, a site for ping-pong cycle in germ cells

The highly abundant piRNAs expressed in the *Drosophila* germline cells rely on a feed-forward amplification cycle named 'ping-pong cycle'. This amplification cycle requires two members of the PIWI family, the Aub and Ago3, supported by other assigned components to process the enormous amount of piRNAs for silencing the transposons during germline development.

In the *Drosophila* germline cells, once the precursors span the nuclear envelope and have been delivered to the nuage, maternally inherited piRNAs, known as initiator piRNAs, guided Ago3 or Aub to recognize them. At this initial stage, while PIWI proteins incise precursors between 10nt and 11nt length of the loading guider piRNA, the recognize-cleavage loop in this cycle relies on the complementary matching of sense and antisense of piRNA in the first ten nucleotides from the 5' end of the slicing products. The Aub bound piRNAs are majorly antisense to the transposons, while the Ago3 preference for binding to sensing piRNA. Sequencing analysis revealed that the Aub-bound piRNAs contain a strong 1 Uridine bias, while the Ago3 bound piRNAs have an alternative 10 Adenosine bias (Brennecke et al. 2007; Gunawardane et al. 2007). The 10A bias of Ago3 bound piRNAs depends on the nucleotide preference of Aub at position 1 in the target strand, which is potentially determined by the intrinsic structure of the Aub protein. Aub preferes to bind the sequences that start with

Uridine and further cleaves its RNA substrates on the 10th nucleotide, creating novel 5' ends on the complementary matched target RNA fragments. The following arrived Ago3 proteins bind to the 5' end of these novels-shaped pre-pre piRNAs, thus gaining a bias on the 10^{th} nucleotides of its binding piRNAs due to base-pairing (Wang et al. 2014) (Figure 1.2 B). These 10 nt complementary overlap with the antisense 1U and sense 10 A bias of the piRNAs has been recognized as the feature of piRNAs derived from ping-pong cycle and known as 'ping-pong signature' (Brennecke et al. 2007; Gunawardane et al. 2007). Recently, another model is proposed that the phasing processing has also been unearthed and merged to this ping-pong cycle since there is the majority of Piwi-bound piRNAs with 1 U signature, but not the ping-pong signature, in germline cells (Han et al. 2015; Mohn, Handler, and Brennecke 2015; Wang et al. 2015).

In each egg chamber, 16-cell cysts generated by four rounds of mitosis contain 15 nurse cells that support the only oocyte (Pepling, de Cuevas, and Spradling 1999). The nurse cells produce abundant piRNAs to prevent the transcripts derived transposon from conveying to the oocyte. This highly effective piRNA processing machinery localizes at the perinuclear region of nurse cells, named 'nuage'. The nuage is a high electron-dense cytoplasmic inclusion, where most of the known piRNAs biogenesis pathway components localize (Brennecke et al. 2007; Pek, Patil, and Kai 2012)(Figures 1.3).

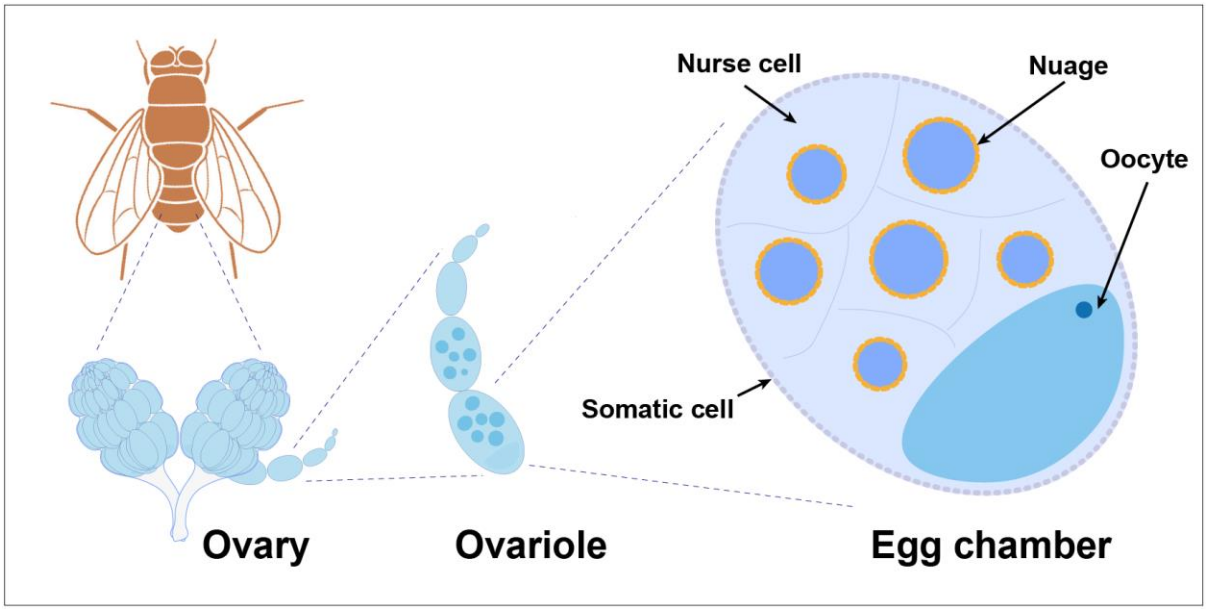


Figure 1.3. Schematic features for female germ cells.

Each *Drosophila* female has one pair of ovaries, containing 16 to 20 of ovarioles covered by muscle sheath. In each ovariole, the germarium, the most anterior structure where the germline stem cells are located, sequentially genarated the egg chambers. Arranged developing egg chambers finally become a mature egg at the most posterior end. Somatic cells surround the oocyte and fifteen nurse cells in each unmatured egg chamber. The Nuage structures are localized in the perinuclear region of the nurse cells and somatic cells.

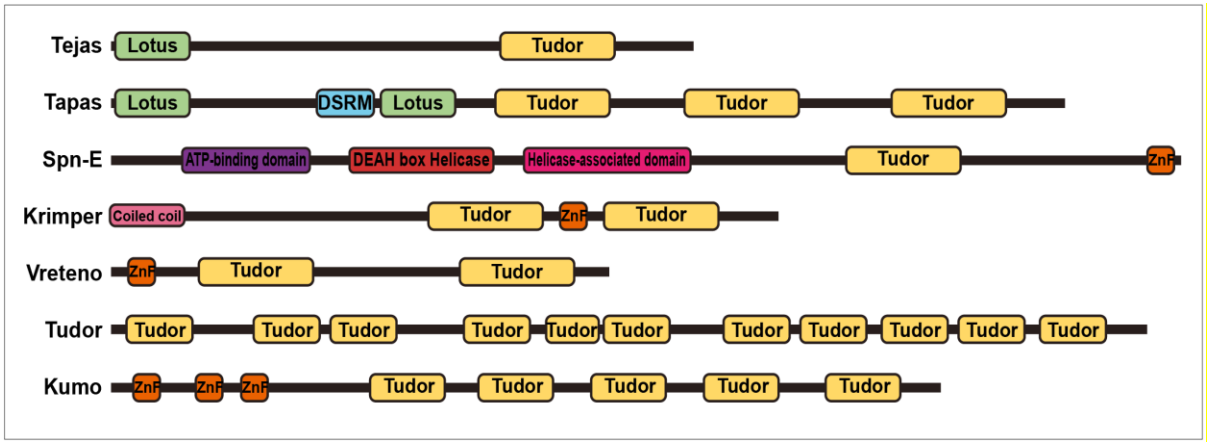


Figure 1.4. Schematic features for nuage proteins.

Nuage proteins in the thesis have one or several number of Tudor domains in the structures.

Studies have shown that the nuage include many components: Vasa (vas), a highly conserved DEAD-box containing helicase, has consistently been recognized as a marker of the germline cells that are localized nuage during early stages of oogenesis and accumulate to the posterior polar plasm at the later stage of the oocyte (Liang, Diehl-Jones, and Lasko 1994); Spindle-E (Spn-E), another DExH-box ATP-binding RNA helicase plays a central role during spermatogenesis and oogenesis (Gillespie and Berg 1995); two PIWI proteins, Aub and Ago3 (Harris and Macdonald 2001; Gunawardane et al. 2007); and the Krimper (Krimp), Tejas (Tej), Tudor (Tud), Tapas (Tap), Kumo, and Vreteno (Vret), a group of Tudor domain-containing proteins (Golumbeski et al. 1991; Lim and Kai 2007; Patil and Kai 2010; Zamparini et al. 2011; Anand and Kai 2012; Patil et al. 2014)(Figure 1.4). Acting as the core factors of the Ping-Pong pathway, the Aub and Ago3 are assisted by the other nuage components to slice the precursor. Most of them were indispensable in the piRNA biogenesis. The loss of these components collapses the piRNA generation massively (Malone et al. 2009).

Studies about Molecular function have gradually revealed the features of these proteins in the complex. For example, Vas functions with UAP56 to direct the precursors span the nuclear envelope (Zhang et al. 2012). It is reported that the ATP hydrolysis activity of Vasa is critical for disassembling the Siwi-Ago3 RNA-protein complex, which facilitates the proper Siwi-Ago3 ping-pong cycle in the BmN4 cell line established from the silkworm ovary (Xiol et al. 2014). Remarkably, the krimp is a well-studied protein that tightly connects with the two slicers Aub and Ago3. The krimp harbors two distinct Tudor domains in the C-terminal identified as the Tud1 and Tud2. The Tud1 strongly binds unmethylated piRNA-unloading Ago3 protein, while the Tud2 specifically recognizes methylated arginines posited at the N-terminal of the Aub. The mature piRNA loading causes a conformational change of Aub protein, thus exposing its inaccessible N terminal for the binding to Krimp. This binding is

regulated by arginine methylation and relies on the organization of a proper Ping-Pong cycle (Webster et al. 2015; Wang et al. 2015; Sato et al. 2015; Huang et al. 2021). Kumo ensures the heterotypic ping-pong pathway by preventing the loading of Ago3 cleavage piRNA onto the Piwi, leading a correct loading of these products to the Aub (Wang et al. 2015) (Figure 1.5 A). However, enough knowledge is missing about the function of other components in the complex, such as the Tej and Spn-E, though the loss of each of them causes a massive reduction of the piRNAs, providing the evidence that they are firmly involved in the piRNA biogenesis.

1.8.2 Yb body, a site for piRNA biogenesis in Somatic cells

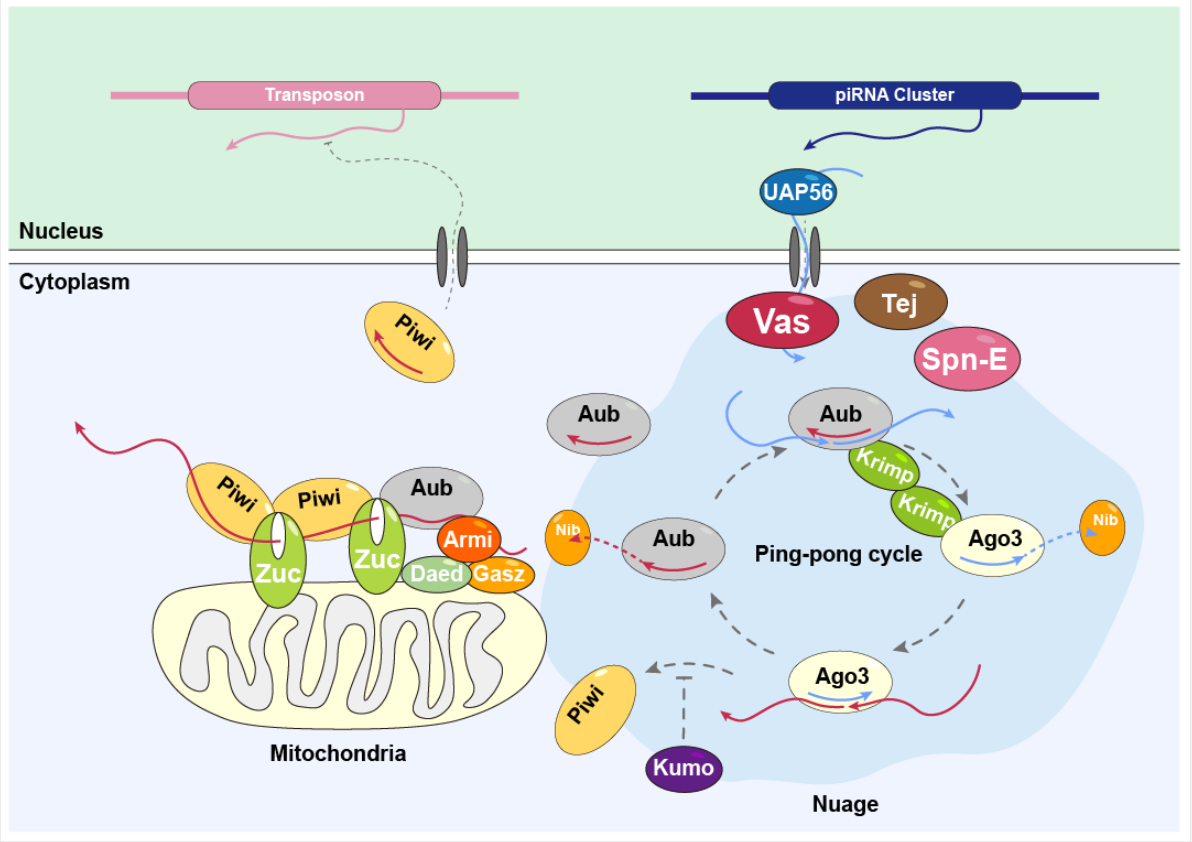
Unlike the ping pong cycle-dependent piRNA biogenesis pathway in nurse cells, piRNA phasing processing controlled by Piwi-Zuc in the somatic cells relies on the Yb-bodies localized around perinuclear cytoplasmic regions where the endonuclease Zucchini (Zuc) is localized to the surface of mitochondria with several co-factors.

The Yb bodies formed by several characterized proteins support the piRNA biogenesis: Sister of Yb (SoYb), Vreteno (Vret), Shutdown (Shu), and Armitage (Armi) (Saito et al. 2010; Olivieri et al. 2010; Zamparini et al. 2011; Handler et al. 2011; Preall et al. 2012; Huang et al. 2014). The Yb protein, a DEAD-box containing helicase, is a core component of the Yb-body that licenses the further processing of precursors and recognize the cis-acting Tj-cis RNA elements harbored in *flamenco* cluster-derived transcripts (Homolka et al. 2015; Ishizu et al. 2015; Pandey et al. 2017). The precursors are further loaded to Piwi proteins with the help of Shutdown that interacts with the HSP83 as and cochaperone (Preall et al. 2012; Olivieri et al. 2012).

After processed by Yb-bodies, the precursors need to be translocated to the Zucchini-dependent phasing machinery. The newly characterized factor Daedalus (Daed) interacts with

the germ cells specific protein Gasz provides an anchoring platform located on the mitochondrial outer membrane, allowing Armi to provide stability. The Armitage-bound RNAs containing the 2′,3′-cyclic phosphate at their 3′ end show they are immature piRNA loaded on the Piwi protein, which contains phosphorylated 5’ end (Saito et al. 2010). Once the Armi brings the intermediate Piwi–piRISC precursor (pre-piRISC) from Yb bodies to mitochondria, the phasing step starts with the machinery in the proximal site to Zucchini (Nishimasu et al. 2012; Wang et al. 2014; Matsumoto et al. 2016; Munafò et al. 2019). Armi performs as a helicase and bounds explicitly to piRNA precursors in the *Drosophila* ovarian somatic cells (OSCs), like the transcripts from the *flamenco* piRNA cluster. The ATP hydrolysis-defective mutant Armi fails to unwind the RNA, leading to the failure of piRNA production. This evidence suggests that Armi is directly involved in the phasing processing of piRNAs with binding and unwinding the piRNA intermediate (Ishizu et al. 2019) (Figure 1.5 B).

A


Transposon
piRNA Cluster
Nucleus
Cytoplasm
UAP56
Vas
Tej
Spn-E
Piwi
Aub
Armi
Zuc
Daed
Gasz
Krimp
Ago3
Nib
Ping-pong cycle
Mitochondria
Kumo
Nuage

B

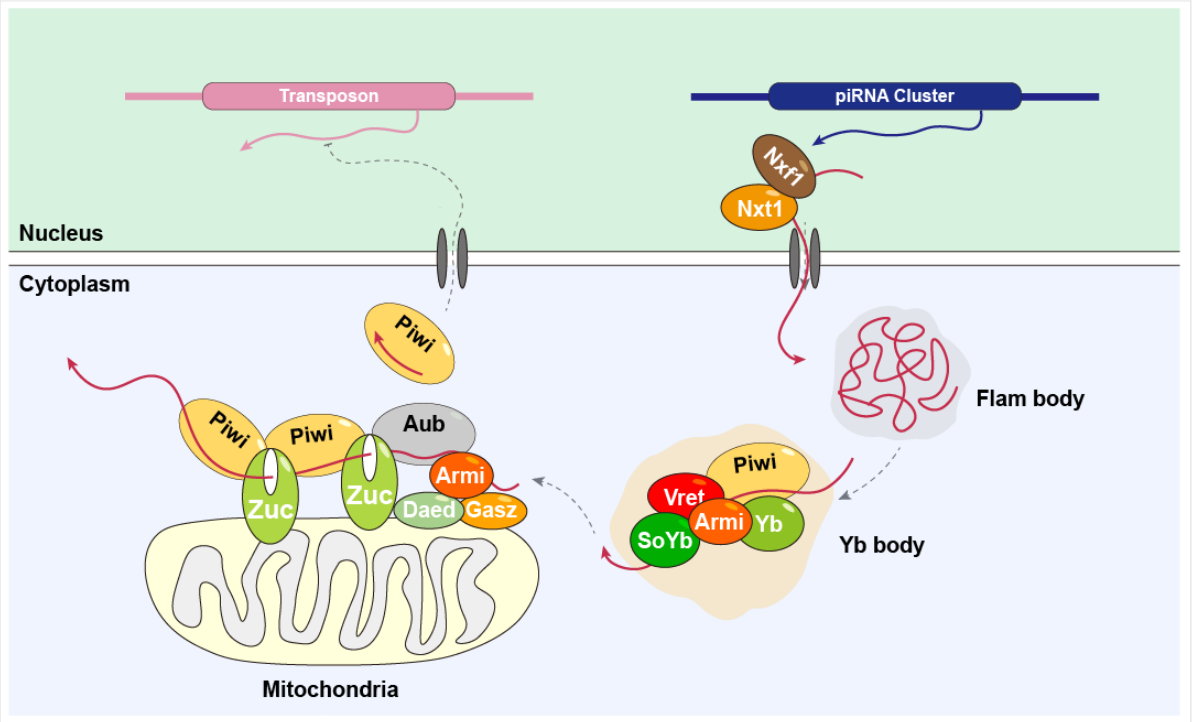

Transposon
piRNA Cluster
Nucleus
Cytoplasm
Nxf1
Nxt1
Piwi
Aub
Armi
Zuc
Daed
Gasz
Flam body
Vret
SoYb
Yb
Yb body
Mitochondria

Figure 1.5. piRNA Biogenesis Machinery.

(A) piRNA biogenesis machinery in *Drosophila* germline cells. In the nuclear pore, once precursors span the nuclear envelope and have been delivered to the nuage by coordinating UAP56 and Vas, Ago3 or Aub recognize and cleavage precursors alternatively. The Zuc and Nbr participate in intermediate piRNAs' 3’ end formation. Krimp connects two slicers Aub and Ago3, as a scaffold protein. Kumo ensures the heterotypic ping-pong pathway by preventing the loading of Ago3 cleavage piRNA onto the Piwi. Spn-E and Tej act as the Ping-Pong pathway's assistant factors, which is indispensable in the piRNA biogenesis, but the function remains elusive. The ping-pong loop contact with the mitochondria-located phasing pathway through the translocation of sliced precursor carried by Piwi and Armi. Mitochondria surface anchored Gasz/Daed complex recruits Armi-RNA complexes to mitochondria and comply with the precursor phasing pathway.
(B) piRNA biogenesis machinery in *Drosophila* ovarian somatic cells. The exported uni-strand precursors will be moved and stored in the Flam bodies, then been processed to mature piRNAs in adjacent Yb bodies. The Yb, SoYb, Armi, Vret proteins composed the Yb body. Yb binds cis-acting RNA elements in flamenco precursors further loaded to Piwi proteins. Phasing of piRNA precursors will happen in the mitochondria, similar to the germline cells. Adopted and modified from (Yamashiro and Siomi 2018)

1.9 The Tudor domain-containing proteins in nuage

Notably, the nuage components included many Tudor domain-containing proteins, the Krimp, Tej, Tud, Tap, Kumo, Vret, and even the ATP-binding helicase Spn-E harbors one to several Tudor domains (Figure 1.4). The Tudor domain was firstly identified from the protein Tudor in the *Drosophila* (Boswell and Mahowald 1985; Golumbeski et al. 1991; Ponting 1997), as a highly conserved motif, the Tudor domain can recognize symmetrically dimethylated arginines (sDMA) (Friesen et al. 2001; Brahms et al. 2002; Côté and Richard 2005; Kim et al. 2006). It was further reported that the Tudor domain proteins recognize sDMA modification on both Aub and Ago3 induced by methyltransferase dPRMT5 (Kirino et al. 2009). Besides the well-known sDMA binding ability, the Tudor domain has also been revealed to promote aggregated formation through sDMA binding, forming non-membranous cellular compartmentalization that contributes to the functional subcellular organization. (Courchaine et al. 2021).

Since the Tudor domain-containing proteins are frequently attendant in the nuage and Yb bodies, the sDMA interaction-mediated aggregates by Todor proteins may contribute to the organization of the granulated non-membranous organelle in both nuage and Yb bodies.

1.10 Lotus domain-containing proteins in piRNA biogenesis

The LOTUS domain, a highly conserved domain in bacteria and eukaryotes (Anantharaman, Zhang, and Aravind 2010), is another eminent domain commonly present in germline cells. The known Lotus domain proteins in the *Drosophila* are Oskar, dMrf1, Tej, and Tap. The dMrf1, Tej, and Tap are evolutionary highly conserved, homologs of them have been identified in the vertebrates as Mrf1, TDRD5, and TDRD7. The posterior oocyte pole localized Oskar mRNA was translated into long and short isoforms, gathering and germ plasm proteins and up to 200 maternal mRNAs as a core factor of the pole granules that contribute to the

germ cell and body axis formation. The dMrf1 is the *Drosophila* homolog of Meiosis arrest female 1 (MARF1), essential for gametogenesis as a regulator in oocyte maturation (Zhu et al. 2018; Kawaguchi, Ueki, and Kai 2020). Tej and Tap are also germline cell-specific proteins located to the nuage in the nurse cells that work synergistically for piRNA production in *Drosophila* germline cells for the maintenance of the germline, though the details remain unknown (Patil and Kai 2010; Patil et al. 2014). Both Tej and Tap are the Tudor domain protein and harbored one Lotus domain in their N terminal and involved in the piRNA biogenesis pathway (Figure 1.4) (Patil and Kai 2010; Patil et al. 2014).

Currently, it has been demonstrated that the Lotus domain of various species has a strong interaction with the C-terminal RecA-like domain of DEAD-box RNA helicase Vasa, which can stimulate the ATP hydrolysis activities of Vas *in vivo* (Jeske, Müller, and Ephrussi 2017). Crosslinking immunoprecipitation also revealed that TDRD5, the mouse homolog of Tej, directly binds to pachytene piRNA precursors (Ding et al. 2018). With the *In vitro* binding experiment, two of the three Lotus domains in the TDRD5 were explicitly bound to RNA G-quadruplexes (G4), a widespread guanosine-rich stacking helical RNA structure that presents on the pachytene piRNA precursors (Ding et al. 2020). These current studies also pointed out that the Lotus domain in the *Drosophila* protein Tej, Tap, and Osk have similar ATP hydrolysis activities that stimulate helicase Vas. Meanwhile, the Tej and Osk lotus can bind to RNA G4 structure, but not Tap. Analyses about these domains provide insight into the molecular functions of the mysterious nuage components.

1.11 Liquid-Liquid phase separation in fly oogenesis

The numerous canonical organelles in eukaryotic cells are isolated from the cytosol by the lipid bilayer membrane. Recently, many so-called non-membranous organelles have been recognized, such as supramolecular assemblies of proteins and RNA molecules formed ’phase

separation' driven by physicochemical forces. Phase separation is a well-studied phenomenon in polymer chemistry. The supersaturated components spontaneously separate into two phases by forming the liquid-liquid phase separation (LLPS), isolating the dense and dilute phases, and stably coexisting (Gleason et al. 2018; Sankaranarayanan and Weil 2020). Well-studied examples are the P granules at the embryos of *C. elegans* and the nucleoli of *X. laevis* oocytes which show liquid-like properties due to the phase separation mediated formation. The round appearance of these liquid-like structures results from minimizing surface tension. The fusion/fission and components exchange events indicate their high dynamics (Brangwynne et al. 2009; Brangwynne, Mitchison, and Hyman 2011; Boeynaems et al. 2018). Several phase separation regulating granular formation had been unearthed during the oogenesis of *Drosophila*: P bodies enriched at the anterior of the oocyte that facilitates RNA storage and translational regulation(Lin et al. 2008; Weil et al. 2012); the sponge bodies package and transport translationally repressed maternal mRNAs (Wilsch-Bräuninger, Schwarz, and Nüsslein-Volhard 1997; Snee and Macdonald 2009); polar granules located at the posterior of the oocyte which compartmentalizes factors that required for the proper formation of the embryonic germ cells (Harris and Macdonald 2001; Snee and Macdonald 2004); and the perinuclear localized nuage in nurse cells as a processing site for the piRNA biogenesis (Liang, Diehl-Jones, and Lasko 1994; Snee and Macdonald 2004; Sankaranarayanan and Weil 2020).

Enriched RNA molecules and RNA-binding proteins facilitate the Liquid-liquid phase separation that establishes many biomolecular condensates. Many proteins that contain multiple RNA-binding domains, as well as intrinsically disordered regions (IRDs), which have a low amino acid sequence complexity, had been identified in the RNA containing granules of

the neuron disease studies (Dunker et al. 2001; Li et al. 2012; Kato et al. 2012; King, Gitler, and Shorter 2012).

The IDR can be found in the polar granules and the nuage components like Osk, Vas, Aub, Ago3. The N terminal harbored long IDR in Vas, and its homologs were reported to ensure the phase separation. The condensate formation is regulated by the methylation status of IDR and shows different solubilization to double-strand or single-strand DNA (Nott et al. 2015). Additionally, the core component of polar bodies, Osk protein, contains IDR in its middle part and exhibits liquid-like or hydrogel-like properties that coordinate Vas (Kistler et al. 2018). Recent studies in *C. elegans* identified several Lotus or Tudor domain proteins that maintain their interaction and proper localization to the P granules for piRNA processing like nuage in *Drosophila* (Cipriani et al. 2021; Price et al. 2021; Marnik et al. 2021).

As the nuage is an assembly of many Lotus or Tudor domain-containing proteins that gather RNAs or form condensates, it is plausible that liquid-liquid phase separation contributes to the orderly processing of piRNA precursors in the nuage. However, to support this hypothesis, further investigation will be needed.

1.12 Thesis Overview

Through the current studies, the process of the piRNA precursor production and piRNA maturation, most of the molecular mechanisms of piRNA biogenesis have been highly investigated. The details of this stepwise processing of this pathway have been characterized. However, the detailed molecular functions of assistant factors in the ping-pong cycle remain elusive. In this thesis, I focused on the molecularly unknown function protein Tejas which sustain the proper piRNA biogenesis in germline cells of *Drosophila melanogaster*.

To explore the molecular function of Tejas in the piRNA biogenesis pathway, firstly, I made the Knock-In and transgenic fly lines that encoded the fluorophore-conjugated Tejas

and relevant proteins. Moreover, I performed further domain analysis of Tejas *in vivo* and *in vitro* and newly identified a unique motif in Tejas that recruits Spn-E, highlighting Tejas functions as an aggregator for the helicase Vas and Spn-E in the nuage granules. Lastly, I found that the intrinsically disordered region in Tej contributes to the dynamic organization of nuage.

This work specified Tejas's molecular-level function and gave some clues for further understanding of nuage organization. Meanwhile, it provided powerful toolkits for applying biochemical or genetic experiments to decipher the molecular function of nuage components.

2. MATERIALS AND METHODS

2.1 Fly stocks

D.melanogaster was used in the current study. Either *y w* or the respective heterozygote was used as a control. The mutant alleles combinations used in the study were *tej*$^{48–5}$*(Patil and Kai 2010), vas*^{PH165}*/Df* (Styhler et al. 1998), *spn-E*^{616}*/Df* (Ott, Nguyen, and Navarro 2014), *krimp*^{f06583} (Lim and Kai 2007), *nxf3*$^{\Delta}$ (Kneuss et al. 2019). The fly lines that express *NGT40*-Gal4, *nos*-Gal4 VP16, and *Traffic jam*-Gal4 deriver were obtained from the Bloomington *Drosophila* Stock Center. Knock-In fly lines *vas*$^{mCherry.HA.KI}$ (DGRC# 118618), *vas*$^{EGFP.KI}$ (DGRC# 118616) and *aub*$^{EGFP.KI}$ (DGRC# 118621) (Kina et al. 2019) were obtained from the *Drosophila* Genetic Resource Center at Kyoto Institute of Technology, Japan. All stocks were maintained at 25°C.

2.2 Generation of Transgenic Fly Lines

Generation of Tej-GFP,mKate2-Ago3, and Spn-E-mKate2 Knock-In fly lines through CRISPR-Cas9 induced double-strand breaks restored by the homology-directed repair (HDR) in the presence of donor plasmids. Two guide RNAs were designed to direct the Cas9 proteins to the regions flanking the start/stop codon of the target genes to induce big scale of double-strand breaks (Guide RNA sequences: Tej-GFP gRNA1: GATCGCTCATAGAAACTGGT, gRNA2: GTGCATAGATTTCTATTATA; mK2-Ago3 gRNA1: TAATAAAAATGCTGGCAATA, gRNA2 TGTGTGTTTCAGAGCATGTC; Spn-E-mK2 gRNA1: GATCACGATGCAATATGGTC, gRNA2: GAACGATGTAACCATTCTTAT). Donor vectors contain the GFP or mKate2 coding sequence flanked by 1kb homology arms adopted from both 3’ and 5’ sides of the insertion site. Plasmids containing coding sequences for GFP and mKate2 were obtained from Addgene. Guide RNA/Cas9 expression plasmid pDCC6 (Gokcezade, Sienski, and Duchek 2014) and donor plasmid were injected into the cleavage stage *y w* embryos with a final concentration of 120

ng/ul for each plasmid. The Knock-In events positive founder were confirmed by single fly genome PCR genotyping and DNA sequencing, then crossed with double balancer fly line *w; Pin/CyO; TM3 Sb/TM6B Tb* for fly line estimation. Tej-GFP, mKate2-Ago3, and Spn-E-mKate2 Knock-In flies were crossed with corresponding loss-of-function allele $tej^{48\text{-}5}$, $Ago3^{t2}$, and $Spn\text{-}E^{616}$ for checking the functionality of endogenies fusion proteins.

The GFP-Tej, GFP-Tej ΔLotus, GFP-Tej ΔTudor, GFP-Tej ΔIDR, GFP-Tej ΔSRS, GFP-Spn-E, and GFP-Spn-E ΔNLS transgenic fly lines were generated by PhiC31 integrase-mediated transgenesis system. Transgenic constructs for injection were generated using the cDNAs obtained by reverse transcription from ovarian RNA of *y w* flies. DNA fragments encoded the Fluorescence-protein conjugated target proteins were generated by Phusion High-Fidelity PCR kit (New England Biolabs) and In-Fusion HD Cloning kit (Takara Bio). Transgenes were recombined into the pUAS-K10 *attB* plasmid backbone, which linearized restriction endonuclease. After the transgenic constructs were injected into the embryo of *attP*-containing strains, the target sequences were directed and integrated to docking sites on the 2nd or 3rd chromosome (BDSC #25709 and BDSC #25710). Rescue lines were generated by crossing or recombinant the transgenic construct into the $tej^{48\text{-}5}$ and $Spn\text{-}E^{616}/Df$ background and derived by the germline-cell specific drivers *NGT40-Gal4*, *nos-Gal4-VP16,* and ovarian somatic cell driver *Traffic jam-Gal4 (Tj-Gal4)*. The plasmid constructs used for generating the Knock-In and transgenic fly lines are shown in Table 3.

2.3 Antibody generation

Antibodies for HeT-A-Gag protein, Ago3, and Tej were generated in this study. For generating antibodies against HeT-A-Gag protein, a DNA fragment that encoding 201 amino acids HeT-A-Gag antigen peptide was amplified from the cDNA which derived from *krimper* mutant ovaries. This fragment was cloned into pENTR/D-TOPO plasmids and recombined into

the pDEST17 plasmid to express the His-HeT-A-Gag antigen peptide. The DNA fragments that encode the N terminal of Tej (amino acids 1 to 110) and Ago3 (amino acids 1 to 150) were cloned into pENTR/D-TOPO plasmids and recombined into the pDEST17 and pDEST15 (Invitrogen), respectively (Primer sequences for the cloning HeT-A-Gag antigen peptide, forward primer: CACCCCCTACTGGAAAAGCTGAAC, reverse primer: CTACAGGGCATCCTTTGT ACGCGCT. Primer sequences for the cloning Tej peptide, forward primer: ATGGATGATGGAG GGGAGTT, reverse primer: CTCGGAGGCGTAGCAATA. Primer sequences for the cloning Ago3 peptide, forward primer: ATGTCTGGAAGAGGAAA, reverse primer: TTACACTTCGTAATTAA AAA). All the His-HeT-A-Gag , HIS- or GST-tagged Tej and Ago3 antigen peptides were expressed by the E. coli expression system. The plasmid was transformed into E. coli strain BL21 (DE3) and cultured at 37 °C in LB medium. Add IPTG at a final concentration of 0.2 mM to induce protein expression when the OD600 reached 0.6-0.8, culture the cells in 25°C. Centrifuged cells were resuspended with 0.5M NaCl containing PBS buffer.

The soluble His-HeT-A-Gag protein were used to immunize rabbits (TLL Animal Facility). The anti-serum was separated from bleeds and stocked on -20°C with presences of 50% (v/v) glycerol and 0.1% (w/v) sodium azide. The soluble GST-Tej (1-110aa) peptides were extracted from the sonication homogenized cells lysate supernatant by the GST-Accept beads (COSMOGEL). The insoluble HIS-Tej (1-110aa), GST-Ago3 (1-150aa), and HIS-Ago3 (1-150aa) was extracted from cell debris after urea denaturing. GST-Tej (1-110aa) containing solution and SDS-PAGE electrophoretic separated gel band containing Ago3 (1-150aa) antigen peptides were used to immunize the rats to generate the antiserum (Eve Bioscience). The antibodies were further purified from the antiserum using the His-tagged Tej and GST-tagged Ago3 antigen peptide. His-Tej (1-110aa) and GST-Ago3 (1-150aa) were loaded into SDS-PAGE gel for electrophoresis separation, then transferred to the PVDF membrane (WAKO). The His-

Tej (1-110aa) and GST-Ago3 (1-150aa) peptide-containing region of PVDF membrane were sliced into pieces and incubated with the GST-tagged Tej and HIS-tagged Ago3 serum at 4°C overnight with rotation. After incubation, the membrane pieces were collected and washed in 1% (v/v) PBST for 2hrs at room temperature. Then the pieces were quickly rains with 0.1M acid-glycine (pH2.5), and the antibody-containing elution was collected. The antibodies containing elution were neutralized to pH7.0 by NaOH and stocked in 50% (v/v) glycerol at -20°C. Both the rat anti-Tej and rat anti-Ago3 worked for immunostaining and immunoblotting. The plasmid constructs used for expressing the peptides are shown in Table 3.

2.4 Western blotting

Ovaries were collected and homogenized in lysis buffer containing 30 mM HEPES (pH7.4), 80mM KOAc, 2mM DTT, 10% (v/v) glycerol, 2mM Mgcl2, 0.1% (v/v) TritonX-100. After centrifugation at 200,000 ×g for 10 min at 4°C, take the supernatants as samples. Protein samples were loaded into each lane of 10-15% SDS-PAGE gels. After electrophoresis, western blotting was performed with standard protocols with the following antibodies. The primary antibodies used in this study are listed up in Table 1. Secondly antibodies were HRP-conjugated goat anti-guinea pig (DAKO, 1:1000), anti-rat (1:1,000), anti-mouse (Invitrogen, 1:1,000), and goat anti-rabbit (BioRad, 1:3000), diluted and stored in the Signal Enhancer reagent HIKARI (NACALAI TESQUE). Chemiluminescence was induced by the Chemi-Lumi One reagent kit (NACALAI TESQUE), and Immunoreactive bands were detected using Chemi Doc (Touch Bio-Rad Laboratories).

2.5 Small RNA Immunoprecipitation

For IP of Aub- and mKate2-Ago3 protein-piRNA complexes, 200 ovaries were dissected manually from adult flies in chilled PBS and homogenized lysis buffer containing 20 mM Tris-HCl (pH7.4), 200 mM NaCl, 2mM DTT, 10% (v/v) glycerol, 2mM Mgcl2, 1% (v/v) TritonX-100,

1x cOmplete protease inhibitor cocktail (Roche) and 1% (v/v) RNaseOUT recombinant ribonuclease inhibitor (Invitrogen). After centrifugation at 200,000 ×g for 10 min at 4°C, the supernatant was collected in new ep tubes and kept on ice. The same centrifugation was repeated 3 times to remove the lipid contamination. The antibody was mixed with the purified lysate and incubated at 4°C for 2 h with rotation. Then the Dynabeads Protein G/A (Invitrogen) 1:1 mixture was added to the lysate-antibody mixture and incubated at 4°C for 1 h with rotation. Mouse anti-Aub antibody (1:20) (Patil and Kai 2010) and mouse anti-mKate2 (Evrogen, 1:200) were used to IP the Aub-protein-piRNA complexes and mKate2-Ago3 protein-piRNA complexes from the ovary lysate, respectively. After incubation, the magnet beads were collected and washed at least 4 times using washing buffer contains 20 mM Tris-HCl (pH7.4), 400 mM NaCl, 2mM DTT, 10% (v/v) glycerol, 2mM Mgcl2, 1% (v/v) TritonX-100, 1x cOmplete protease inhibitor cocktail (Roche) and 1% (v/v) RNaseOUT recombinant ribonuclease inhibitor (Invitrogen). 10% Sample mixed with the SDS containing sample buffer and heated at 95°C for 5 mins and then loaded onto SDS-PAGE gels for WB checking the protein immunoprecipitation efficiency. 90% of suspension contains beads-Protein complexes then mixed with TRIzol LS (Invitrogen), and the protein-binding small RNA will be extracted according to the standard manufacturer's protocol of TRIzol LS.

2.6 Radioisotope

For visualization with radioisotope labeling, small RNAs from immunoprecipitation were treated by phenol-chloroform extraction and ethanol precipitation. Purified small RNAs were labeled with 32P-γ -ATP using T4 polynucleotide kinase (Thermo Fisher Scientific). Purified Long RNA fragments were treated with alkaline phosphatase (New England Biolab), which nonspecifically catalyzes the dephosphorylation of 5´ ends before 32P-γ -ATP labeling. Then the solution was filtered with the G-25 column (GE Healthcare) to remove the excess 32P-γ -

ATP and then treated by phenol-chloroform extraction twice. Finally, the radioisotope labeled RNA was purified twice by ethanol precipitation and resuspended in RNase-free water. After electrophoretic separation by 15% urea-containing denaturing polyacrylamide gel in ×0.5 TBE, radioisotope signals were captured to the radiosensitive film plates (Fuji Film) and analyzed by Amersham Typhoon scanner (GE).

2.7 Crosslinking Immunoprecipitation

Crosslinking IP was performed for Tej-GFP, Spn-E-mKate2, and Vas-GFP to detect the protein interactions. Ovaries were dissected manually from adult flies in ice-chilled PBS and fixed using 0.1% (w/v) paraformaldehyde for 20 mins on ice. Quenching the fixed ovaries by 125 mM glycine for 20 mins and then homogenized in CLIP lysis buffer that contains 50 mM Tris-HCl (pH 8.5), 150 mM KCl, , 5 mM EDTA, 1% (v/v) TritonX-100, 0.1% (w/v) SDS, 0.5 mM DTT and 1x cOmplete protease inhibitor cocktail (Roche). The lysate was incubated at 4°C for 20 min with rotation for sufficient lysis, followed by sonication with Bioruptor (Sonicbio). After centrifugation at 200,000 ×g for 10 min at 4°C, the supernatant was collected in new Eppendorf Protein LoBind tubes and diluted by adding equal volumes of CLIP wash buffer that contains 25 mM Tris-HCl (pH 7.5), 150 mM KCl, 5 mM EDTA, 0.5% (v/v) TritonX-100, 0.5 mM DTT and 1x cOmplete protease inhibitor cocktail (Roche). The diluted lysate was pre-cleaned by Dynabeads Protein G/A (Invitrogen) 1:1 mixture for 1 h at 4°C and incubated with antibody overnight at 4°C. The mouse anti-GFP (Thermo Fisher Scientific, 3E6, 1:500) and mouse anti-mKate2 (Evrogen, 1:500) were used to IP the GFP- or mKate2 fused proteins. CLIP washing buffer equilibrated Dynabeads Protein G/A (Invitrogen) 1:1 mixture was added to the lysate-antibody mixture and incubated at 4°C for 3 hrs with rotation. After incubation, the magnet beads were collected and washed at least 4 times using CLIP washing buffer. When required a harsh binding and washing condition, the potassium salt concentration of the CLIP washing

buffer was adjusted up to 1M. Beads were mixed 1:1 with the SDS containing 2x sample buffer, heated at 95°C for 5 min, then loaded onto 12% SDS-PAGE gels for Western Blotting.

2.8 RT-qPCR

Total RNA was extracted from the ovaries dissected from the two days yeast fatten upped female *Drosophila* with TRIzol LS (Invitrogen). Operation is according to the standard manufacturer's protocol. DNase I (Invitrogen) treated RNAs are reverse transcribed using the SuperScript III system (Invitrogen). Oligo d(T)20 and hexadeoxyribonucleotide mixture primer were used for the reverse transcription reaction. qPCR was performed using KAPA SYBR Fast qPCR Master Mix (KAPA biosystems). All the target expression was normalized to rp49. The primer sequences for detecting transposon transcripts and piRNA cluster transcripts are shown in Table 2.

2.9 Cell culture and plasmid transfection

Drosophila Schneider S2 cells were grown at 26°C in 10% (w/v) Fetal Bovine Serum (FBS) supplemented Schneider medium, with the presence of 50-100 U penicillin and 50-100 µg streptomycin. Plasmids used for transfection were generated using the Gateway cloning system (Life technologies). Utilizing the Gateway LR in vitro recombination reaction and the *Drosophila* Gateway Vector Collection (DGVC) destination vectors, a series of plasmids were generated, which express the N terminal fluorescence protein fusion proteins driven by the Actin5C promoter. Modified based on the pAGW, the destination vector pAKW suitable for expressing the mKate2 N-terminal fusion protein was generated in this research. S2 cells were grown to 50-60% surface cover rate before being transfected by using HilyMax (Dojindo Molecular Technologies, Inc) according to the manufacturer's instructions. The optimized ratio of the plasmid: HilyMax was 1:5 for all the plasmid transfection. The plasmid constructions used in this study are shown in Table 3.

2.10 S2 cell fluorescence observation

After the plasmid transfection, the fluorescence-tagged proteins well expressed cells were resuspended and moved to the concanavalin A pre-coated coverslips, then cultured at 26°C for at least 20mins for an efficient adhesion. Then fixed for 15 min in 4% (w/v) paraformaldehyde, permeabilized for 10 min in PBX (PBS with 0.2% (v/v) TritonX-100) and washed for 10 min by PBX twice. Apply DAPI (1:1000) staining at room temperature for 10mins, and rains with PBS to remove the DAPI solution. Equilibrated in Fluoro-KEEPER Antifade Reagent (NACALAI TESQUE) for 10mins before mounting.

2.11 Immunofluorescence staining

Ovaries were dissected in ice-chilled PBS, fixed in 4% (w/v) paraformaldehyde for 10mins on ice, and washed by PBX (PBS with 0.2% (v/v) TritonX-100) 2 times. Ovaries were then blocked by 4% (w/v) BSA in PBX for 30mins and incubated with primary antibody diluted in 0.4% (w/v) BSA containing PBX at 4°C overnight. Ovaries were washed 3 times with PBX after overnight antibody incubation. Then incubated with secondary antibody diluted in 0.4% (w/v) BSA containing PBX at room temperature for at least 1hrs, followed by three washes with PBX. Continuously, DAPI (1:1000) staining at room temperature for 10mins, and rains with PBS to remove the DAPI solution. The ovaries will be equilibrated in Fluoro-KEEPER Antifade Reagent (NACALAI TESQUE) for 10mins before mounting. The antibodies used for immunostaining are listed in Table 1. Secondary antibodies were Alexa Fluor 488-, 555-, 633-conjugated goat anti-rabbit, anti-mouse, anti-rat, or anti-guinea pig IgG(Molecular Probes, Eugene, Oregon, USA), 1:200 diluted in 0.4% (w/v) BSA containing PBX as working solution.

2.12 RNA *in situ* hybridization chain reaction (HCR)

The probes were designed to target the transcripts derived from the unique regions at *cluster 38C* (Chr2L: 20104855..20107574) and *42AB* (Chr2R: 6322410..6323756). A

commercial kit with a customized DNA probe set and essential reagents were purchased from Molecular Instruments, Inc. The protocol was modified according to the published article (Slaidina et al. 2020). Ovaries were fixed in 4% formaldehyde for 20 minutes and washed twice with PBST (PBS with 0.1% (v/v) Tween-20) at room temperature. Fixed samples were further dehydrated by a sequentially washing of 25%, 50%, 75%, and 100% (v/v) methanol in PBS for 5 min each on ice. Dehydrated ovaried were stored at −20°C overnight, then rehydrated by sequential washes with 100%, 75%, 50%, and 25% (v/v) methanol in PBS on ice on day2. Permeated for 2 hours in PBX (PBS with 0.2% (v/v) TritonX-100) at room temperature, followed with a post-fixation by 4% (w/v) paraformaldehyde for 20 min at room temperature. Washed twice with PBST (PBS with 0.1% (v/v) Tween-20) for 5 min on ice; washed with 50% (v/v) PBST and (v/v) 50% 5× SSCT (5× SSC with 0.1% (v/v) Tween-20) for 5 min on ice; washed twice with 5× SSCT for 5 min on ice. After washing steps, ovaries were equilibrated in the probe hybridization buffer for 5 min on ice; prehybridized in the probe hybridization buffer for 30 min at 37°C; 4 pmol of probe mixture was added to 0.5 mL of pre-warmed probe hybridization buffer, infiltrate the sample, and then hybridized overnight at light-avoiding 37°C shaker. After hybridization, ovaries were washed 4 times with probe wash buffer for 15 minutes each at 37°C shakers, then washed with 5× SSCT for 5 minutes each at room temperature. Next, the ovaries were equilibrated in a prewarmed amplification buffer for 5 min at room temperature. 30 pmol of dye conjugated hairpin were heated up separately to 95°C for 90 sec, then cool down at room temperature for 30 min at the light-avoiding box. Then cool the hairpins on ice for 10 sec and mix with 500 µL amplification buffer at room temperature. The ovaries were incubated with the freshly prepared hairpin solution overnight in a light-avoiding container at room temperature. Terminate the hairpin chain reaction by twice washing with 5× SSCT for 5 min, then wash twice with 5× SSCT for 30 min at room

temperature. DAPI (1:1000) and Wheat Germ Agglutinin (5μg/ml, Alexa Fluor 488 Conjugated, Thermo Fisher Scientific) were added to 5× SSCT in the first 30-min wash. Ovaries were equilibrated in Fluoro-KEEPER Antifade Reagent (NACALAI TESQUE) at room temperature before mounting. Further observation was done by LSM 900 with Airy Scan 2 (Choi et al. 2018; Slaidina et al. 2020). The sequences for designing the probe sets are shown in Table 4.

2.13 S2 cell Live imaging

Resuspend fluorescence positive cells by fresh pre-warmed growth medium after ~48 hours transfection, culture the S2 cells in a concanavalin A precoated multi-well glass-bottom culture chamber (MATSUNAMI) for over 30 min at 26°C. All images were recorded at 26°C by incubation modules advanced ZEISS LSM 900 with Airy Scan 2 using 63× oil NA 1.3 objectives.

2.14 Fluorescence Recovery After Photobleaching (FRAP)

S2 cells expressing N terminal GFP or mKate2 fused Tej, Tej truncations, Vasa, and Spn-E proteins were prepared as described above. The GFP signals were repeatedly bleached using a pulse of 488 lasers 50 times, and images were taken every second to record fluorescence strength. Initial 10 images were acquired to establish the levels of pre-bleach fluorescence. One single nuclear granule was bleached per cell to FRAP granulated GFP signals in S2 cells.

Imaging was performed as described above in ZEISS LSM 900. Fluorescence strength in timelines was analyzed using an online FRAP analysis tool easyFRAP-web (Rapsomaniki et al. 2012). A full-scale normalization procedure that corrects differences in bleaching depth among different experiments was used to normalize recovery curves. Maen curves of normalized datas fit a double term exponential fitting equation that was used to calculate the half time to full fluorescence recovery t1/2 (s) and show the percentage of maximum fluorescence recovery (Rapsomaniki et al. 2012).

2.15 Protein disorder prediction and conservation analysis

The intrinsically disordered region was defined using the IUPRED server (https://iupred2a.elte.hu/). The region containing residues with IUPRED scores more prominent than 0.5 was classified as an intrinsically disordered region (Dosztányi et al. 2005).

2.16 Analysis of small RNA libraries

tej$^{48\text{-}5}$ flies and their heterozygote siblings were used in this analysis. To extract the Ago3-bound small RNAs, the mK2-tagged Ago3 substituted the endogenous Ago3 in the *tej*$^{48\text{-}5}$ flies and their heterozygote siblings. Ovaries of female *Drosophilas* dissected on ice and snap-frozen in liquid nitrogen, then stock on -80°C. Small RNAs were extracted from the immunoprecipitated protein complexes that pull down from the ovary lysate. Purified Small RNAs were used for library generation for deep sequencing. Deep sequencing was performed on HiSeq3000 at Genome Information Research Center, Research Institute for Microbial Diseases of Osaka University. Libraries were normalized with noncoding RNAs including snoRNAs,snRNAs, miRNAs, and tRNAs. After trimming and removing the rRNA, snoRNAs, snRNAs, miRNAs, and tRNAs, only 23- to 29- nt reads that mapped to the piRNA cluster or transposable elements with best alignments were included in the analysis of precursor derived piRNA. The libraries were mapped to the *Drosophila* piRNA cluster sequence and transposable elements with up to three mismatched bases by Bowtie (Langmead et al. 2009) in Linux. piRNA cluster definition references to the (Brennecke et al. 2007), TE sequences were adopted from the Flybase (Release 6.32). The normalized numbers of piRNA clusters mapping reads were distributed to the position of the cluster sequence and visualized with pyGenomeTracks (Ramírez et al. 2018).

2.19 Tables

Table 1. List of primary antibodies used in this study.

Antibody Name	Source
Anti-Tej Rat polyclonal antibody	Current study
Anti-Vas Guinea pig polyclonal antibody	(Patil and Kai 2010)
Anti-Spn-E Rat polyclonal antibody	(Patil and Kai 2010)
Anti-Ago3 Rat polyclonal antibody	Current study
Anti-Piwi mouse monoclonal antibody	(Saito et al. 2006)
Anti-Ago2 Guinea pig polyclonal antibody	(Iki, Takami, and Kai 2020)
Anti-Myc Mouse monoclonal antibody	Wako, Cat.# 017-21871
Anti-eGFP Mouse monoclonal antibody	Invitrogen, Cat.# A-11120
Anti-mKate2 Mouse monoclonal antibody	Evrogen, Cat.# AB231
Anti-HetA Rabbit polyclonal antibody	Current study
Anti-Fibrillarin Rabbit polyclonal antibody	Abcam, Cat.# ab5821
Anti-Aub Guinea pig polyclonal antibody	(Lim et al. 2022, Being revised in 'Frontiers in Molecular Biosciences')

Table 2. List of primers used for qRT-PCR in this study.

Primer label	Primer sequence	Source
Rp49_Fw	ATGACCATCCGCCCAGCATAC	Current study
Rp49_Rv	CTGCATGAGCAGGACCTCCAG	Current study
Tubulin_Fw	GGTAACCGTCGAAATCAGTGTT	Current study
Tubulin_Rv	TGGCTTTTCTGCTATACGTGTC	Current study
38C #1_Fw	GAGACTTGCCGTTCCTTAG	Current study
38C #1_Rv	CCATCTGGAATGCAAACG	Current study
38C #2_Fw	TCCGTGACGGTTTAGCCCA	Current study
38C #2_Rv	AGGTTTCAAACCTTCCAG	Current study
42AB #1_Fw	CGTCCCAGCCTACCTAGTCA	(ElMaghraby et al. 2019)
42AB #1_Rv	ACTTCCCGGTGAAGACTCCT	(ElMaghraby et al. 2019)
42AB #2_Fw	CGCTGTTGAAAGCAAATTGA	(ElMaghraby et al. 2019)
42AB #2_Rv	GAGACCTTCGCTCCAGTGTC	(ElMaghraby et al. 2019)
Flam #1_Fw	ACGCTCAGGAAGGGATTTCA	Current study
Flam #1_Rv	AAACATGTCGTCTATCCATC	Current study
Flam #2_Fw	TCTCGGATAGAACTCTTCCC	Current study
Flam #2_Rv	TTGAACCTGTAGGCTAGGTA	Current study
HetA_Fw	ACAGATGCCAAGGCTTCAGG	(Piñeyro et al. 2011)
hetA_Rv	GCCAGCGCATTTCATGC	(Piñeyro et al. 2011)
TART_Fw	TTCTATCAACAGGCTGTCCACAGGTT	(Savitsky et al. 2006)
TART_Rv	CCTTCGTAGTCGGGTAGGATTATTCGT	(Savitsky et al. 2006)
TAHRT_Fw	CTGTTGCACAAAGCCAAGAA	(Chen et al. 2016)
TAHRT_Rv	GTTGGTAATGTTCGCGTCCT	(Chen et al. 2016)
I-element	TGAAATACGGCATACTGCCCCCA	(Klenov et al. 2011)
I-element	GCTGATAGGGAGTCGGAGCAGATA	(Klenov et al. 2011)

Table 3. List of constructs used in this study.

Name of construct	Source
pDEST-15 Tej	Current study
pDEST-17 Tej	Current study
pDEST-15 Ago3	Current study
pDEST-17 Ago3	Current study
pDEST-17 HeT-A-Gag	Current study
pDEST-mkate2	Current study
pDCC6-Tej sgRNA #1	Current study
pDCC6-Tej sgRNA #2	Current study
pGEM-Tej-GFP doner	Current study
pDCC6-Ago3 sgRNA #1	Current study
pDCC6-Ago3 sgRNA #2	Current study
pGEM-mK2-Ago3 doner	Current study
pDCC6-Spn-E sgRNA #1	Current study
pDCC6-Spn-E sgRNA #2	Current study
pGEM-Spn-E-mK2 doner	Current study
pENTR-Tej	(Lim and Kai 2007)
pENTR-Vas	(Lim and Kai 2007)
pENTR-Spn-E	(Patil and Kai 2010)
pENTR-Spn-E ΔNLS	Current study
pENTR-Tej ΔLotus	Current study
pENTR-Tej ΔTudor	Current study
pENTR-Tej ΔSRS	Current study
pENTR-Tej ΔIDR	Current study
pENTR-Tej 140-559	Current study
pENTR-Tej 295-559	Current study
pENTR-Tej 363-559	Current study
pENTR-Tej 101-337	Current study
pENTR-Tej 101-349	Current study
pENTR-Tej 101-362	Current study

pAGW-Vas	Current study
pAKW-Vas	Current study
pAGW-Spn-E	Current study
pAKW-Spn-E	Current study
pAGW-Spn-E ΔNLS	Current study
pAMW-Tej FL	Current study
pAGW-Tej FL	Current study
pAGW-Tej ΔLotus	Current study
pAGW-Tej ΔTudor	Current study
pAGW-Tej ΔSRS	Current study
pAGW-Tej ΔIDR	Current study
pAKW-Tej FL	Current study
pAKW-Tej ΔLotus	Current study
pAKW-Tej ΔTudor	Current study
pAKW-Tej ΔSRS	Current study
pAKW-Tej ΔIDR	Current study
pAKW-Tej 140-559	Current study
pAKW-Tej 295-559	Current study
pAKW-Tej 363-559	Current study
pAKW-Tej 101-337	Current study
pAKW-Tej 101-349	Current study
pAKW-Tej 101-362	Current study
pUASp-miniTurbo GFP-Tej FL	Current study
pUASp-miniTurbo-GFP-Tej ΔLotus	Current study
pUASp-miniTurbo-GFP-Tej ΔTudor	Current study
pUASp-GFP-Tej ΔSRS	Current study
pUASp-GFP-Tej ΔIDR	Current study
pUASp-GFP-Tej FL	Current study

Table 4. Regions of HCR-FISH probe sets targeting in this study.

Cluster 38C (Chr2L: 20104855..20107574) for probe sets designing.
1.CCTTGATTACGACTCTGTAGGCTCTCATCATAAACATAATAATAAACATAACTACATTAAATGTGATA
2.TATTAACTTCTGTAAACGGCGGCTAAAATCGGTACCCAGGGAACACCACGGGGAGG
3.TAAATGATTTTTCCATAAGGGGGAAAAAATTCCTGGCCTCTCTCGTGAACGAGTTGTCTGACCGAGAAGA
4.GTAGATGGGTCGGAGTTTTCATTAATTATCTGAACATCTGACTTGAAACTCTTTACCCTCCGTAAAGGATTTG
5.ATGAGATCAGGGCCTGGAGACTTGCCGTTCCTTAGTTTGAAAAATGGAGAGTATATCTGCTTGGTTGGTCTTGCAATAAAGCTCGTCCTCGGGTTTGTTGTTGATAAGGCCG
6.ATCGCGAAGCTATTATCCGCTAGAGCAGCTTGGTGGTGGGTGGGTGTGGTGGGTCCGTCTATTGGATGAGAGTG
7.TAAACTATAGCGACAGCTATGTCGATACTGGTTCTACGGTGCGAAAATACGAACGCACAATGCT
8.TTTTGTCAACCCCTGGGCTTAATGGGTAGTTTCATACTGTCACTATTTTTTTTATCAGCGTTGATCACGATGTTCCAGCG
9.ACGTTTCAATCCATTGGAATTCGTCCACGGCTGCGCACCTGGGAGAAGATGCGTCCTCGCCAGGTATTGAAGACGCCTCTGCGCCCTTATAACCACTATTTAGACGATTCACCAGACCACCTCATGCGAGTCCCCGATTAAGATTGGTGTCATTCCGGTGCAGCTCCAGGAATGCTAAATGGAGTTTTTCTGCATAATATTTATCCGATTGGTATGATTCTGCTGCTGATCAATGAGAAGAGGTGATATTTTTTTTCTCTTCACTCTGGACCTTGCG
10. AATCGAAGATGTCGTCCTGGACACCGCCTTCTCGGCCGGCAAGTTTACCTCATCCGAT
11. GAGCAGATGGCCGAGGAAATCCCGGATCGACATGGCGCAGGTAGAGCTCGTTTGATTTTGA

Cluster 42AB (Chr2R: 6322410..6323756) for probe sets designing.
TGACTCGCTACTGCGAGGGTTTCCCCTTTTATTCATTCTTACATAGGTTCTTATCATCTAATTATATA ATGTAGCGCTTGATGTCAAGTCGAATATAGGACGGTGGCAGATCCGTCCGTTTCTCCGTCAAAAG GCTTCGGAAAAATTTTTTTATATACGTCTTATTCCGGCACTCATTAAAGCTGAGCTTATTCCATTAT TCAATTCCAAGGAAAATTCAGATACAACCTTCAGTGGAAAATTGCCTGATTTTAATATAAATCTGC TCAGCCAAAACTGTATGCTCATACTTTTAATCATTATGCCCTCATACTTTTTAAGGTATAATGATGA CTGGATCTCGAACCTTCGTGATTCCTAGGTGTCTGTAGGTTGCCTTTTCACAAAGTTTCCGAAGTG TTGCTAACTTTGGTATGTAATTTGGTTTTGGGTCATCTTTCATATTTTCGGTGTCTTCCTTCCTCCAA AGCAGCTTTTTTTAATCAATTCCAAATTTTAAGTCGTCTGGATCCAGTGCACCGTCGTTTTTTATCA TTATTGGCACTGCTATTATTCTATTATTGGCACTGCTATCCAGCACATCAATGCCACATCGTTCTTTC ACCACAGTCATTCCGATTGATTCTGGCATAGAAATTAATGGTCTTCGCGCTAATTGGTCCTTTTAAT TGGTTACTTATATCCATCCAAATCGCTTTATCGCTTCTTGGCTTAAGAGTTCCAGTCCGTGTAAAAA TATTCGCACTGCAAAATACATCCCCTATGGTGGGGTTTGACTATGTTTTTTATGTTTTTTTATTTATA TATTTTAGGCCGCTAAATGAAACGGCTACTAAAACAAAAAGTATTTTTTTATACTAAGGAATAATC CATAAAATAAAAATATAAGTATAAACGGTATATCTTGACGTCCCCTTTTCCTTCAGCTTGGAGGAA ACCAGCATCGGCTAGTCCGAAATTTATGAAAAGTTCTTTTTCACAACACATATGGCTTGTTGCGCC ACTGTCTAAGCACTGCACTCTTTAAGTCGCTGACGTCACCAGCATTTATTATGCTACTTTGGCGGTT CTCTGAGTGTTTTATTACCTTTTTGCCATCTCCGTAACTAAAACTTGAGCACTGTGCTTTTATGTGCC CTTCTTCACCACATTTGAAACGATGAGTTTTCGTTTTTTCTGTTTCGCGAATAAGTTTGGCTTCTGT GCTGCGACGAATGCCTTTGTGCTATCGCTTCGCGCCTATTTTAACTTTATGATGAAAACATTGAAA GAAGGCATCTGATCCCGGGTTTCTATAGACACCACGAAATGTTCAAACTGTTCTGACAGGCTTGA CAGCAACAATAAGGATCGAAGCTCCTC

3. RESULTS

3.1 Tejas Maintain the Organization of the Nuage

3.1.1 Refurbish the dynamic of nuage with fluorescent-conjugated nuage components

In the nuage, several cofactors assisted the proper ping-pong cycle between Aub and Ago3. The previous research found that the Tej is one of the essential factors in the nuage for the piRNA biogenesis in the germ cells.

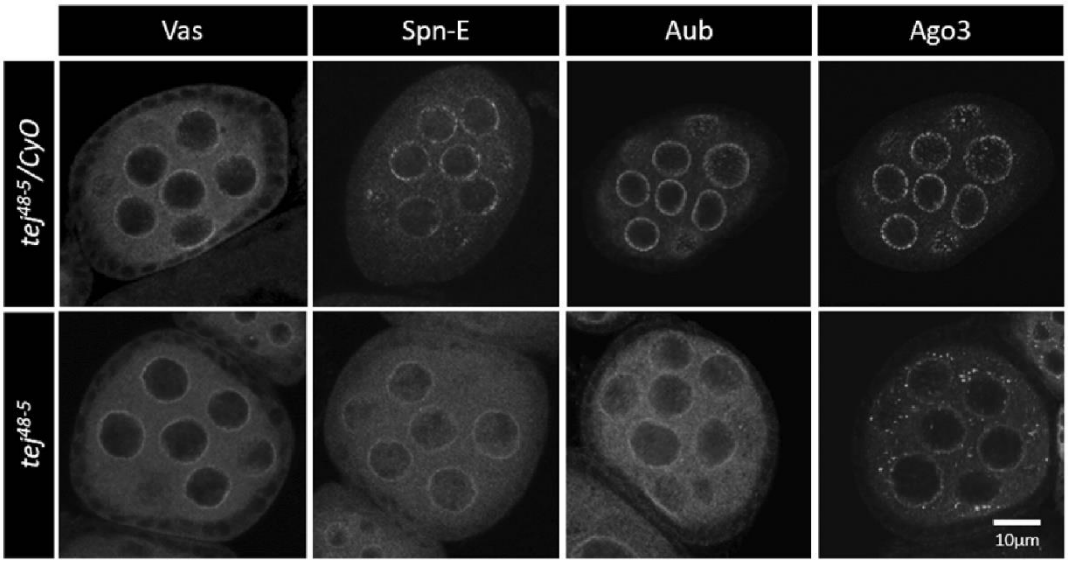
The previous observation about the nuage components with the antibody immunostaining gives information on the behavior of these proteins. However, depending on the specificity and purity of the antibodies, some of the results remain ambiguous (Figure 3.1.1 A). To gain easy-operating tools for precisely observing the nuage components' subcellular dynamics, I utilized the CRISPR- Cas9 genome editing system and generated Knock-In fly lines introducing either GFP- or mKate2- tag on the endogenous protein. The fluorophore-conjugated protein Tej-GFP, Spn-E-mK2, and mK2-Ago3 under the control of endogenous promoters of target genes reflected the practical expression level and behaviors of proteins with high accuracy. All the Knock-In fly strains, including the gifted Vas-GFP, Aub-GFP, and Vas-mCherry lines from other researchers, can be maintained as homozygotes. It indicates minimum impacts on the function of the target genes. These knock-In fly lines allowed us to obtain the images with higher resolution so that it enables us to observe the more meticulous behavior of targeted proteins in *tej* mutant ovaries.

Simultaneous scanning of the two proteins conjugated with distinct fluorophores showed Tej was localized to the prominent nuage granules tightly with Vas and Spn-E. In contrast, Vas forms two distinct structures, partially distributed to the smooth layer surrounding the nuclear, and majorly colocalized to the significant kernels with other nuage factors. Spn-E, Tej, and Ago3 tend to form more concentrated particles, while Vas have a

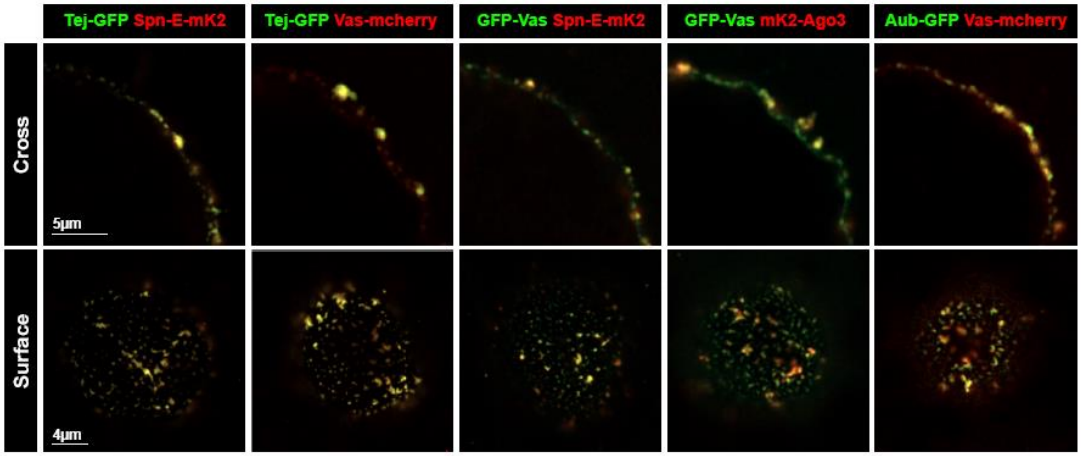
smooth distribution in accordance with Aub (Figure 3.1.1 B). Vas foci is motley distributed around the nuclear surface, relatively separated from the Spn-E, Tej, and Ago3 involving prominent nuage granules (Figure 3.1.1 B).

To further understand the role of Tej in the nuage, I crossed the knock-In fly lines with a loss-of-*tej* fly line to observe the impact of missing in Tej among the nuage components. Consistently with the previous report (Patil and Kai 2010), the nuage granules were disassembled in *tej* mutant. Aub-GFP drastically dispersed to the cytoplasm, but fewer parts of Aub still can be traced as a smooth perinuclear distribution, while mK2-Ago3 is wholly detached from the nuage as foci (Figure 3.1.1 C). Vas-GFP was disappeared from the nuage granule in the *tej* mutant ovaries, but a smooth layer of dotty signals of Vas remains in the perinuclear region (Figure 3.1.1 C), like Aub-GFP. The Spn-E-mK2 are highly granulated and colocalized with other nuage components in the wild-type ovaries (Figure 3.1.1 C). Surprisingly, I observed the endo-promotor controlled Spn-E-mK2 in loss-of-*tej* ovaries lost all their granulation at the perinuclear region of nurse cells, and most of them are localized in the nucleus when Tej is missing (Figure 3.1.1 C). These further precise observations renewed the knowledge we had previously performed with immunostaining. Although the Vas remains fractionally perinuclear localized, the failure of Vas granulation in Tej mutant may lead to the fall of nuage formation. Strikingly, unexpected nuclear localization of Spn-E raises the possibility of its potential role as an intrinsically nuclear protein.

A


Vas
Spn-E
Aub
Ago3
tej48-5/CyO
tej48-5
10µm

B


Tej-GFP Spn-E-mK2
Tej-GFP Vas-mcherry
GFP-Vas Spn-E-mK2
GFP-Vas mK2-Ago3
Aub-GFP Vas-mcherry
Cross
Surface
5µm
4µm

C

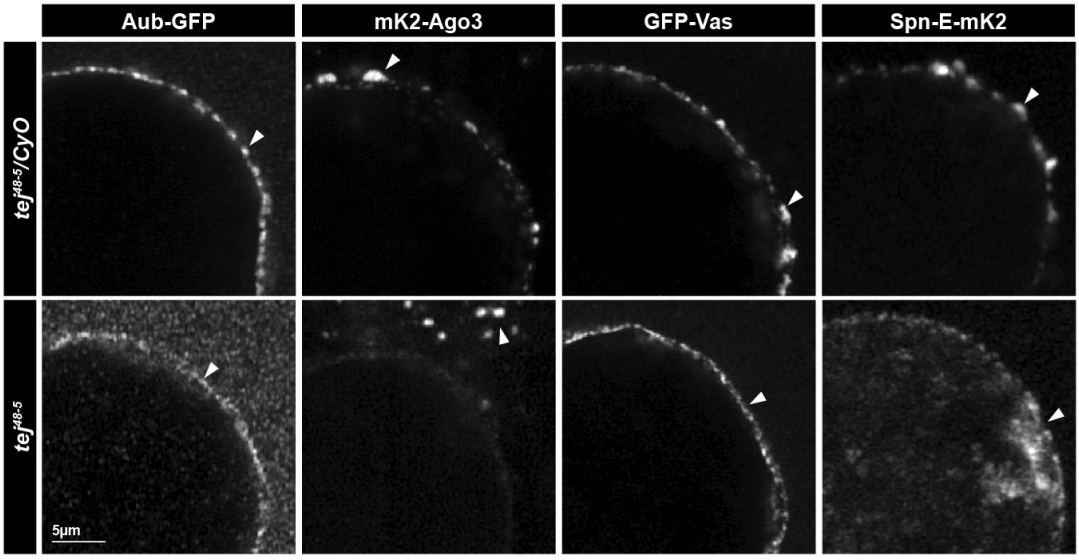

Aub-GFP
mK2-Ago3
GFP-Vas
Spn-E-mK2
tej48-5/CyO
tej48-5
5µm

Figure 3.1.1. Refurbish the dynamic of nuage with fluorophore-conjugated nuage components.

A) Polyclonal antibody immunostaining gives unpolished localization information of nuage components. *tej*$^{48\text{-}5}$*/CyO* and *tej*$^{48\text{-}5}$ / *tej*$^{48\text{-}5}$mutant ovaries were immunostained with polyclonal antibodies against nuage component proteins through the optimized experiment method. All the immunofluorescence signal has been converted to grayscale.

B) The nuage components formed uneven nuage structures. Different fluorophore-tagged endogenous nuage components were presented in the ovary with pairwise combinations. The localization pattern of nuage proteins in the *Drosophila* ovary was recorded by super-resolution imagination. The distribution of proteins was shown in both cross-section (top panel) and surface view (bottom panel).

C) Tej is required for the proper nuage formation. Fluorescent-tagged nuage components were presented in both *tej*$^{48\text{-}5}$*/CyO* and *tej*$^{48\text{-}5}$ / *tej*$^{48\text{-}5}$ mutant ovaries. A general dissemble of the nuage granules was shown (bottom panel). All the immunofluorescence signal has been converted to grayscale.

3.1.2 Tej plays a central role in complex formation with RNA helicases, Vas and Spn-E

Simultaneous scanning of the proteins conjugated with specific fluorophores showed that the Vas-GFP and Spn-E-mK2 have distinct localization and dissembling of the nuage foci when tej is absent. Since Vas and Spn-E proteins remain segregated (Figure 3.1.2 A) at the perinuclear region in Tej mutant, these scattered patterns suggest the potential repulsion between Vas and Spn-E.

To detect the physical interaction between the Tej and two important helicases, Vas and Spn-E *in vivo*, I performed the cross-linking immunoprecipitation, which captured the transient physical interaction among the highly dynamic proteins in ovaries. The immunoprecipitation with fluorophore tagged Tej, Vas and Spn-E successfully co-purified the target multiprotein complexes from the ovary lysate. I can detect several known nuage components in the immuno-precipitated as Tej multiprotein complexes, including both Vas and Spn-E (Figure 3.1.2 B), but not Ago2 protein. This result suggests that Ago2 is probably due to irrelevance to the piRNA biogenesis pathway but involves siRNA biogenesis and RISC assembly (Wei et al. 2012) (Figure 3.1.2 A-C). Interestingly, although Tej and other cofactors were discovered in both Vas multiprotein complexes and Spn-E multiprotein complexes, barely a few amounts of either Vas or Spn-E in the immunoprecipitation with the opponent of proteins can be detected (Figure 3.1.2 C, D). These results suggest a mutual exclusion between Vas and Spn-E with robust reproducibility in the experiment by several biological replicates (Figure 3.1.2 E, F).

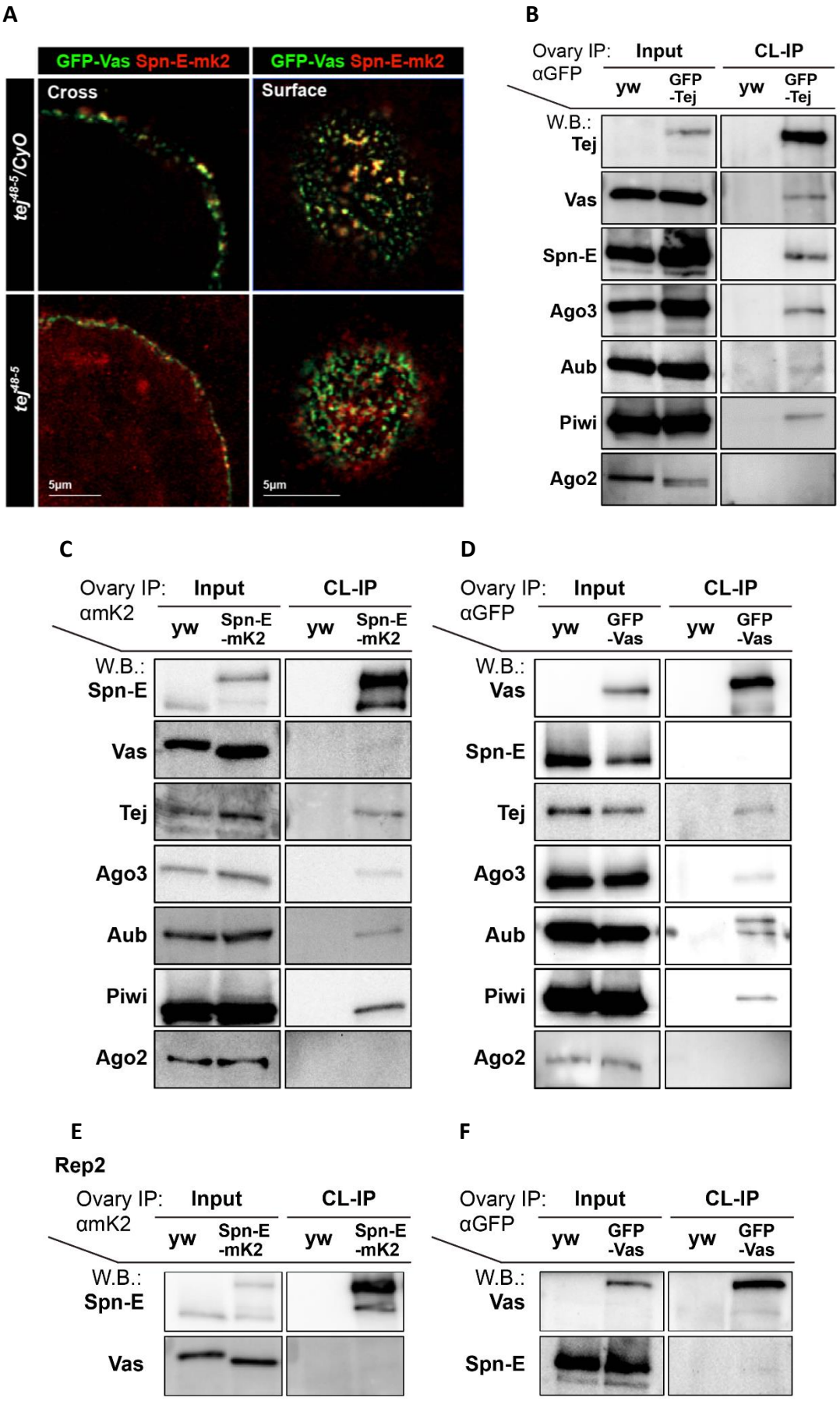

A
GFP-Vas Spn-E-mk2
GFP-Vas Spn-E-mk2
Cross
Surface
tej48-5/CyO
tej48-5
5µm
5µm
B
Ovary IP:
αGFP
Input
CL-IP
yw
GFP -Tej
yw
GFP -Tej
W.B.:
Tej
Vas
Spn-E
Ago3
Aub
Piwi
Ago2
C
Ovary IP:
αmK2
Input
CL-IP
yw
Spn-E -mK2
yw
Spn-E -mK2
W.B.:
Spn-E
Vas
Tej
Ago3
Aub
Piwi
Ago2
D
Ovary IP:
αGFP
Input
CL-IP
yw
GFP -Vas
yw
GFP -Vas
W.B.:
Vas
Spn-E
Tej
Ago3
Aub
Piwi
Ago2
E
Rep2
Ovary IP:
αmK2
Input
CL-IP
yw
Spn-E -mK2
yw
Spn-E -mK2
W.B.:
Spn-E
Vas
F
Ovary IP:
αGFP
Input
CL-IP
yw
GFP -Vas
yw
GFP -Vas
W.B.:
Vas
Spn-E

Figure 3.1.2. Tej plays a central role in complex formation with Vas and Spn-E.

A) Vas and Spn-E detached when Tej in missing. Fluorophore-tagged Vas (green) and Spn-E (red) were simultaneously presented in *tej*$^{48\text{-}5}$/CyO and *tej*$^{48\text{-}5}$/ *tej*$^{48\text{-}5}$mutant ovaries. The nurse cell nuclear's cross-section (left panel) and surface view (right panel) were shown.
B-D) Vas and Spn-E show physical interaction with Tej, respectively. The endogenies Tej-GFP, GFP-Vas, and mK2-Spn-E complexes were fixed and immunoprecipitated from the *Drosophila* ovary lysate. Ovaries from *y w* flies were used as control. Major piRNA biogenesis factors and Ago2 were detected by western blot in each complex. 10% of total extracted complexes were loaded as Input.
E, F) Additional biological replicates of mK2-Spn-E or GFP-Vas cross-linking immuno-precipitation.

3.1.3 Tejas recruit Vas and Spn-E with distinct domains

Results in the immunoprecipitation suggested that the Tej forms distinct complexes with either Vas or Spn-E, and they are localized to each other in a mutually exclusive way. To know how the Tej is involved in the organization of the Vas and Spn-E further, I use the S2 cells that have no expression of major nuage components, Tej, Vas, and Spn-E, as an experiment platform and test the dynamics of nuage components. First, I expressed the GFP-tagged versions of Vas or Spn-E and the mK2-tagged versions of Tej in the S2 cell. Vas forms aggregation, distributing in the cytoplasm (Figure 3.1.3.1 A). It has been reported previously that the N terminal disordered region of Vas contributes to forming aggregation (Nott et al. 2015). Since the Spn-E shows a unique but stable nuclear accumulation in the tej mutant (Figure 3.1.3.1 A), which has never been observed in the wild-type germ cells, it indicates that they are intrinsically nuclear proteins that can be imported to the nuclear for some reason. The full length of Tej protein forms a prominent sturdy cytoplasmic droplet-like structure when expressed in S2 cells. Unlike the perfectly nuclear surrounding nuage granules in the ovary, these granules are distributed in the cytoplasm and never show notable perinuclear localization (Figure 3.1.3.1 A) in S2 cells. Co-expression of Vas and Spn-E with the full length of Tej have colocalized the Tej granules, respectively. Notably, Tej can robustly recruit the nuclear-localized Spn-E to the cytoplasm (Figure 3.1.3.1 A).

Then, to understand how Tej recruit Vas and Spn-E simultaneously, I analyzed the domain architecture of Tej using S2 cells. Two crystal structure characterized domains, Lotus (6-74 aa) and Tudor (377-488 aa), posited at the N- and C-terminal Tej protein, respectively. The 304 amino acid peptide contains a predicted disordered region between these two domains (Figure 3.1.3.1 A, top panel).

Firstly, I made mKate2-tagged truncations in each of the Lotus and Tudor domains of Tej and then co-expressed them with GFP-tagged Vas or Spn-E in the S2 cells (Figure 3.1.3.1 A). The deletion of the lotus domain of Tej(Tej ΔLotus)forms porous cytoplasmic condensates (Figure 3.1.3.1 A). Consistent with the previous report that Lotus domain is engaged in the interaction of Vas the, these granules of Tej ΔLotus failed to gather the Vas aggregates. Instead, it showed the ability to completely exclude Vas from the condensates (Figure 3.1.3.1 A). Unlike Vas, it was observed that a firm colocalization of Spn-E and Tej ΔLotus in the condensates (Figure 3.1.3.1 A). Remarkably different from the Tej ΔLotus, the deletion of Tudor of Tej (Tej ΔTudor) lost granulation ability and was dispersed in the cytoplasm (Figure 3.1.3.1 A). Moreover, it preserved the full ability of Vas binding and still vigorously enforced the nuclear-localized Spn-E to the cytoplasm (Figure 3.1.3.1 A). It has been reported that Spn-E interacts with Tej through the C terminal (Patil and Kai 2010) of Tej.

Next, to refine the structure-based function of Tej about Spn-E precisely, I made truncated versions of Tej stepwise from the N terminus of Tej (Figure 3.1.3.1 B, top). I cotransfected them with Spn-E in the S2 cells and found that loss of 295-362 aa fragments deprived the ability of Tej to recruit the Spn-E (Figure 3.1.3.1 B, bottom). As the middle part of Tej between the Lotus and Tudor domain retains the ability to tether the Spn-E to the cytoplasm (Figure 3.1.3.2 A, right), it is speculated that this region contains the putative motif that recruits Spn-E and localizes in the cytoplasm. I further truncated the middle part of Tej (101-349 aa) to refine the tethering region of Spn-E. The series of truncations in the middle part of Tej is co-expressed with SpnE in S2 cells, and it was found that deletion of 338-349 aa fragment in Tej lost the recruitment of Spn-E significantly (Figure 3.1.3.2 A, right). By aligning the amino acid sequence of the Tej homologs among other species, including the several kinds

of *Drosophila* and vertebrates, I identified a highly conserved motif located adjacent to the N terminus of the Tudor domain.

Further step-wise truncation of Tej and co-transfection with SpnE in S2 cell revealed that 341-348 aa region of tej is critical (Figure 3.1.3.2 B) for the Spn-E recruitment as a minimum motif and named it 'Spn-E Recruit Site', SRS in brief (Tej ΔSRS). Thus, unlike GFP-Vas, most GFP-Spn-E remains in the nucleus in the cotransfection with Tej ΔSRS (Figure 3.1.3.1 A). The particles formed by mK2-tagged Tej ΔSRS are lesser and smaller than the Tej-FL formed particles, despite the S2 cells being transfected with the same amount of plasmid, indicating the depletion of SRS motif in Tej may impair Tej's stability (Figure 3.1.3.1 A). Finally, for further analysis about this motif, I tested whether the point mutation in highly conserved amino acids would cancel the Spn-E recruitment. Highly conserved Proline, Aspartic or Arginine in the SRS motif of Tej was substituted with alanine and expressed in the S2 cells together with the GFP-tagged Spn-E. Both P343A and R348A show a weak cessation of nuclear-localized Spn-E, but not in D345A, suggesting that the number 343 Proline and 348 Arginine residues in Tej potentially act as a critical point for Spn-E recruitment (Figure 3.1.3.2 C).

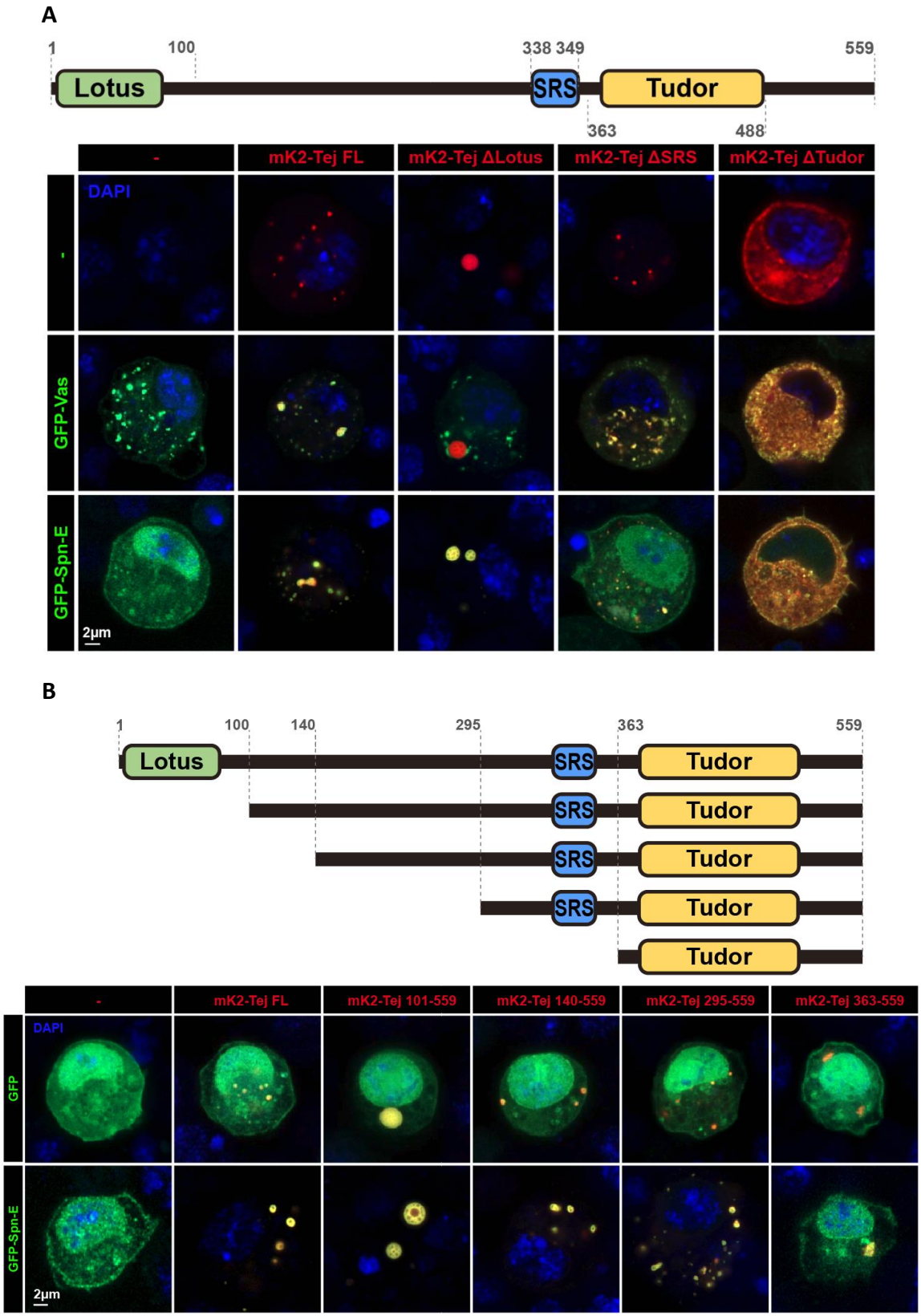
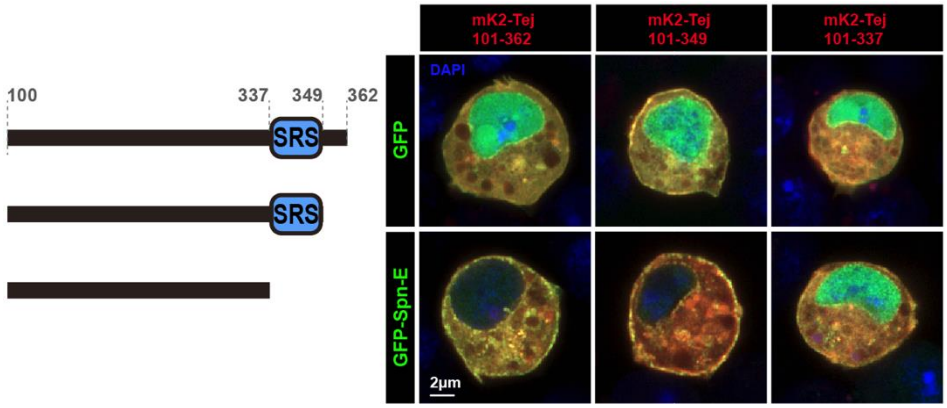

A
1
100
338 349
559
Lotus
SRS
Tudor
363
488
-
mK2-Tej FL
mK2-Tej ΔLotus
mK2-Tej ΔSRS
mK2-Tej ΔTudor
DAPI
-
GFP-Vas
GFP-Spn-E
2μm
B
1
100
140
295
363
559
Lotus
SRS
Tudor
-
mK2-Tej FL
mK2-Tej 101-559
mK2-Tej 140-559
mK2-Tej 295-559
mK2-Tej 363-559
DAPI
GFP
GFP-Spn-E
2μm

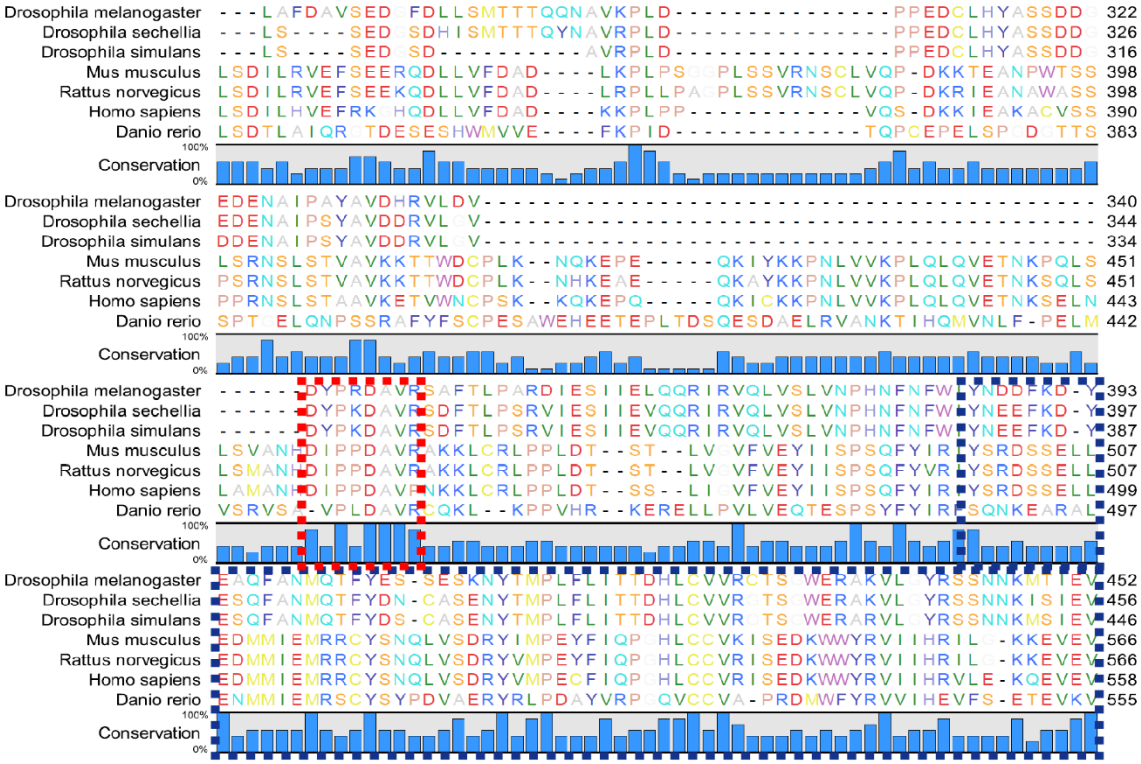
Figure 3.1.3.1. Tejas recruits Vas and Spn-E to nuage via distinct domains.

A) Recruitment of Vas and Spn-E by distinct domain or motif of Tej in S2 cells. The GFP-tagged Vas or Spn-E (green) was co-expressed in S2 cells along with the mKate2-tagged Tej FL or Tej truncated variants (red): Tej ΔLotus (delete Lotus domain), Tej ΔSRS (delete SpnE recruitment motif), and ΔTudor (delete Tudor domain). A single expression of all the fluorescent-tagged proteins was shown in the top and left panels, DNA is stained with DAPI (blue).
B) Particular region of the Tej is necessary for recruiting the Spn-E *in vivo*. The GFP-tagged Spn-E (green, bottom panel) was co-expressed in S2 cells along with the mKate2-tagged Tej FL or Tej stepwise truncated variants (red, bottom panel). GFP protein was expressed as control along with the same Tej stepwise truncated variants (green, middle panel). Schematic representation of the Tej truncation on above. DNA is stained with DAPI (blue).

A


mK2-Tej 101-362
mK2-Tej 101-349
mK2-Tej 101-337
100
337
349
362
SRS
SRS
DAPI
GFP
GFP-Spn-E
2μm

B


Drosophila melanogaster
Drosophila sechellia
Drosophila simulans
Mus musculus
Rattus norvegicus
Homo sapiens
Danio rerio
Conservation
100%
0%

C

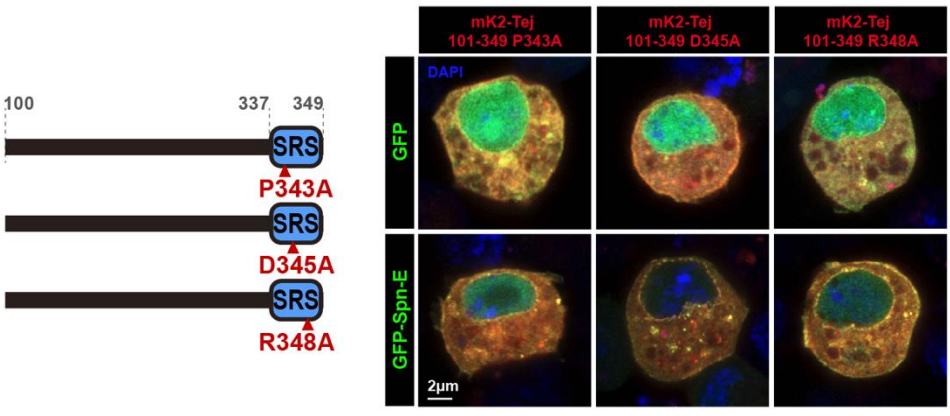

mK2-Tej 101-349 P343A
mK2-Tej 101-349 D345A
mK2-Tej 101-349 R348A
100
337
349
SRS
P343A
SRS
D345A
SRS
R348A
DAPI
GFP
GFP-Spn-E
2μm

Figure 3.1.3.2. Highly conserved core amino acid residues are essential to Spn-E recruitment.

A) A highly conserved SRS motif is essential for the capacity of Tej to recruit Spn-E. The GFP-tagged Spn-E (green) was co-expressed in S2 cells along with the mKate2-tagged Tej middle part variants (red): 101-362 aa, 101-349aa, and 101-337aa. GFP protein was expressed as control along with the same Tej stepwise truncated variants (green, middle panel). Schematic representation of the Tej truncation settled on the left. DNA is stained with DAPI (blue).
B) Sequence conservation in a multiple sequence alignment of Tej homologs. The amino acid sequence of the Tej homologs from *Drosophila* species and vertebrates were aligned. The residues were colored depending on the conservation. Blue bars show the conservation of amino acid residues between the aligned homologs. The red frame cycled the SRS motif, and the blue frame cycled part of the identified Tudor domain.
C) Same as shown in A), GFP-tagged Spn-E (green) was co-expressed in S2 cells along with the mKate2-tagged Tej middle part variants containing single substituted residue P343A, D345A, and R348A (red).

3.1.4 Tejas aggregates Vas and Spn-E and segregates them into unique phases *in vitro*

Analysis with truncation mutant of Tej revealed Tejas recruits Vas and Spn-E through the distinct domains or specific motifs. To answer why I detected the mutually exclusive binding of Vas and Spn-E to the Tej *in vivo* (Figure 3.1.2 A, C, D). I expressed GFP-tagged Vas, mKate2 tagged Spn-E, and Myc-tagged Tej to monitor their subcellular localization.

Consistent with the behavior described above, Tej aggregated with the Vas and Spn-E and formed large cytoplasmic granules. Surprisingly, the unique structure can be stably observed as Spn-E was concentrated to the central part surrounded by Vas and Tej is localized the entire granule with encircling signal of Vas (Figure 3.1.4). When replaced the Myc-Tej full length with the Vas interaction deficient variant Tej ΔLotus, the Tej-Spn-E particles were no longer surrounded by Vas, and Vas was detached from the particles and formed the distinct cytoplasmic aggregates (Figure 3.1.4).

These peculiar behavior patterns of Tej, Vas, and Spn-E in S2 cells raise the possibility of the Tej acting as a central role of aggregation to gather the two helicases and also segregate them into separate phases. This can be a piece of evidence supporting that Vas and Spn-E were segregated in distinct subcomplex in the nuage, and Tej is an essential factor that sustained their restricted contact.

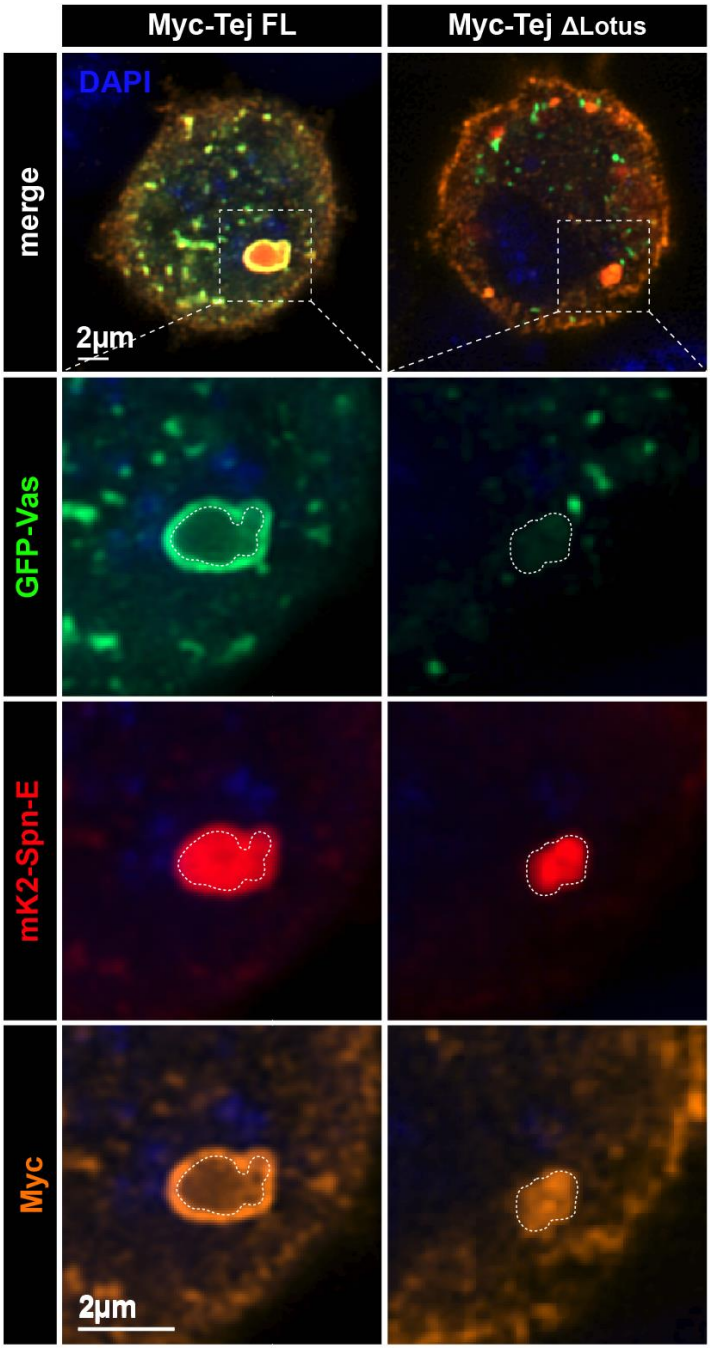


Figure 3.1.4. Tejas aggregates meanwhile isolated Vas and Spn-E into unique phases *in vitro.*

Tej recruits and segregates Vas and Spn-E in S2 cells. Myc-tagged full-length Tej, GFP-tagged Vas (green), and mKate2-tagged Spn-E (red) were expressed in S2 cells. Tej was stained with a Myc antibody (orange). DNA is stained with DAPI (blue).

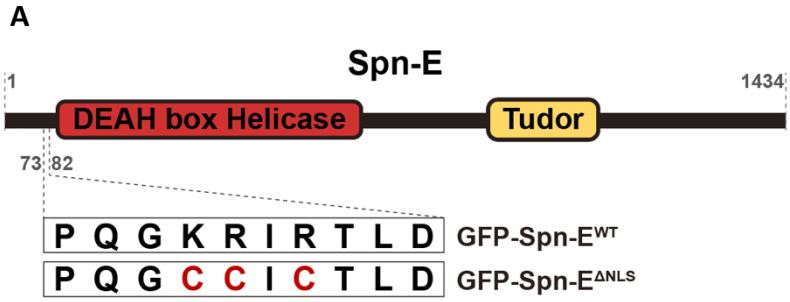
3.2 Subcellular localization of Spn-E is regulated by its NLS motif

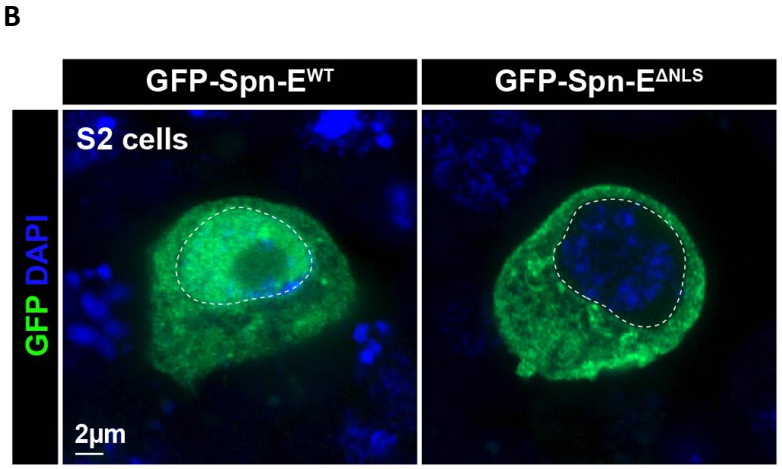
3.2.1 Spn-E is an intrinsically nuclear protein controlled by NLS

In vitro experiments in S2 cells showed the nuclear-accumulated Spn-E observed in the *tej* mutant ovaries. I further checked the behavior of Spn-E in the *Drosophila* ovaries. By utilizing the Spn-E-mK2 Knock-In flies, I observed the nuclear localization of Spn-E in the ovarian somatic cells and the nuage localization in the nurse cells (Figure 3.2.1 C). Using a mapping tool for detecting the nuclear localization signal, I identified a putative novel class II monopartite NLS in the N terminal of the Spn-E (73-82 aa) (Figure 3.2.1 A).

To verify whether the Spn-E nuclear localization is dependent on this predicted NLS, I built plasmids that encode the GFP-tagged Spn-E full length (Spn-E FL) and the NLS disrupted mutant in Spn-E (Spn-EΔNLS), which contains amino acid substitution, K76C, R77C, and R79C in the core region of NLS (Figure 3.2.1 A). Expression of Spn-E ΔNLS in the S2 cell impeded to import of Spn-E to the nucleus though the Spn-E FL was localized dominantly in the nucleus and cytoplasm(Figure 3.2.1 B).

To investigate further how the NLS controls the subcellular localization of Spn-E *in vivo*, I generated the transgenic fly lines expressing the GFP-tagged full length (Spn-E FL) and NLS disrupted variant of Spn-E (Spn-EΔNLS). Then I observed the dynamics of either GFP-tagged Spn-E FL or Spn-E ΔNLS driven by the *tj*-Gal4 promotor in the ovarian somatic cells and the *nos*-Gal4 in the nurse cells. The localization of Spn-E WT inside the somatic nucleus was changed when the NLS of Spn-E was mutated and the Spn-E ΔNLS was diffused in the cytoplasm uniformly (Figure 3.2.1 D). On the other hand, perinuclear localization of Spn-E in the nurse cell was not affected due to invalid NLS in Spn-E (Figure 3.2.1 D). Collectively, these results suggested that the NLS on the N terminus of Spn-E was defined as an intrinsically nuclear localization of Spn-E.


A
Spn-E
1
1434
DEAH box Helicase
Tudor
73 82
P Q G K R I R T L D
GFP-Spn-E^WT
P Q G C C I C T L D
GFP-Spn-E^ΔNLS


B
GFP-Spn-E^WT
GFP-Spn-E^ΔNLS
S2 cells
GFP DAPI
2μm

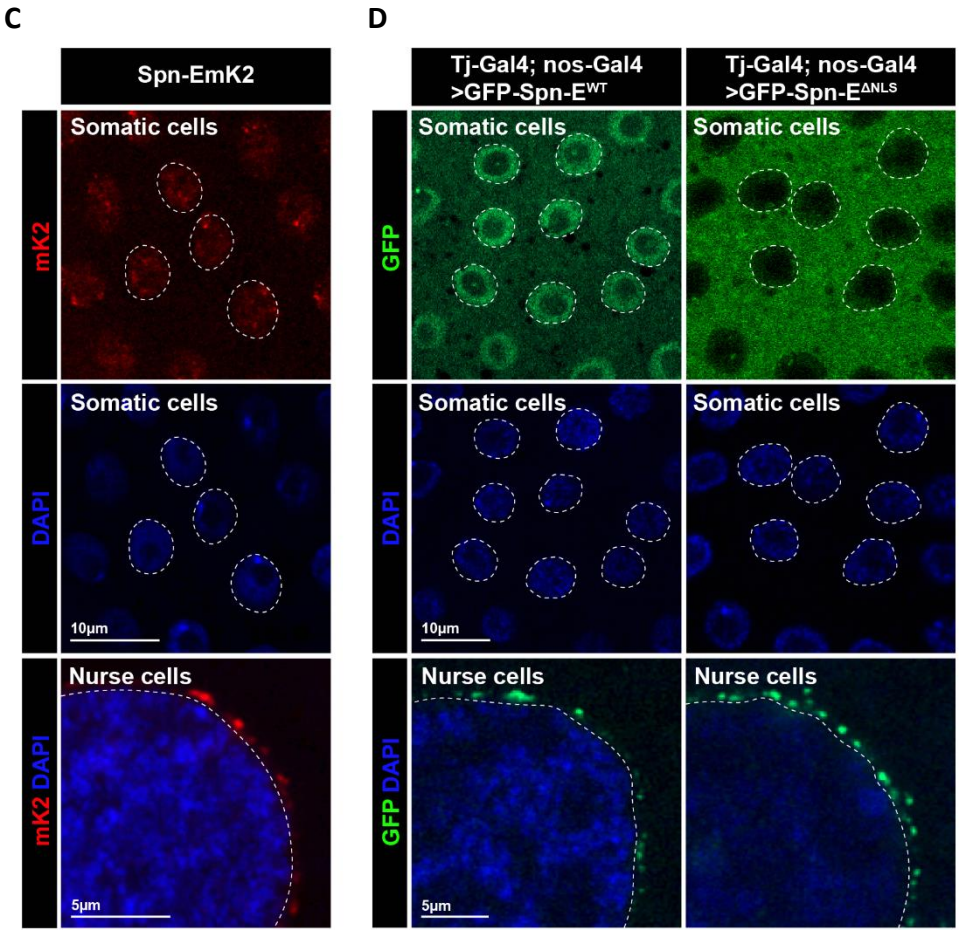

C
Spn-EmK2
Somatic cells
mK2
Somatic cells
DAPI
10μm
Nurse cells
mK2 DAPI
5μm
D
Tj-Gal4; nos-Gal4 >GFP-Spn-E^WT
Tj-Gal4; nos-Gal4 >GFP-Spn-E^ΔNLS
Somatic cells
Somatic cells
GFP
Somatic cells
Somatic cells
DAPI
10μm
Nurse cells
Nurse cells
GFP DAPI
5μm

Figure 3.2.1. Spn-E is intrinsically a nuclear protein controlled by NLS.

A) Creation of the plasmids and transgenic fly line expressed the nuclear localization signal deficient GFP-tagged Spn-E. Schematic representation of the residue substitution in the putative NLS of Spn-E.

B) The mutated NLS impedes the nuclear import of Spn-E *in vitro*. The GFP-tagged Spn-E full length (Spn-E FL) and the Spn-E ΔNLS, which contains substitution mutations K76C, R77C, and R79C in the core amino acids of its NLS, were expressed in the S2 cells (green). DNA is stained with DAPI (blue), dotted line cycled the nuclear.

C) Spn-E has distinct localization in ovarian somatic cells and nurse cells. mKate2-tagged Spn-E (red) were presented in wildtype *Drosophila* ovaries. DNA is stained with DAPI (blue), dotted line cycled the nuclear.

D) NLS responsible for the nuclear localization of Spn-E in ovarian somatic cells. GFP-tagged Spn-E WT and Spn-E ΔNLS were expressed respectively in the somatic cell and nurse cells of *Drosophila* ovaries under the control of promotor UASp, driven by the driver *tj*-Gal4 and *nos*-Gal4. DNA is stained with DAPI (blue), dotted line cycled the nuclear.

3.2.2 Nuclear-localized Spn-E does not associate with the piRNA precursors

Combining the facts that Tej- dependent cytoplasmic prenuclear localization of Spn-E and NLS-dependent nuclear localization in the absence of Tej, the Spn-E was thought to be regulated by distinctive localizations. I wanted to understand dynamics of Spn-E are associated with its function.

Firstly, I quantified the nuclear fraction of Spn-E in control and *tej* mutant ovaries. Direct observation of the fluoresce signal of Spn-E-mK2 shows a significant proportion of the Spn-E was smeared inside the nuclei in tej mutant compared to the wild type(Figure 3.2.2 A, B). Upon the fact that the Spn-E was localized in the nucleus of nurse cells in the tej mutant, it is hypothesized that Spn-E is engaged in the transcription or translocation of piRNA precursor inside the nuclear. I used Fibrillarin, a component of several ribonucleoproteins that include a nucleolar small nuclear ribonucleoprotein (SnRNP) and small nucleolar ribonucleoproteins (snoRNPs), and DAPI as makers of nucleole and DNA respectively and stained with mK2-tagged Spn-E in the *tej* mutant ovaries. The Spn-E signal overlapped neither fibrillarin nor DAPI (Figure 3.2.2 C). Since Spn-E is a DExH box helicase that can bind RNA and switches its localization between the nuclear and perinuclear region, I questioned whether Spn-E captures and transports the piRNA precursors from the transcription site to the perinuclear nuage. I traced the piRNA precursors derived from the piRNA clusters *38C* and *42AB* by hairpin chain reaction fluorescence *in situ* hybridization (HCR-FISH) and observed the location of endogenous mK2-tagged Spn-E in the loss-of-*tej* ovaries. The nuclear localization of Spn-E induced by the absence of Tej was segregated from the HCR-FISH-stained precursors (Figure 3.2.2 D). Taken together, these results suggested that these ectopically nuclear-accumulated Spn-E have fewer possibilities to approach the DNAs or nascent RNAs and involve in the transcription or translocation of piRNA precursors.

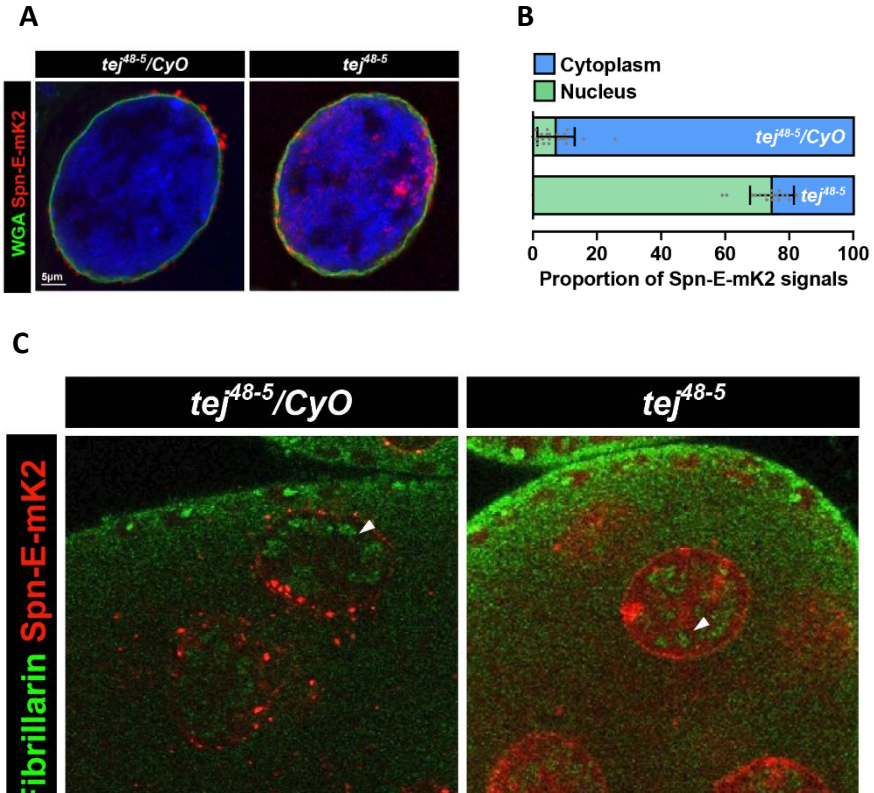

A
tej48-5/CyO
tej48-5
WGA Spn-E-mK2
5μm
B
Cytoplasm
Nucleus
tej48-5/CyO
tej48-5
0
20
40
60
80
100
Proportion of Spn-E-mK2 signals
C
tej48-5/CyO
tej48-5
Fibrillarin Spn-E-mK2
10μm
D
tej48-5/CyO
tej48-5
38C Spn-E-mK2
42AB Spn-E-mK2
5μm

Figure 3.2.2. Nuclear-localized Spn-E does not cooperate with the piRNA precursors.

A) Spn-E is massively localized to the nuclear in *tej* mutant ovaries. mKate2-tagged Spn-E (red) were presented in wildtype *Drosophila* ovaries. DNA is stained with DAPI (blue), the membrane is stained with (green) Alexa Fluor 488 Conjugate WGA.

B) The proportion of nuclear and cytoplasmic fraction was quantified for $tej^{48\text{-}5}/CyO$ (n = 20) and *tej* mutant (n = 20) nurse cell nuclei.

C) Nuclear localized Spn-E does not overlap with nucleole. Nucleole marker fibrillarin is stained by antibody immunostaining (green) and shown with the endogenous mKate2-tagged Spn-E (red). The white arrow marked the fibrillarin signals.

D) Nuclear localized Spn-E does not overlap with *clusters 38C,* and *42AB* derived piRNA precursors. The distribution endogenous mKate2-tagged Spn-E (red) and HCR-FISH stained piRNA precursors (green) was scanned in $tej^{48\text{-}5}/CyO$ and *tej* mutant nurse cell nuclear.

3.3 Tej functions processing of piRNA precursor via recruitment of Vas and Spn-E to nuage

3.3.1 Tejas is necessary for the ordinary loading of piRNA to Aub and Ago3.

As localization of Aub and Ago3 show has been severely affected by the loss of Tej (Figure 3.1.1 C), I addressed whether Aub and Ago3, the PIWI family members, are the two core components in ping-pong cycles, have the feasibility of loading piRNAs in the tej mutant ovaries. Small RNAs extracted from the immunoprecipitated with either Aub or mKate2-Ago3 complexes were visualized by radioisotope labeling. With the equal amount of proteins in each condition was verified by western blot analysis (Figure 3.3.1 A, B), a massive depletion of Aub and mKate2-Ago3 bound piRNA was shown by autoradiography (Figure 3.1.1 A, B). Then, I performed small RNA deep sequencing for those Aub and Ago3-bound piRNAs in the condition of either control or *tej* mutant ovaries. Aub and Ago3 binding piRNAs were plotted over the germline-specific dual-stranded piRNA clusters *38C* and *42AB*. The number of reads from each library was normalized dependent on the number of reads of piRNA irrelevant small RNA reads: rRNA, snRNA, snoRNA, tRNA, and miRNA. Both sence and antisence piRNA populations mapped to the Clusters *38C* and *42AB* are significantly downregulated when Tej is missing (Figure 3.3.1).

Taken together, these data indicate that Tej is required for ordinary loading of piRNA to Aub and Ago3, which sustains the ping-pong amplification cycle, which digests the piRNA precursors into the mature piRNAs.

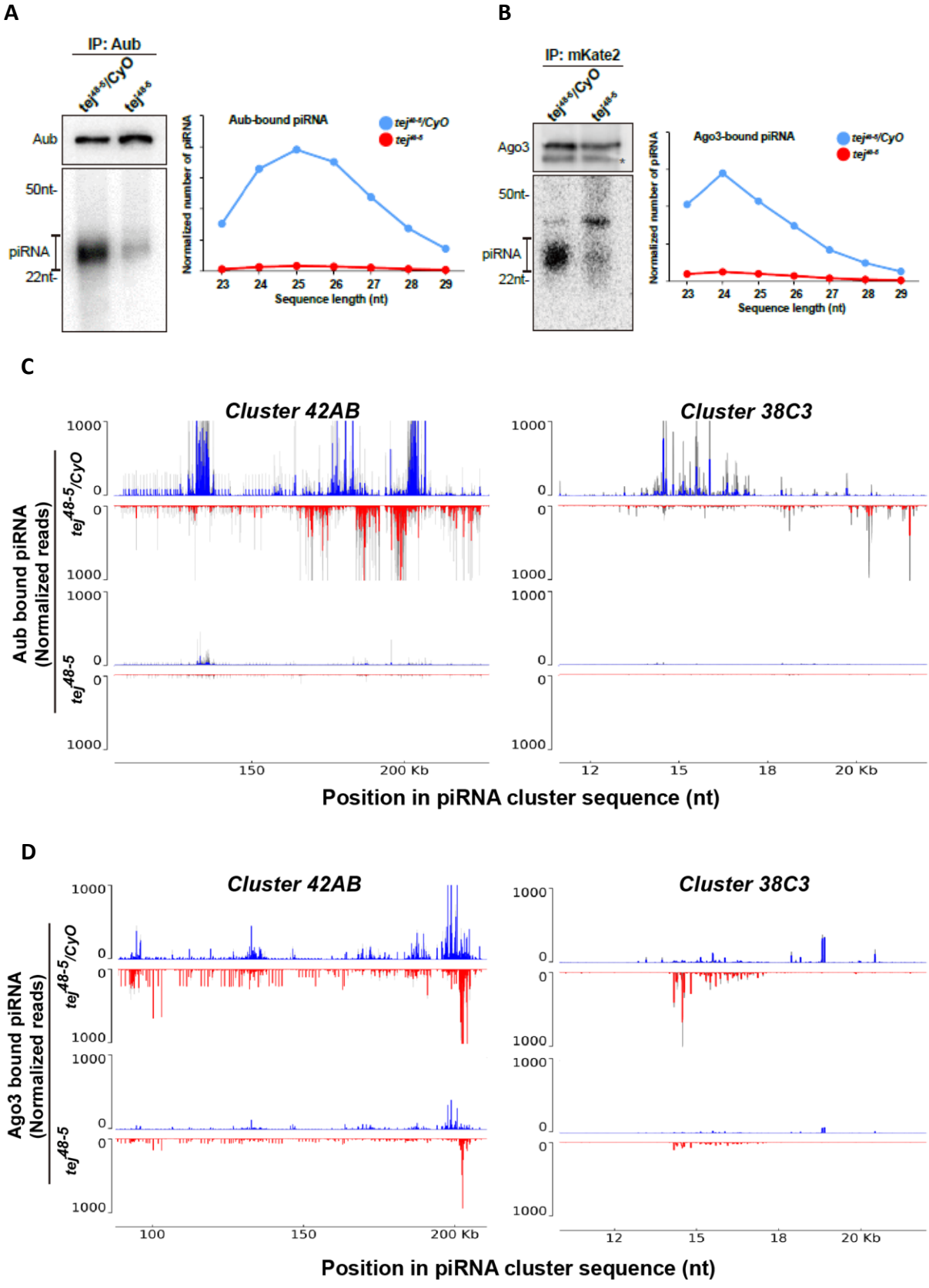

A
IP: Aub
tej48-5/CyO
tej48-5
Aub
50nt-
piRNA
22nt-
Aub-bound piRNA
tej48-5/CyO
tej48-5
Normalized number of piRNA
23 24 25 26 27 28 29
Sequence length (nt)
B
IP: mKate2
tej48-5/CyO
tej48-5
Ago3
50nt-
piRNA
22nt-
Ago3-bound piRNA
tej48-5/CyO
tej48-5
Normalized number of piRNA
23 24 25 26 27 28 29
Sequence length (nt)
C
Cluster 42AB
Cluster 38C3
Aub bound piRNA
(Normalized reads)
tej48-5/CyO
tej48-5
1000
0
150
200 Kb
12
15
18
20 Kb
Position in piRNA cluster sequence (nt)
D
Cluster 42AB
Cluster 38C3
Ago3 bound piRNA
(Normalized reads)
tej48-5/CyO
tej48-5
1000
0
100
150
200 Kb
12
15
18
20 Kb
Position in piRNA cluster sequence (nt)

Figure 3.3.1. Tejas is Necessary for the Ordinary Loading of piRNA to Aub and Ago3

A, B) piRNAs are massively depleted from the Aub and Aog3 complexes. The piRNAs extracted from the immunoprecipitated PIWI protein-complexes between the control and *tej* mutant are visualized by radioisotope labeling. The amount of the PIWI proteins in elution was detected and equalized by western blot. Line plots show the abundance of Aub and Ago3 bound piRNA in control (blue) and *tej* mutant (red) ovaries. The reads number was normalized to the read numbers of small RNAs that exclude piRNA between all libraries

C, D) PIWI proteins binding piRNAs mapping against the piRNA cluster are significantly downregulated when Tej is missing. Aub- or Ago3-bound piRNAs mapping against the germline cell-specific clusters *38C* and *42AB* are decreased drastically. Uniquely mapping piRNAs are plotted over the germline-specific dual-stranded piRNA clusters *42AB* and *38C3*. Sense (blue) and antisense (red) piRNAs are indicated with upward and downward peaks, respectively. The displayed region of each cluster was marked at the bottom. Libraries were normalized to the read numbers of small RNAs that exclude piRNA between all libraries.

3.3.2 Absence of Tejas leads to the accumulation of piRNA precursor

It has been reported that Aub and Ago3 alternately cleave the long precursors into mature piRNAs in the ping-pong pathway. The massive depletion of Aub- and Ago3-bound piRNA in tej mutant ovaries indicates that the upstream mechanisms of the ping-pong cycle failed to process the precursors properly. To understand how the loss of Tej causes the malfunction of the precursor processing in germline cells, I quantified the RNA levels of precursor in *tej* mutant ovaries by RT-qPCR.

The piRNA precursors derived from *cluster 38C* and *cluster 42AB* that are the predominant piRNA precursor in germline cells (Brennecke et al. 2007), significantly accumulated in the *tej* mutant when compared with heterozygotes of Tej (Figure 3.3.2 A). In contrast, the precursor expression of ovarian somatic cell-specific cluster *flamenco* was not drastically affected. To know how this precursor accumulation is associated with the piRNA biogenesis, I also tested the levels of piRNA precursors in *vas*, *spn-E*, *krimp* and nxf3 mutants. Except for the *nxf3*, all tested loss-of-function mutants in the *vas*, *spn-E,* and *krimp* ovaries had an accumulation of *cluster 38C* and *cluster 42AB* derived piRNA precursors (Figure 3.3.2 A).

This accumulation of precursors delivered obvious evidence that precursors were not correctly processed when Tej was missing. Ordinary, the precursors are transported to the perinuclear region by the Nxf3-Nft1 pathway after transcription (ElMaghraby et al. 2019; Kneuss et al. 2019). To identify further the defective step during processing of precursors, I visualized the transcripts of precursors with Hairpin Chain Reaction *in situ* Hybridization (HCR-FISH), a high-resolution RNA-FISH that can detect the small amount of the piRNA precursors by an amplification method. Firstly, I designed probe sets targeted explicitly in the unique regions of *cluster 38C* and *cluster 42AB*. Probe pairs were placed in the junction sites of two

distinct transposons harbored in the cluster loci to avoid off-target detection of transposon transcripts (Table 4). To define the precise location of the accumulating precursors inside or outside the nuclear membrane, I used the Alexa Fluor 488 Conjugate Wheat Germ Agglutinin (WGA) to stain the nuclear membrane of the nurse cells (Figure 3.3.2 B). As a result, transcripts derived from *cluster 38C* and *cluster 42AB* have accumulated both the nucleus and perinuclear region, more concentrated around the nuclear envelope, in *tej* mutant ovaries (Figure 3.3.2 B). It was significantly different from the wild-type germline cells where the precursor transcripts are appropriately processed and only barely located in the perinuclear region, reflecting a transitory existence before processing (Figure 3.3.2 B, left panel). To understand the behavior of piRNA precursors in other mutants, I further applied the HCR-FISH in the *spn-E*, *vas, krimp* and *nxf3* mutants, respectively. A similar accumulating pattern about the piRNA precursors derived from *cluster 38C* and *cluster 42AB* can be observed in loss-of-*vas* and loss-of-*spn-E* ovaries (Figure 3.3.2 B), but not the *nxf3* mutant ovaries, which consist with the qPCR results above and the previous study (Kneuss et al. 2019; ElMaghraby et al. 2019) (Figure 3.3.2 B). Although I can detect milder increased precursors in the *krimp* mutated ovaries by qPCR quantification, the HCR-FISH shows transcripts of that the precursors are generally enriched inside the nuclei but not close to the cell membrane (Figure 3.3.2 B). These comparable accumulations of precursors in the Tej, Vas, or Spn-E insufficient ovaries indicated that Tej might involve in the piRNA precursor processing in coordination with two helicases Vas and Spn-E.

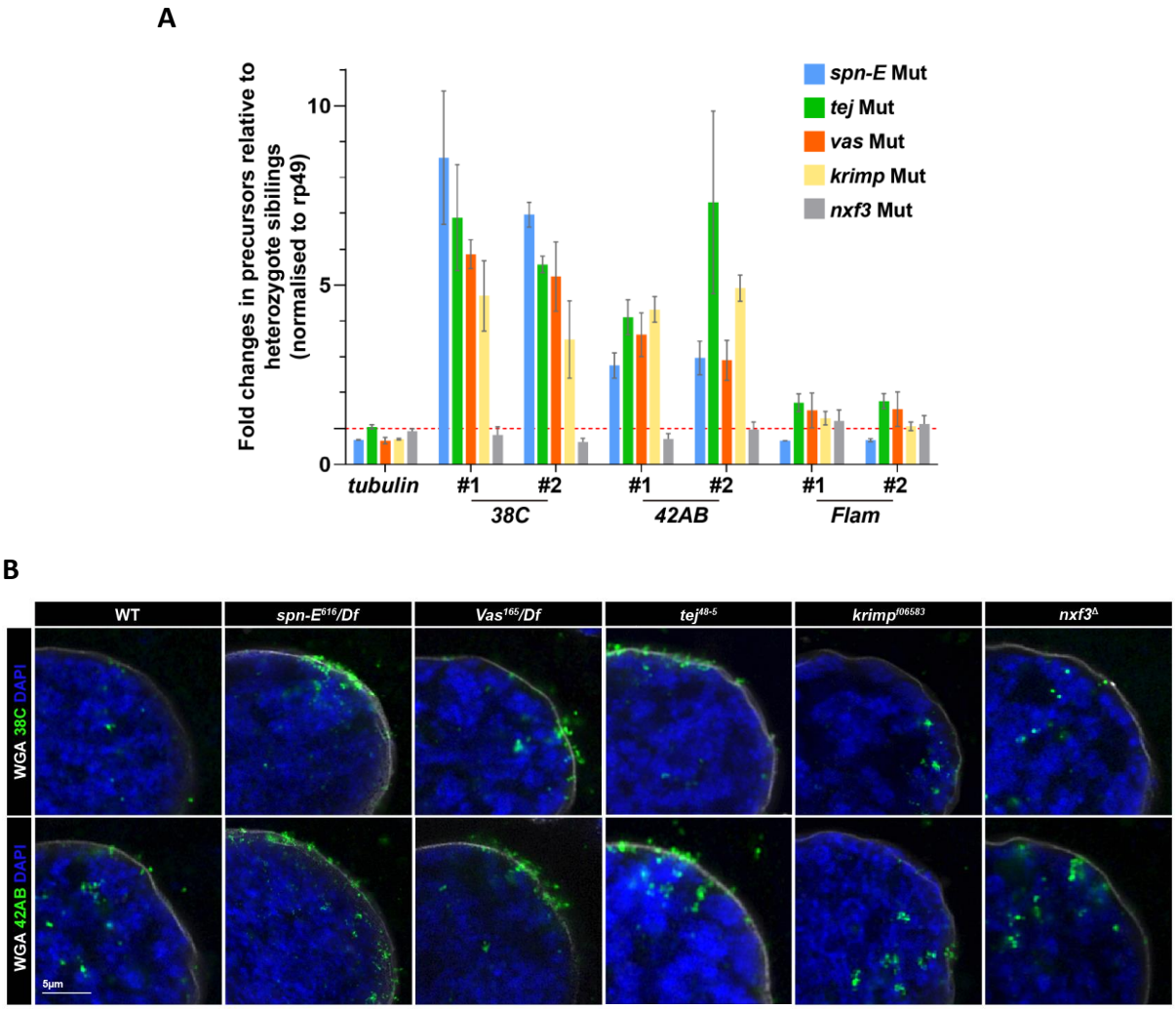


Figure 3.3.2. The absence of Tejas leads to the accumulation of piRNA precursors.

A) Absence of Tejas leads to the accumulation of piRNA precursors. Bar graphs showing fold changes in RNA levels of *tubulin*, piRNA precursors derived from *cluster 42AB, 38C*, and *flam* in the total RNA extracted from the *Drosophila* ovaries with indicated genotypes. Error bars indicate standard deviation, n=3.

B) piRNA precursors are accumulated with different distributions in nurse cells. piRNA precursors accumulated in piRNA biogenesis associated proteins deficient *Drosophila* ovaries. *cluster 42AB* and *38C* transcripts are shown by HCR-FISH (green) in wild-type and mutant ovaries of indicated genotypes. The nuclear envelope is marked by Alexa Fluor 488 conjugated WGA (white), DNA is stained with DAPI (blue).

3.3.3 Tejas controls the nuage localization of Vasa and Spn-E *in vivo*

The previous study shows that the absence of Tej leads to massively disassembles of the nuage granules, which leads to difficulties for studying how Tej is involved in the precursor processing. As I have identified the domains of Tej that recruit Vas or Spn-E, I defined the Tej-dependent spatial localization for Vas or Spn-E individually.

I generated transgenic fly lines expressing the N terminal mini-Turbo and GFP-fused Tej full length (Tej-FL) and Tej truncation variants, Tej ΔLotus (deleted LOTUS domain), Tej ΔSRS (deleted SpnE recruitment domain), and Tej ΔTudor (deleted TUDOR domain). Tej-FL and all the transgenic Tej-truncated variants were expressed by the UAS-Gal4 system, driven by *NGT40*-Gal4 and *nos*-Gal4-VP16 for the expression in the germline cells, in the t*ej* mutant background. When either Tej FL, Tej ΔLotus, or Tej ΔSRS in *tej* mutant were expressed, they are formed similar perinuclear foci, while the majority of the Tej ΔTudor distributed to the cytoplasm uniformly and only a few of the Tej ΔTudor forms a smooth layer surrounding the nucleus (Figure 3.3.3 B, top panel). The subcellular localization of these truncated Tej variants is similar to what was observed in the S2 cells above, and the localization of Tej FL, Tej ΔLotus, and Tej ΔTudor are also partially consistent with the previous study about the Tej (Patil and Kai 2010). Notably, the faint signal of Tej ΔTudor that remains in the nuclear envelope was newly unearthed. For knowing the details of the nuage formation in the expression of each truncation in *tej* mutant, I co-stained with Vas and Spn-E and observed their localization by antibody staining. Tej-FL successfully recruited both Vas and Spn-E on perinuclear nuage granules, which suggests it completely rescued the formation of the nuage granules (Figure 3.3.3 B). Tej ΔLotus significantly eliminated Vas from the nuage granules, while the Spn-E nuage localization remains unaffected (Figure 3.3.3 B).

In contrast, the Tej ΔSRS rescue of Tej mutant ovaries experiment shows lesser attraction to the Spn-E in the missing of the interactive region. Spn-E was still partially tethered to the nuage granules even in the absence of the SRS motif of Tej, although other fractions of Spn-E were observed in the nucleus or cytoplasm (Figure 3.3.3 B, top panel). Unexpectedly, the Vas recruitment was also affected when the 'SRS' motif of Tej was missing (Figure 3.3.3 B, top panel). Surprisingly, Tej ΔTudor lost its granulation and cannot form the proper nuage foci, overlapping weakly with the Vas and Spn-E in the perinuclear region (Figure 3.3.3 B, top panel).

Despite the missing of the SRS motif in Tej, the localization of Spn-E in nuage was only partially impaired. This result suggests that the SRS motif in Tej is not the only factor for tethering Spn-E to the nuage granules *in vivo*. These results indicated that the truncated Tej variants inducted the incomplete formatted nuage granules *in vivo*.

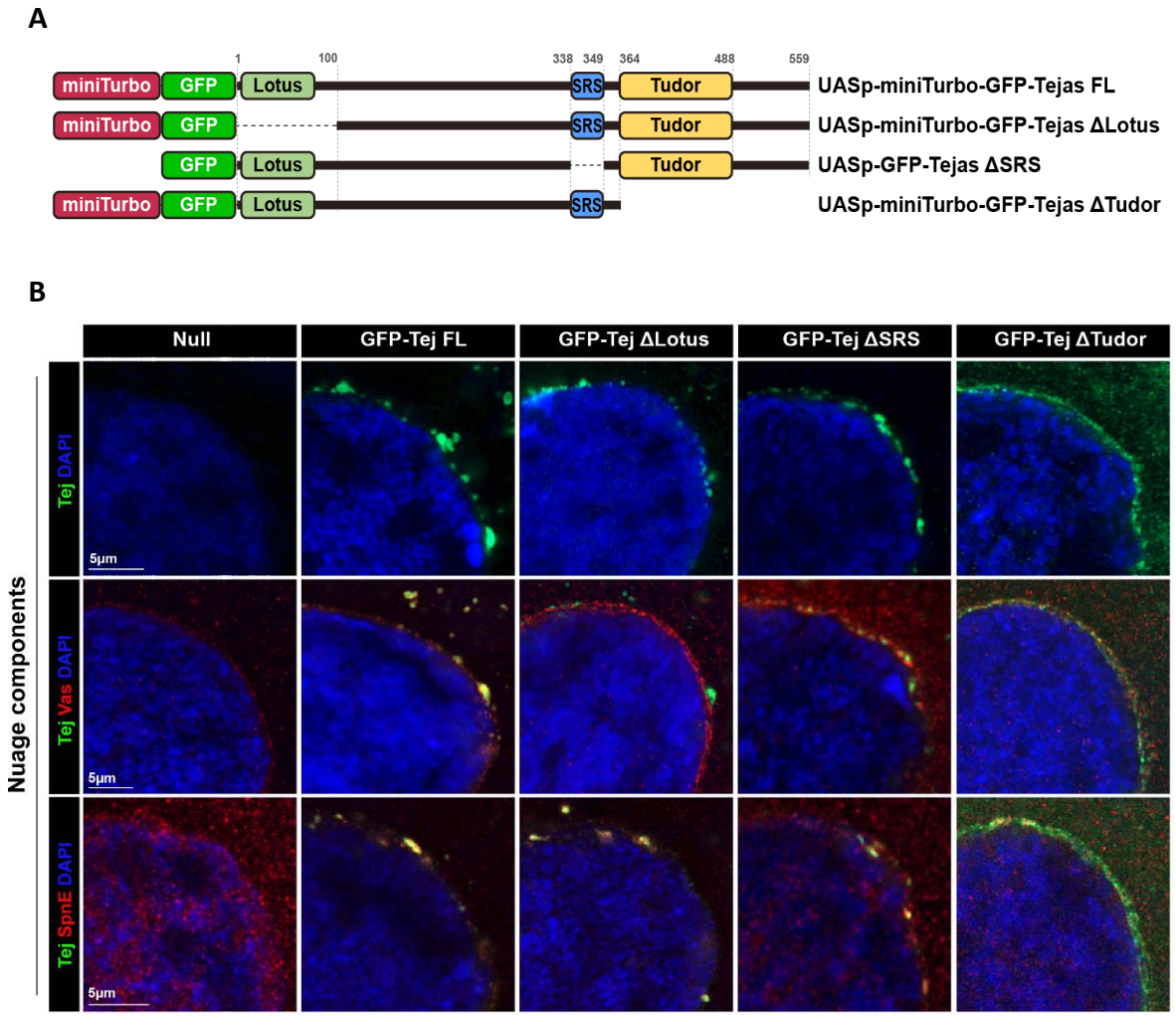


Figure 3.3.3. Tejas control the nuage granule localization of Vasa and Spn-E *in vivo* via distinct domains.

A) Creation of the transgenic fly line expressing GFP-tagged Tej truncation variants. Schematic representation of the Tej truncation variants transgenes.

B) Tejas control the nuage granule localization of Vas and Spn-E in vivo via distinct domains. miniTurbo-GFP-tagged Tej-FL, Tej ΔLotus, Tej ΔTudor, and GFP-tagged Tej ΔSRS (green) were expressed in *tej* mutant *Drosophila* ovaries under the control of promotor UASp, driven by the driver *tj*-Gal4 and *nos*-Gal4. Endogeneous Vas and Spn-E were immunostained with antibodies (red), respectively. DNA is stained with DAPI (blue).

3.3.4 Incomplete nuage granules lead to accumulation of piRNA precursor

It was revealed that the deletion of the Lotus domain or SRS motif of the Tej impaired the localization of either Vas or Spn-E in the nuage, respectively. I further checked the behavior of precursors in the ovaries rescued by these Tej truncation variants with HCR-FISH. In the *tej* mutant ovary, the precursors have accumulated proximity in both sides of the nuclear membrane as described previously (Figure 3.3.2 B). Only little amount of precursors observed inside the nucleus shows that they were probably transported and processed. Moreover, the rarely observable perinuclear precursor signals were separated from the significant nuage granules, which suggested they are the transiently existing precursors before further processing in the wild-type ovaries(Figure 3.3.4 A).

Interestingly, Tej ΔLotus could not rescue the accumulation of piRNA of precursors as seen in the absence of Tej. The piRNA precursors were concentrated around the incomplete nuage granules, detaching Vas from the nuage (Figure 3.3.4 A), which is similar, but the precursors were also accumulated milder when Tej ΔSRS was expressed in the *tej* mutant ovaries (Figure 3.3.4 A). This entangled precursor at the perinuclear region indicated that the procession of precursors was arrested in these foci in the absence of the interactive region of Tej. While the loss of the Lotus domain or SRS motif in Tej caused the different piRNA precursors accumulation, the Tej ΔTudor domain rescued the accumulation of precursors. This result suggests that although the granule formation was lost, Tej ΔTudor was enough for dredging traffic-jammed precursors from the perinuclear region though it holds the Lotus domain an SRS motif (Figure 3.3.4 A). Then I performed qRT-PCR to measure the level of the accumulation of precursors, giving consistent results with the accumulation phenotype revealed by the HCR (Figure 3.3.4 C).

To examine further whether the suppression of the transposon by the piRNA pathway was impaired in the absence of the interactive regions of the Tej, I immunostained the ovaries with the antibody that recognizes the Gag protein derived from the transposon *Het-A*. The massively expressed Het-A can be observed in the oocyte at the posterior of the egg chamber of the *tej* mutant ovaries (Figure 3.3.4 B). The Tej-FL completely suppressed this expression of HetA while the Tej ΔLotus failed to suppress in the *tej* mutant background. (Figure 3.3.4 C). As the Tej FL rescued this precursor accumulation phenotype and indicated that the N terminals conjugated tags have a negligible effect on the biological function of Tej variants. A relatively high expression level of Het-A was observed in the Tej ΔLotus rescued ovaries, whereas the Tej ΔSRS and Tej ΔTudor variants only show a much milder defect in transposon suppression (Figure 3.3.4 B). The results of qRT-PCR for the transposons show a corresponding trend with the HetA staining in the expression of Tej truncated variants (Figure 3.3.4 D). These data supported a conclusion that the Lotus domain bears greater importance than the SRS motif and Tudor domain in Tej in sustaining the precursor processing pathway.

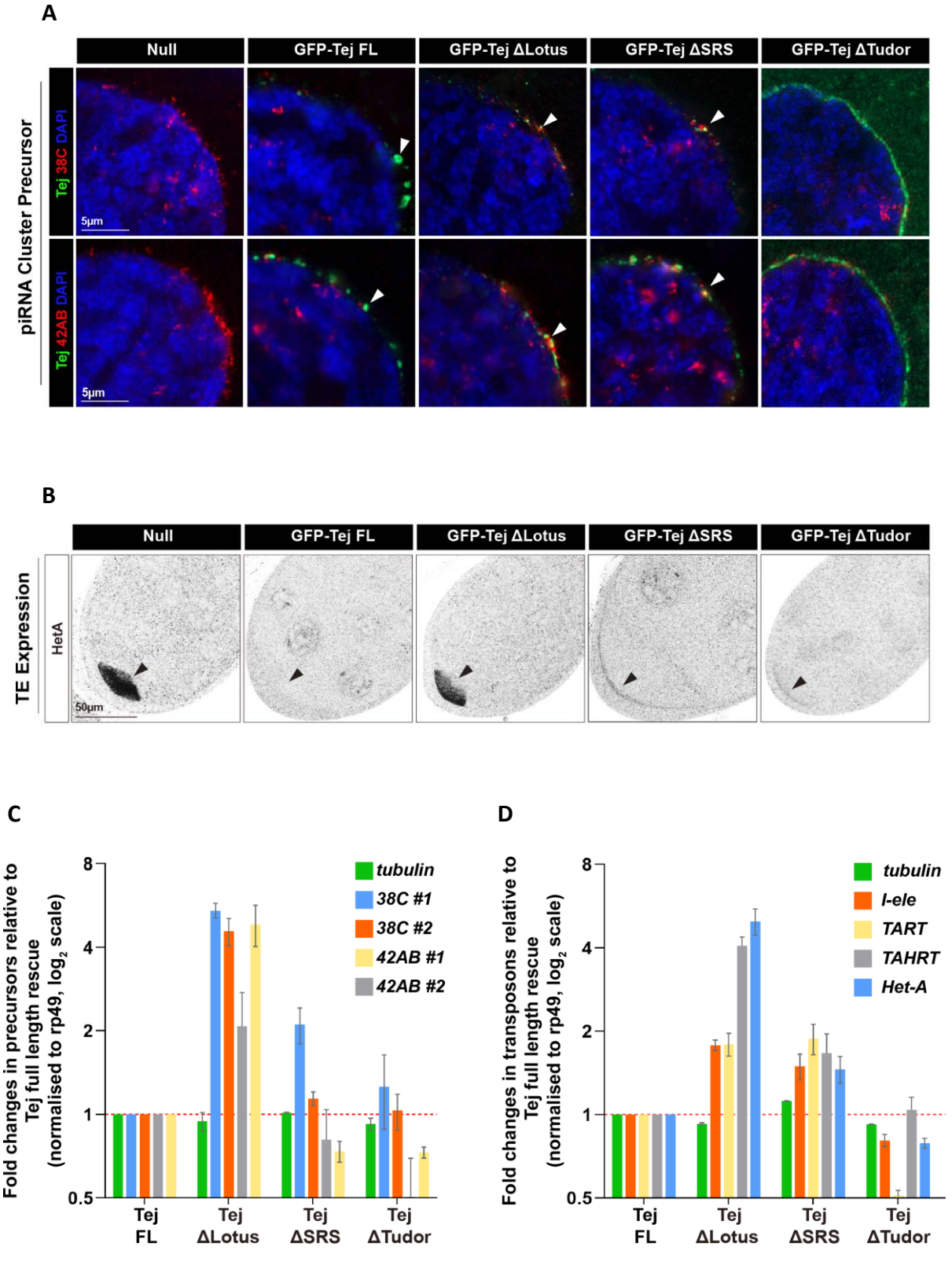

A
Null
GFP-Tej FL
GFP-Tej ΔLotus
GFP-Tej ΔSRS
GFP-Tej ΔTudor
piRNA Cluster Precursor
Tej 38C DAPI
Tej 42AB DAPI
5μm
B
TE Expression
HetA
50μm
C
Fold changes in precursors relative to Tej full length rescue (normalised to rp49, log2 scale)
tubulin
38C #1
38C #2
42AB #1
42AB #2
Tej FL
Tej ΔLotus
Tej ΔSRS
Tej ΔTudor
D
Fold changes in transposons relative to Tej full length rescue (normalised to rp49, log2 scale)
tubulin
I-ele
TART
TAHRT
Het-A

Figure 3.3.4. Tejas functions in piRNA precursors processing pathway via recruitment of Vas and Spn-E to nuage.

A) piRNA precursors derived from *cluster 42AB* and *38C* (red) were stained by HCR-FISH. White arrows marked the congested precursors around nuage granules. DNA is stained with DAPI (blue).

B) Transposable element *Het-A* expression partially rescued by Tej truncation variants. Gag protein (black, black arrow marked) derived from the ovaries specific *Het-A* transposon transcripts were immunostained. All the immunofluorescence signal has been converted to grayscale with inverse hue.

C) Bar graphs showing fold changes of transcripts derived from *cluster 42AB* and *38C* in the total RNA extracted from the Tej truncation variants rescued *tej* mutant ovaries, *tubulin* as a control. Relative RNA level to the Tej-FL in rescued ovaries by each truncated variant, all RNA levels are normalized to *rp49*. Error bars indicate standard deviation, n=3.

D) Bar graphs showing fold changes of transposon transcripts in the total RNA extracted from the Tej truncation variants rescued *tej* mutant ovaries. Relative RNA level to the Tej-FL in rescued ovaries by each truncated variant, all RNA levels are normalized to *rp49*. Error bars indicate standard deviation, n=3.

3.3.5 The Tej is insufficient for the nuage localization of Spn-E.

In contrast with the behavior of Tej ΔSRS in the S2 cell, the depletion of SRS on Tej did not weaken the Spn-E nuage recruitment in the nurse cells of ovaries as strongly as I expected. This phenomenon suggested that Tej is not the only factor required for the proper localization of the Spn-E to the nuage. To get further explicit information about their organization in the nuage, I visualized the localization of Vas, Spn-E and Vasa in the *vas* and *spn-E* mutants.

In the *vas* mutant, most Tej saved their nuage localization at the stage 4 egg chambers and gradually detached from the perinuclear region towards the cytoplasm at the later stage, robustly correlated with the reduced expression level of Spn-E (Figure 3.3.5, top). When Spn-E is missing, the Tej was massively detached from the nuage and formed a droplet-like doty cytoplasmic focus together with the Vas, while some amounts of Vas remain on the nuage (Figure 3.3.5, bottom), consistent with the behavior described above in S2 cells (Figure 3.1.3.1 A). These recruiting priorities revealed the dependence of different nuage components in establishing the nuage formation.

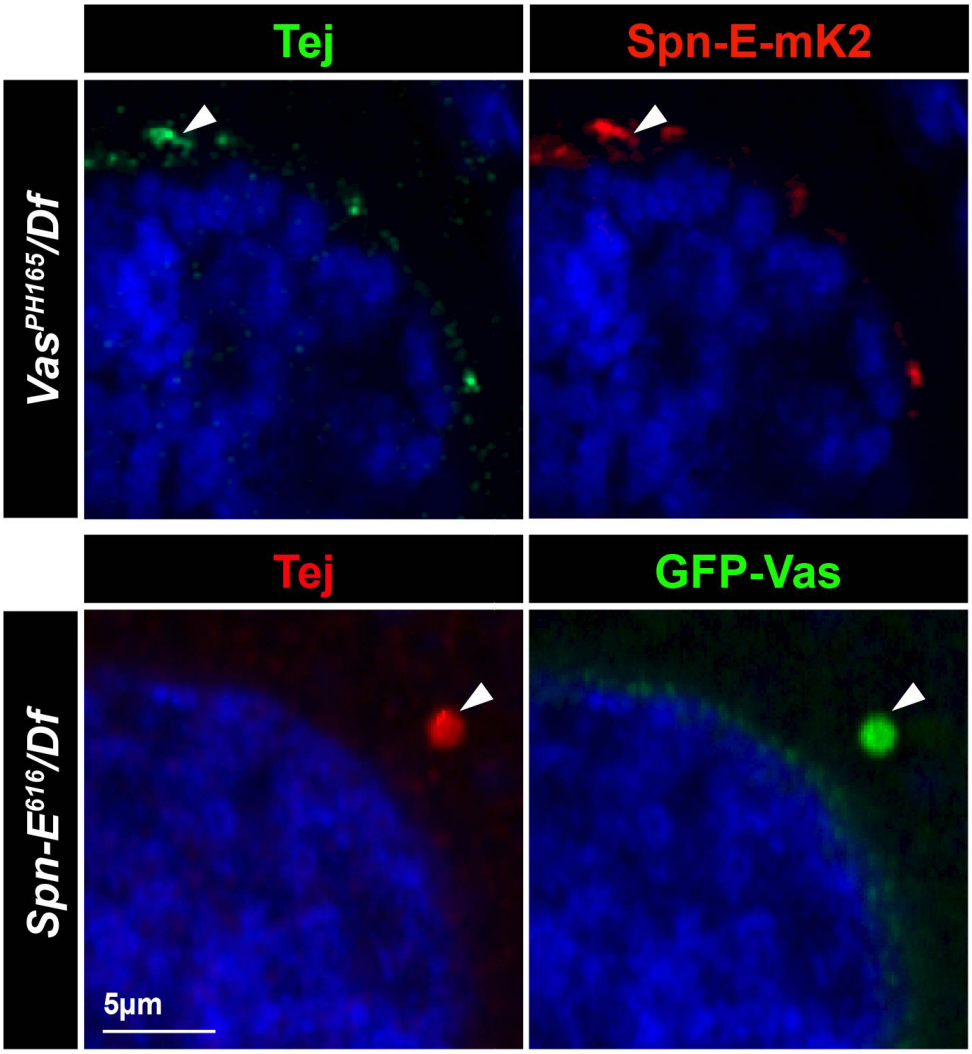


Figure 3.3.5. The recruiting priorities of nuage components.

Tej and Spn-E mutually sustain their localization of each other. The *vas* or *spn-E* mutant ovaries were immunostained with polyclonal antibodies against Tej protein (top, green; bottom, red), with the presence of fluorescence-tagged Spn-E proteins (top, red) or Vas (bottom, green). The white arrow marked the overlapped signal. DNA is stained with DAPI (blue).

3.4 Tej contributed to the kinetics of nuage

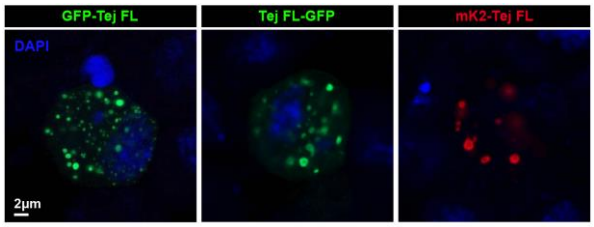
3.4.1 Disordered region endows the formation of condensates to Tej

During the development of *Drosophila* oogenesis*,* several phase separation-regulated granules had been unearthed (Lin et al. 2008). The nuage is also known as one of the non-membranous structures that assemble various proteins and RNAs, speculated to be organized by the liquid-liquid phase separation(ref). GFP-tagged Tej formed significant droplet-like condensates in the cytoplasm of S2 cells (Figure 3.4.1 A). Either switching the fluorophore from N-terminus to the C-terminus of Tej or replacing the fluorophore from GFP to mKate2 did not affect the formation and localization pattern of the Tej-formed condensates (Figure 3.4.1 A). This behavior indicates the possibility of Tej being involved in phase separation. It has been reported that the harboring of intrinsically disordered regions (IDRs) endows the phase separation feature to the protein, and I tried to identify the IDRs about Tej protein by using IUPred2 (Figure 3.4.1 B). As a result, the middle region of Tej (100-362 aa) between the Lotus and Tudor domain displays a highly disordered feature that can be defined as an IDR. The location of the disordered region on Tej is distinct from its vertebrate homologs, which carried a major IDR on their C terminal (Figure 3.4.1 B).

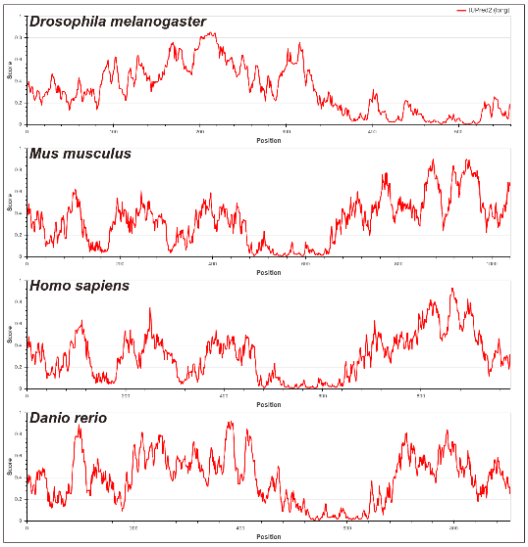
I expressed the N-terminus GFP-tagged Tej Middle part (101-362 aa) in S2 cells and observed that the Tej by GFP fluorescence was broadly smeared in the cytoplasm (Figure 3.4.1 C). Moreover, adding the Tudor domain on the middle part of Tej rescued the distribution and formed the condensate but not adding the Lotus domain instead(Figure 3.4.1 C). The reciprocal experiment showed that deletion of a highly disordered region of Tej (140-294 aa) from the Tej FL (Tej ΔIDR) formed a looser structure compared with the impacted drop-like condensates by Tej-FL (Figure 3.4.1 D, E). The notable point is that Tej ΔIDR retained the ability to recruit both Vas and Spn-E in the S2 cells, respectively (Figure 3.4.1 E). These results

suggested that the IDR facilitates the formation of droplet-like Tej condensates while the Tudor domain provides a strong aggregation ability.

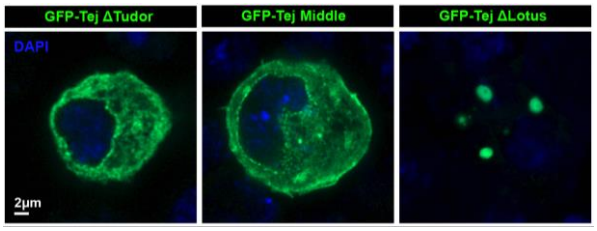
A


GFP-Tej FL
Tej FL-GFP
mK2-Tej FL
DAPI
2μm

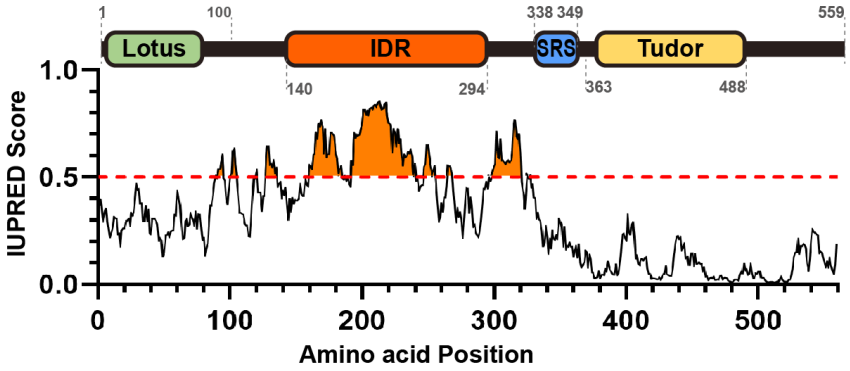
B


Drosophila melanogaster
Mus musculus
Homo sapiens
Danio rerio
Position

C


GFP-Tej ΔTudor
GFP-Tej Middle
GFP-Tej ΔLotus
DAPI
2μm

D


1
100
338 349
559
Lotus
IDR
SRS
Tudor
140
294
363
488
IUPRED Score
1.0
0.5
0.0
0
100
200
300
400
500
Amino acid Position

E

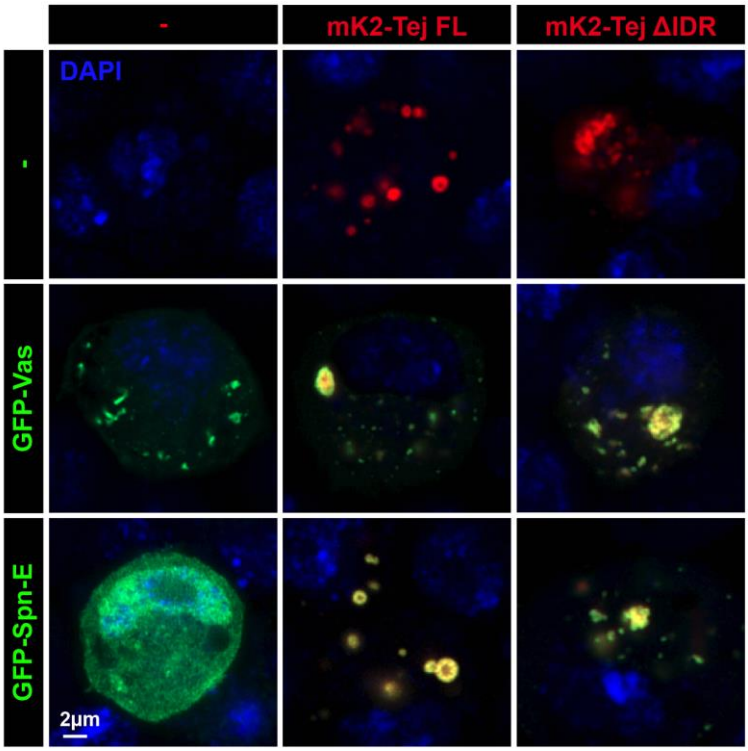

-
mK2-Tej FL
mK2-Tej ΔIDR
DAPI
GFP-Vas
GFP-Spn-E
2μm

Figure 3.4.1. Tej contains a large intrinsic disordered region.

A) Tej formed cytoplasmic droplet-like condensates unaffected by the fluorophore. N or C terminal GFP-tagged and N-terminal mKate2-tagged Tej were expressed in the S2 cells. DNA is stained with DAPI (blue).

B) Tej contains large intrinsic disordered region (IDR). Disordered protein regions prediction results of the Tej and vertebrate homologs were listed.

C) IDR and Tudor of Tej are essential for forming the condensate. GFP-tagged Tej truncated variants were expressed in S2cells. DNA is stained with DAPI (blue).

D) Schematic structure of Tej and the position of intrinsically disordered regions (IDR) by the IUPRED prediction. The central part with an IUPRED score > 0.5 was defined as the IDR of Tej.

E) The Vas and Spn-E recruitment of Tej is independent of the IDRs of Tej. mKate2-tagged full-length Tej (Tej FL) or IDR truncated Tej (Tej ΔIDR) is co-expressed with the GFP-tagged Vas or Spn-E in S2 cells, respectively. Images show the single expression of all the fluorescent-tagged proteins was shown in the top and left panels, DNA is stained with DAPI (blue).

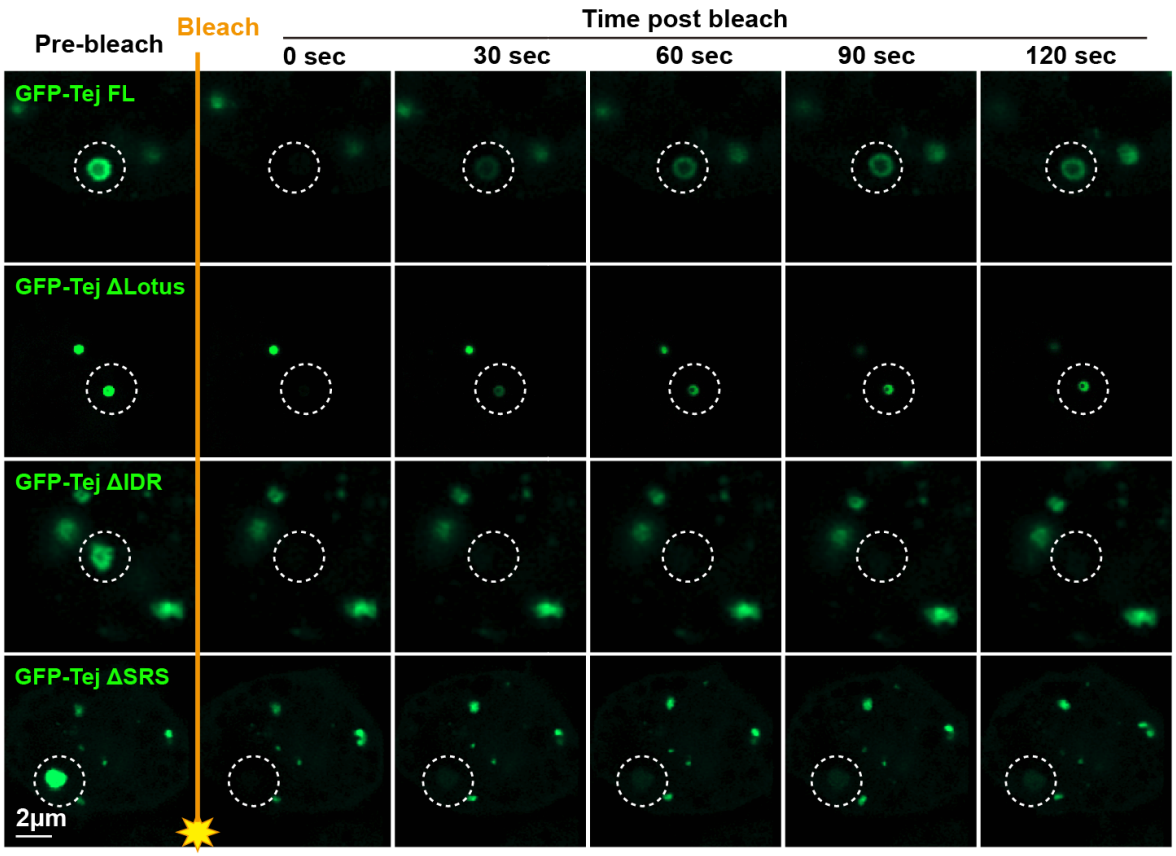
3.4.2 IDR of Tejas controls the mobility of nuage components.

As Tej can form the round droplet-like condensates in the S2 cells like the other non-membranous granules, I further analyzed the dynamics of these aggregated Tej molecules to examine their mobility and determine whether they are exchanged in the condensates according to the environment. GFP-tagged Tej variants that form aggregations, including the Tej-FL, Tej ΔLotus, Tej ΔIDR, and Tej ΔSRS, were expressed in the S2 cells. Under a fluorescent confocal microscope, these variants were observed as granules in the living S2 cell. Furthermore, by utilizing the fluorescent recovery after photobleaching (FRAP), I photobleached the granules expressed in each GFP-tagged Tej variant in the S2 cell during the live imaging and monitored the fluorescent recovery the mobility of Tej molecules (Figure 3.4.2 A). After photobleaching, the Tej-FL rapidly exchanged with the bubble-like granule, and 65 percent of fluorescent is recovered with a half time ($t_{1/2}$) of 47.62 seconds to reach the saturation fluorescent recovery. I also recorded a similar FRAP recovery for the Tej ΔLotus granules with 75 percent recovery in a half time ($t_{1/2}$) of 30.52 seconds, and even it was recovered quicker than the full-length Tej. Distinctively, the slightly deformed granules in Tej ΔIDR showed a significantly reduced recovery rate to 10 percent. Unexpectedly, Tej ΔSRS also showed a low recovery rate, 33 percent of fluorescent was recovered after bleaching (Figure 3.4.2 B, C). These tests revealed that the IDR of Tej facilitates the dynamic of the Tej-formed granules.

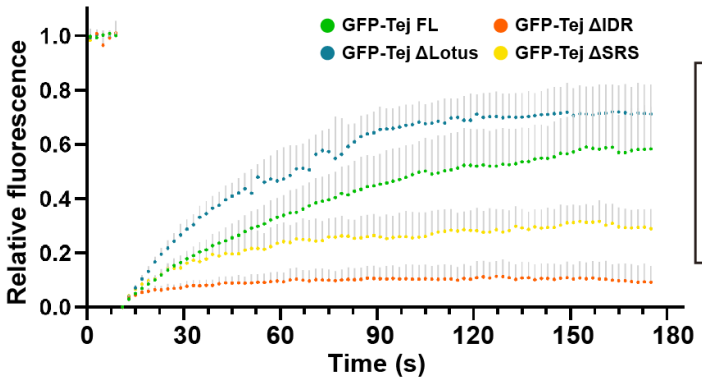
Given the highly dynamic behavior of the Tej formed granules, I further hypotheses that this feature contributes to the factor exchange of the other nuage components. Thus, I examined the mobility of the helicase Vas and Spn-E that are colocalized in the Tej formed granules in the S2 cells. Either mK2-tagged Tej-FL or Tej ΔIDR were co-expressed in the S2 cells with the GFP-tagged Vas or Spn-E. As seen above, both mK2-tagged Tej-FL and Tej ΔIDR

formed sites for the cytoplasmic localization of Vas and Spn-E S2 cells (Figure 3.4.2 D). I performed FRAP for the GFP channel and recorded the recovery status of Vas or Spn-E in GFP fluorescence. GFP-Vas colocalized with Tej-FL showed a rapid recovery with a 15.63 second half time ($t_{1/2}$), and the recovery rate arrived at 86 percent, while GFP-Vas expressed with Tej ΔIDR lost their mobility drastically, showing only 56 percent of fluorescent was recovered. However, I observed an unusual behavior of Spn-E when expressed with the Tej variants. Spn-E co-expressed with either Tej-FL or Tej ΔIDRs formed granules and showed prolonged and low-level fluorescent recovery when compared to the high mobility of Vas. Collecting these FRAP results *in vitro*, the IDR contributes to Tej molecules' mobility and endows Tej the ability to control the mobility of the other nuage components, including Vas and Spn-E.

A



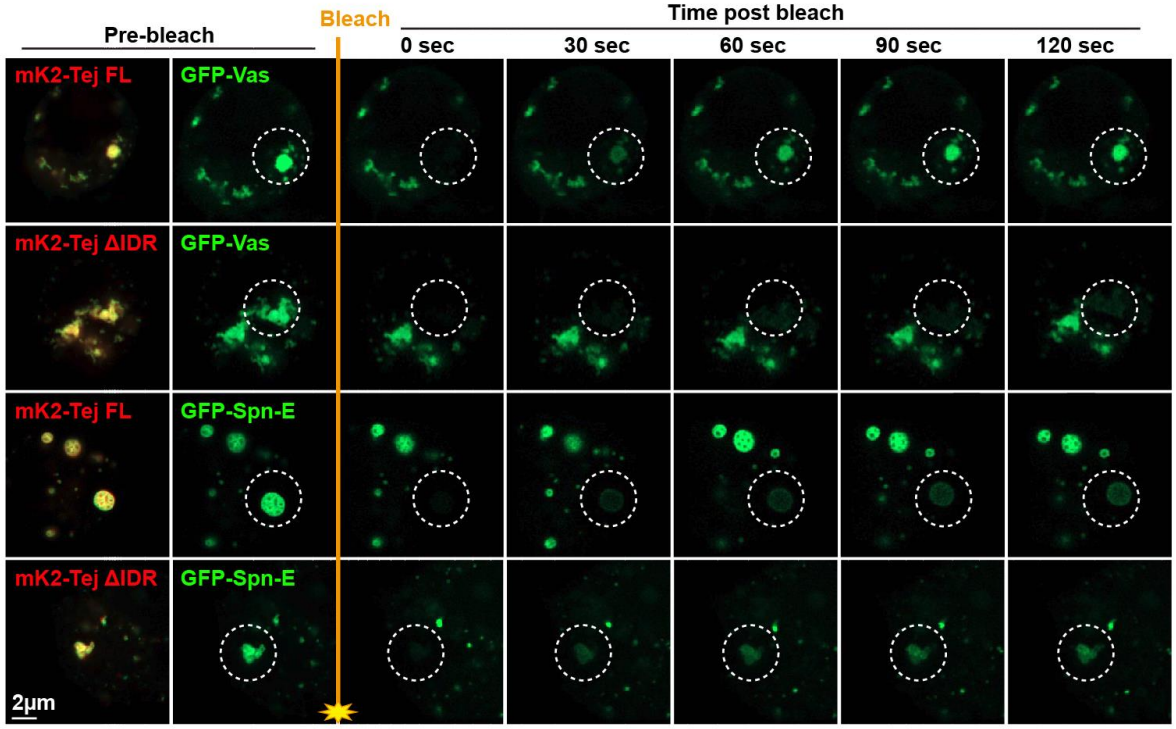
B



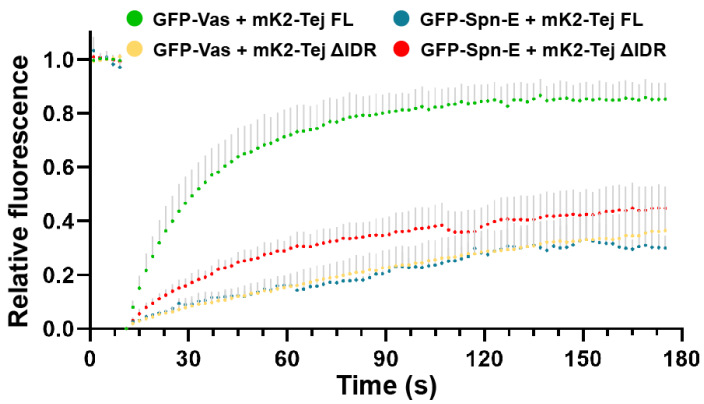
C

	Mobile fraction (%)	$t_{1/2}$ (s)
GFP-Tej FL	0.65	47.62
GFP-Tej ΔLotus	0.75	30.52
GFP-Tej ΔIDR	0.10	2.22
GFP-Tej ΔSRS	0.33	21.71

D



E



F

	Mobile fraction (%)	$t_{1/2}$ (s)
GFP-Vas +mK2-Tej FL	0.86	15.63
GFP-Vas +mK2-Tej ΔIDR	0.56	107.62
GFP-Spn-E +mK2-Tej FL	0.47	95.43
GFP-Spn-E +mK2-Tej ΔIDR	0.45	36.48

Figure 3.4.2. IDR of Tejas controls the mobility of nuage components.

A-F) Images in the graph show fluorescence recovery of GFP signals before and after photobleaching. The white dotty line shows the bleached area. The relative fluorescence strength was recorded per second. The line plot shows the normalized relative fluorescence recovery rate. Colored dots show the mean of all experimental data. The mean ± SD of five condensates measurements is shown in gray bars above. Besides the graph, the proportion of mobile fraction and $t_{1/2}$ derived from the mean value fitting curves are shown.

A) IDR of Tej controls the kinetic of Tej-formed aggregations. Fluorescent recovery after photobleaching (FRAP) experiment applied to the aggregations formed by GFP-tagged Tej-FL, Tej ΔLotus, Tej ΔSRS, and Tej ΔIDR (green) in S2 cells.

D-E) IDR of Tej controls the dynamic of Vas and Spn-E in Tej-formed granules. mKate2-tagged Tej-FL or Tej ΔIDR (red) was expressed in S2 cells along with the GFP-tagged Vas or Spn-E (green).

3.4.3 Investigation of the potential RNA binding ability of Tej

Most phase separation-dependent condensates arise from RNAs and proteins. It is already known that some RNAs can promote its phase separation with RNA-binding proteins (RBP)(Pak et al. 2016). TDRD5, the mouse homolog of Tej, had been reported to bind to the piRNA precursors directly in the mouse male germ cells (Ding et al. 2018). I further investigated whether Tej is also an RBP that interacts with RNAs directly.

I performed crosslinking immunoprecipitation to pull down the Tej-GFP complexes from Tej-GFP-expressing ovaries and the wild-type control. To remove the co-immunoprecipitated factors of proteins? that may influence the results, I followed strict high salt conditions for protein complexes-beads conjugation and beads washing steps. As a result, this experiment led to a high proportion of Tej-GFP remaining in the eluted fraction. Other nuage components can be stained with fainter bands by western blotting compared with the normal washing condition (normal wash condition, Figure 3.1.2 B; strict wash condition, Figure 3.4.3 A). The putative Tej-bound RNAs were extracted from the immunoprecipitated complexes after reverse crosslinking and were labeled with 32P-γ-ATP following to dephosphorylation of 5´ ends. The putative Tej-bound RNAs showed a smearing signal after electrophoresis with denaturing urea polyacrylamide gel, indicating that the elution contains long RNA fragments with a length up to ~300 nucleotides (Figure 3.4.3 B). The result suggested the capability of Tej to bind RNAs directly.

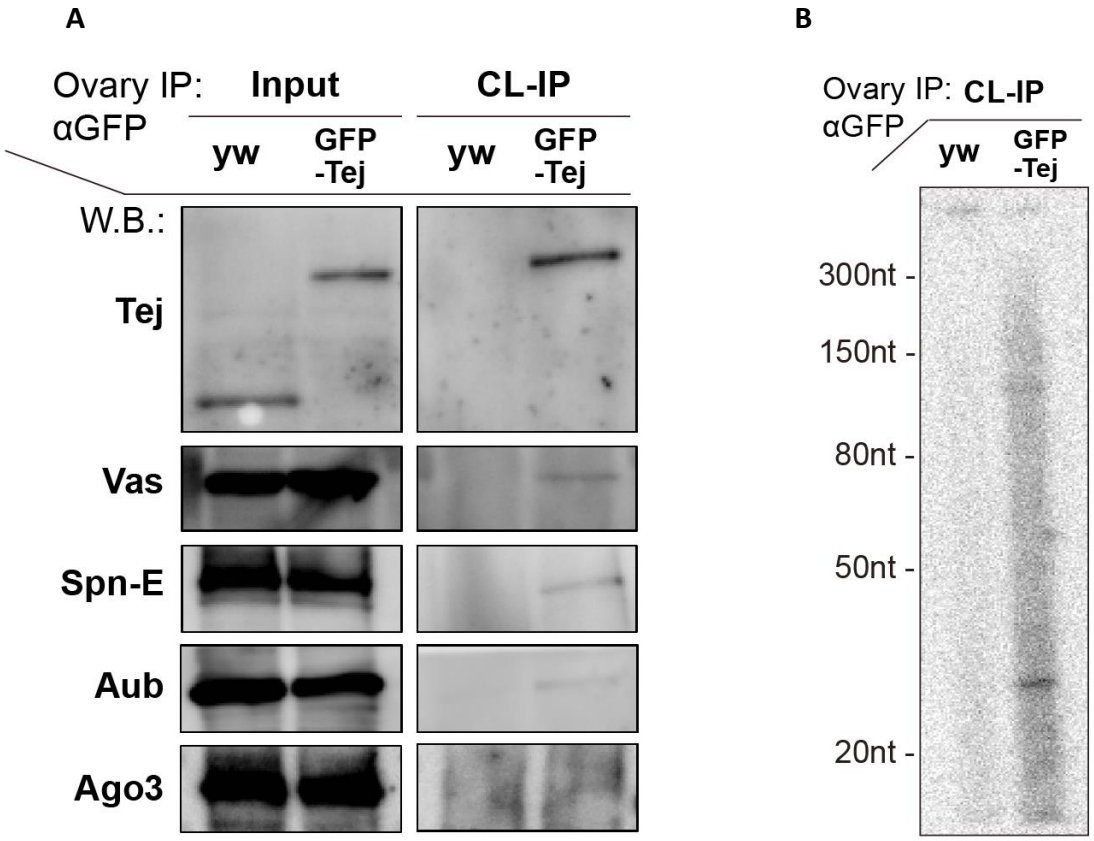


Figure 3.4.3. Tej is an RNA binding protein.

A) Immunoprecipitated Tej-GFP complexes with crosslinking. The endogenous Tej-GFP complexes were immunoprecipitated from the lysate of the *Drosophila* ovary with harsher binding and washing conditions. Ovaries from *y w* flies were used as control. Major piRNA biogenesis factors were detected by western blot in each complex. 10% of total extracted complexes were loaded as Input.

B) Putative Tej binding RNAs weres detected. RNAs extracted from the immunoprecipitated complexes elution was 32P-γ-ATP labeled and electrophoresis with denaturing urea polyacrylamide gel.

4. DISCUSSION

In the *Drosophila* germline piRNA pathway, a significant proportion of piRNAs was produced through the 'ping-pong' amplification pathway, which forms at nuage structure in the perinuclear region of the nurse cells. The nuage requires highly organized bio-molecular machinery in which proper components of proteins and RNAs must be correctly assembled. Tejas (Tej) is an essential protein that assists in operating the piRNA amplification machinery correctly and localized to the nuage. Previously, Tej has been reported to be a core position on the organization of nuage structure (Patil and Kai 2010). Missing of Tej collapses the nuage formation completely. However, the details of the mechanism, how each factor is recruited to this core protein in nuage, and how Tej regulates one of the steps in processing piRNA precursors are yet to be fully understood.

In this research, I further investigated the molecule functions of Tej in aiming to reveal the details for the molecular basis of piRNA biogenesis, which allows it to engage in the organization of the nuage structure in the nurse cells. The results suggested that Tej regulates the dynamics of Spn-E and Vas as a core organizer of nuage, is involved in the processing of piRNA precursors, and utilizes the intrinsically disordered region to control the mobility of nuage.

4.1 Tej plays a crucial role in organizing nuage components as an essential factor controlling piRNA biogenesis

Visualizing the endogenously expressed fluorophore-tagged proteins of nuage enabled me to refurbish the dynamics of nuage. With these Knock-In fly lines for nuage components, I succeeded in observing perdurable impact in the proper organization of the nuage and no longer fluctuated localization depending on the quality of polyclonal antibodies. As a result, dissemble of fluorophore-tagged components at the nuage can be reproduced in the loss-of-

tej ovaries (Figure 3.1.1 C). The cytoplasmic-diffused Aub and Ago3 foci detached from the nuage are consistent with the previous studies (Patil and Kai 2010). The meticulous observation by utilizing the Knock-In fly lines has been achieved the new findings of misregulated Vas and Spn-E. Distinct from the conclusion visualized by the antibody staining, I found the Vas lost their aggregation in the prominent nuage granules, although the distribution at the smooth perinuclear layer remains unaffected (Figure 3.1.1 C). The Vas proteins were segregated into the granulated nuage particles and supernumerary small doty attached to the nuclear membrane. This distinctly localized Vas indicates that the Vas plausibly has exceeded functions other than nuage formation and piRNA processing when presented on the perinuclear region (Figure 3.1.1 B).

No noticeable nuclear fraction of Spn-E was found in the wild-type nurse cells in *Drosophila* (Figure 3.1.1 C), different from mouse TDRD9, a homolog of Spn-E, which presents both germinal granules/nuage and nuclear in spermatids (Shoji et al. 2009; Wenda et al. 2017). In contrast, the loss of Tej led to a nuclear accumulation of the Spn-E. Considering the nuclear-localization of Spn-E homolog, TDRD9 in the mouse spermatids, genetically and physically associated with MIWI2, the mouse homolog of Piwi (Shoji et al. 2009). In addition, the ATPase TDRD9 is known to be dispensable for its nuclear localization (Wenda et al. 2017). It was thought that these abnormally smearing signals of Spn-E in the nucleus might give some clues about the Spn-E function. When expressed in the S2 cell, the Spn-E firmly accumulated inside the nucleus of S2 cells (Figure 3.2.1 B). As a relatively large protein weighing 159.17 kDa, the nuclear Spn-E ought to be translocated depending on the nuclear import machinery. As expected, I verified that a putative novel class II monopartite NLS enabled Spn-E to be located inside the nuclei (Figure 3.2.1 A, B).

Additionally, observation of the mK2-tagged Spn-E in the ovarian somatic cells tells us that a faint but genuine expression of Spn-E was spotted in the nuclei (Figure 3.2.1 C). The malfunctioned NLS on Spn-E (Spn-E △NLS) canceled the nuclear localization of Spn-E in the ovarian somatic cells, while some remain intact in the localization around the perinuclear of the ovarian nurse cells (Figure 3.2.1 D). These results indicated that the Spn-E is intrinsically a nuclear protein that contains NLS but is unusually transported and located to the perinuclear nuage with an unknown recruitment mechanism.

4.2 Tej interacts with Vas and SpnE via its distinct domains/motifs, and Tej forms two exclusive complexes with Vas or Spn-E.

Both Vas and Spn-E are the critical helicases for the piRNA biogenesis and are mislocalized in the *tej* mutant ovaries. Missing of Tej showed the Vas and some Spn-E proteins remained perinuclear region but were segregated from each other while aggregating particles in the wild-type ovaries. These results indicated that Tej is required for their proper connection (Figure 3.1.2 A). Crosslinking immunoprecipitation enables us to capture the transient interactions among Tej, Vas, and Spn-E *in vivo*. Vas and Spn-E were detected in mutually exclusive subcomplexes with Tej (Figure 3.1.2 B-F). A similar type of subcomplexes, respectively containing either Vas or Spn-E, have been reported in the *Bombyx* germ cells, Siwi/BmVasa complex involves neither BmSpn-E nor BmQin/Bmkumo. Unlike *Drosophila*, BmSpn-E was only partially colocalized with BmVasa and displayed ordinary segregated granules (Nishida et al. 2015), while the Spn-E and Vas forms firmly merged nuage granules in the ovaries (Figure 3.1.2 A). If Spn-E and Vas in the *Drosophila* nuage needed to be spatially isolated each other for their proper functions as observed in the *Bombyx*, Tej would perhaps require for creating such a condition. Co-existing Vas and Spn-E were effectively aggregated to the Tej-formed granules in the S2 cells and segregated into two distinct structures, and Vas

encircled the core-localized Spn-E. Spn-E was aggregated when they were co-expressed with Tej and showed the incompatible localization from Vas. It is suggested that aggregated Spn-E brings mutually exclusive physicochemical properties to the complex(Figure 3.1.4). RNA-dependent DEAD-box ATPases (DDXs) have been reported to form non-membranous phase-separated organelles that selectively recruit or release their components, such as RNAs and proteins (Hondele et al. 2019). These results raise the possibility of RNA-dependent DxxD-box ATPases Vas and Spn-E being phase-separated features in the nuage. Vas possesses a large IDR region in its N terminal, facilitating its phase-separated status (Nott et al. 2015). The highly structured Spn-E with several conserved domains but without notable IDRs plausibly gained a phase-separation ability depending on Tej.

As Tej interact with Vas and Spn-E, I wonder which domains are required for these interactions. Lotus and Tudor domains, characterized by crystal structures harbored at N- and C-terminus of Tej protein. By co-expression of Tej truncated variants along with GFP-tagged Vas or Spn-E in the S2 cells (Figure 3.1.3.1 A), I verified that the Lotus domain is responsible for the Vas interaction of Tej (Figure 3.1.3.1 A), which is consistent with the previous study (Patil and Kai 2010). The Lotus domain has been reported to endow the ability of homodimerization for Oskar protein, mediates the interaction between Oskar and helicase Vas, which can enhance the ATP hydrolysis efficiency of Vas. The lotus domain has also been identified as a novel G-rich G4 RNA structure binding domain (Jeske et al. 2015; Jeske, Müller, and Ephrussi 2017; Ding et al. 2020). The Lotus domain deficient Tej variant even more strongly forms round multi-droplet condensates, excluding Vas completely, while absorbing all the Spn-E (Figure 3.1.3.1 A).

Further truncation of the Tudor domain leads to a loose morphology of the Tej-formed condensates, which indicates the Tudor domain is indispensable for the self-aggregation of

Tej. The Tudor domain recognizes and binds sDMA modifications (Kirino et al. 2010; Sato et al. 2015), also been reported to contribute to the formation of condensates in eukaryotes bound to the sDMA containing peptides (Courchaine et al. 2021). The Tej variant lacking the Tudor domain lost its aggregation both *in vitro* (Figure 3.1.3.1 A) and *in vivo* (Figure 3.3.3 B), which supports that the Tudor domain of Tej functions to contribute the condensate formation.

In the analysis by the sequential truncation of Tej, I identified the short but highly conserved motif named 'Spn-E Recruit Site'. This motif contains several evolutionary conserved amino acid residues from *Drosophila* to vertebrates (Figure 3.1.3.2 B). Amino acids substitution on the SRS motif significantly malfunctions the Spn-E recruitment of Tej *in vitro* (Figure 3.1.3.1 A) and partially reduces the Spn-E nuage localization *in vivo* (Figure 3.3.3 B, bottom). The SRS motif on Tej delivers the robust colocalization between Tej and Spn-E described above, giving the colocalization at the perinuclear nuage that Tej controls the subcellular localization of Spn-E by inhibiting its intrinsic feature of nuclear deposition.

Consequently, I questioned whether the intrinsic nuclear localization of Spn-E is due to its potential function in the nucleus or just misregulated when the nuage is disintegrated under the absence of Tej. Spn-E is a Tudor domain-containing DExH helicase that has the ability to bind and unwind double-strand RNA duplex with ATPs hydrolysis (Sengoku et al. 2006). Then I proposed the hypothesis that Spn-E might work as a shuttling for transcribed piRNA precursors between nuclear and perinuclear nuage. I further tried to verify this hypothesis by searching the potential co-factors of the nuclear-localized Spn-E. Neither DNA nor the nucleolus marker, which be considered a prominent region for transcription and RNA processing, is overlapped with the Spn-E signals (Figure 3.2.2 A, C). Next, I tried to visualize the precursor transcripts derived from the piRNA cluster *42AB* and *38C* by HCR-FISH and

combined with the mK2-tagged Spn-E under the *tej* mutant background to ensure whether the Spn-E captures the piRNA precursors directly or stays adjacent to them in the nucleus. However, the results indicate no support for the hypothesis that Spn-E acts as precursors carrier as no overlapping of Spn-E and piRNA precursors can be observed (Figure 3.2.2 D). Thus, further investigation is needed to clarify the potential function of the accumulating Spn-E in the nucleus of tej mutant ovaries.

4.3 Tej sustain the nuage localization of Vas and Spn-E, and this regulation is vital for piRNA precursor processing.

The reduction of the piRNA biogenesis in Tej lacking ovaries was shown by northern blotting (Patil and Kai 2010) and small RNA sequencing (Patil et al. 2014). To gain detailed information to identify the failure step of the piRNA production, I checked the piRNA binding status of the two major components of the ping-pong pathway, Aub and Ago3, which set at the relatively downstream of the piRNA biogenesis pathway. The RNA-immunoprecipitation experiment revealed that both Aub and Ago3 are massively unloaded with the piRNA in the Tej lacking ovaries (Figure 3.3.1 A, B). Both Aub and Ago3 bound piRNAs were mapped against the major piRNA clusters, *42AB* and *38C*, in ovaries are significantly downregulated (Figure 3.3.1 C, D). Collectively, the results suggest that the ping-pong cycle was largely defected when the nuage lost Tej.

Given the failed ping-pong cycle, I further seek answers about the upstream of piRNA production machinery. With the qPCR and HCR-FISH, an accumulation of the precursors was quantitatively and spatially confirmed in the Tej missing ovaries (Figure 3.3.2 A, B). The accumulated piRNA precursors deposited at the bilateral sides of the nuclear membrane, not only the inside of nuclei, suggest that the biological machinery for exporting piRNA precursors

remains functional even in the mislocalization of Spn-E due to the loss of Tej. The accumulated piRNA precursors suggested that they had been 'traffic-jammed' at the perinuclear region, which raised several possibilities: the componens for interacting the precursors were not correctly recruited to the nuage; the improper formated nuage by the interference of the proper RNA folding or unwinding by RNA helicase; the abolished downstream machinery causes the accumulation of unprocessed precursors. As it is known that the unexported precursors are directed to the RNA decay pathway and eliminated from the nucleus (Kneuss et al. 2019), the similar accumulation of precursors observed in the *tej, vas*, and *spn-E* mutant ovaries (Figure 3.3.2 B) suggested that they are likely to be captured and protected from the RNA degradation mechanism by malfunction nuage components.

In order to verify this hypothesis, I performed the rescue experiments in vivo by utilizing the transgenetic flies that expressed the Tej truncated variants mentioned above. GFP-tagged Tej variants that are particularly defective Tej have been expressed in the *tej* mutant ovaries (Figure 3.3.3 B, top) and impaired the processing of piRNA precursors by inducing the incomplete nuage granules (Figure 3.3.3 B, bottom, Figure 3.3.4 B, C, D). Significantly, Vas localized at the central part of nuage granules was eradicated when the Lotus domain dropped from the Tej, while the loss of the SRS motif did not cause the severe misregulation of Spn-E compared to the phenotype in S2 cells (Figure 3.3.3 B). Further tests show that the precursors are dominantly congested around the disassembled nuage in the Tej ΔLotus, and a similar but milder accumulation was observed when the SRS motif was impaired in Tej (Figure 3.3.4 A). Different extents of accumulation of piRNA precursor and failure of transposons suppression have been detected in the rescued ovaries by the Tej variants (Figure 3.3.4 B-D), which reflect the domains function of Tej in the feature of piRNA biogenesis.

As the Lotus domain was reported to stimulate the ATP hydrolysis activity of Vas and promote its releasing activity of the RNAs (Jeske, Müller, and Ephrussi 2017), combined with my experimental results, the Lotus domain deficient Tej failed to recruit and stimulate Vas, which possibly stop unwinding of piRNA precursors or prevent them from binding to other nuage components. The decisive role of Tej to recruit Vas to nuage granules was supplied by the observation in the *vas* mutant, where the Tej remains strongly localized to the perinuclear region together with a portion of the Spn-E (Figure 3.3.5). The milder defect caused by the missing SRS motif indicated that the nuage localization of Spn-E was not absolute depending on the SRS motif harbored in the Tej. Whether other components, including long-chain RNA or proteins, are required for similar and critical tasks need further investigation. As shown in the Spn-E missing ovaries, the proper localization of Vas and Tej has also been influenced (Figure 3.3.5), suggesting that the nuage does not possess the simple hierarchical organization in the nuage since each component mutually sustains their subcellular localization and functions.

4.4 IDRs of Tej define the mobility of nuage components in vitro, indicating the possibility that Tej facilitate the dynamic of nuage granules

During the development of *Drosophila* oogenesis, several liquid-liquid phase separation-regulated granules had been unearthed (Lin et al. 2008). The nuage is one of the non-membranous organelles assembled by various proteins and RNAs (Meikar et al. 2011), but the molecular mechanisms of phase separation in the nuage remain elusive. Depending on the features of condensates formation, I found an intrinsically disordered region (IRDs) harbored in the middle part of the Tej, which is between the Lotus and Tudor domain (Figure 3.4.1 D). Unlike the IDRs in the N terminus of Vas, which directly promotes the spontaneously self-associate, leading to the formation of non-membranous condensates (Nott et al. 2015),

the shorter (~120 aa) IDR of Tej required the Tudor domain for forming the aggregates (Figure 3.4.1 C). Depletion of the IDR loosens the compacted droplet-like structure formed by Tej, while the Vas and Spn-E recruitment remain unchanged (Figure 3.4.1 E). The loss of IDR significantly suppresses the mobility of Tej molecules in the self-associate structure, showing a reduced fluorescent recovery on a large scale after the photobleaching during the live imaging (Figure 3.4.2 A, B). Notably, the deficiency of Lotus domain in Tej unexpectedly promotes the mobility of Tej (Figure 3.4.2 A, B), potentiality due to the IDR part is no longer impeded by the Lotus domain, shows more substantial flexibility. On the other hand, the loss of the SRS motif also inference the ordinary dynamic of Tej (Figure 3.4.2 A, B). As a highly conserved part between the IDR and Tudor domain, the SRS motif truncation potentially affected the conformation of the Tej protein, which further reduced the molecular kinetic.

Moreover, this measurement was applied to know the dynamics of Vas and Spn-E molecules in the Tej-formed aggregates in S2 cells. Remarkably, my results showed that Vas was highly mobile in the Tej-formed condensates (Figure 3.4.2 D-F), while the Tej ΔIDR formed an irregular structure and significantly blocked the dynamics of Vas (Figure 3.4.2 D-F). A distinct behavior was found on the measurement of Spn-E. The relatively lower fluorescence recovery of Spn-E molecules in the condensates merged with Tej, suggesting that Spn-E is a stationary component in Tej-formed granules. The existence of IDR does not essentially rule the mobility of Spn-E (Figure 3.4.2 D-F). Given these results, I found that the IDR of Tej defines the mobility of some of the nuage components, such as Vas, *in vitro*, then the inference comes following that Tej facilitates the dynamics of nuage granules.

Considering the higher dynamics of Vas and the relatively stationary behavior of Spn-E in the Tej-formed granules, the unique surface-core structure formed by Vas, Tej, and Spn-E in S2 cells may indicate that the Vas is a highly dynamic component of nuage while the Spn-E

is compacted in the core part of nuage. This model may explain the different degrees of piRNA biogenesis controlled in the Lotus or SRS deficiency of Tej *in vivo*. However, in vivo experiments by utilizing the transgenic knock-in fly lines will be needed for the validation of this hypothesis

My data suggested that Tej regulates the subcellular dynamics of Spn-E and Vas as an organizer for nuage and potentially sustains the limited contact between the Vas and Spn-E. A newly validated functional NLS in Spn-E and a novel identified Spn-E recruiting motif in Tej cooperate to control Spn-E's subcellular localization. Tej further engaged in the proper formation of the nuage granules with distinct domains that ensure the processing of the piRNA precursors. The Tej also regulations the mobility of the nuage components *in vitro* through its disordered region, indicating the possibility of Tej contributing to the dynamic of nuage granules in a potentially engaging in the liquid-liquid phase separation. Taken together, Tej maintains the *Drosophila* piRNA biogenesis machinery in the germline cells as a multifunctional and core organizer of nuage (Figure 4).

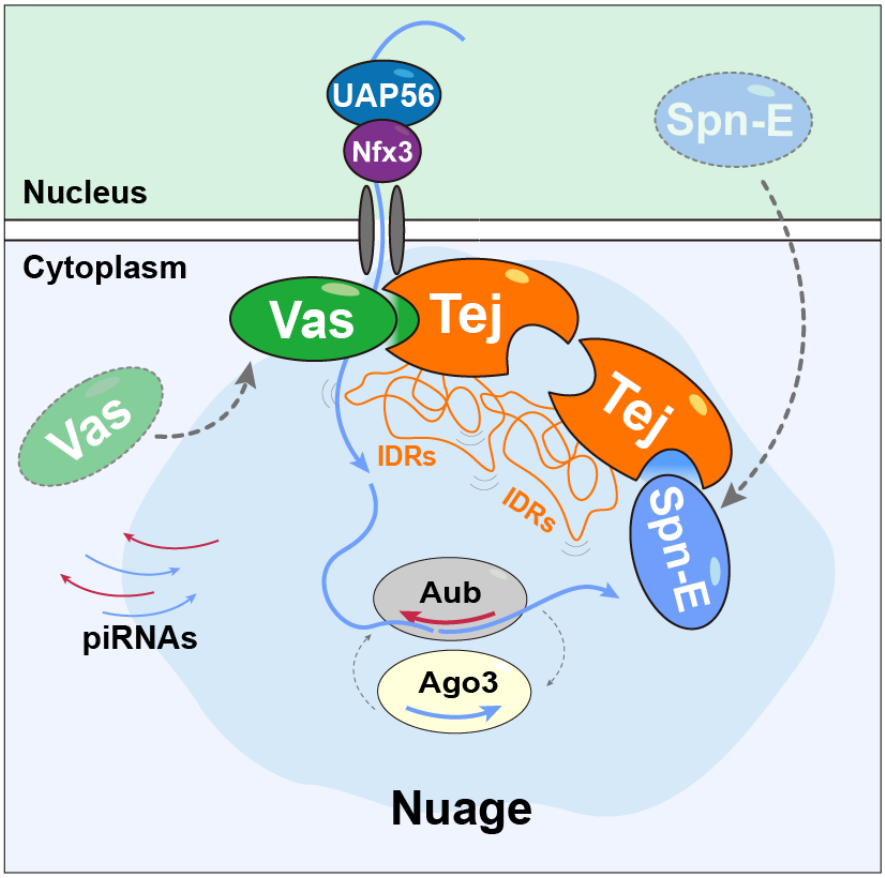


Figure 4. Tej is a multifunctional organizer of nuage.

Tej recruit cytoplasmic diffusing Vas and intrinsically nuclear-localized Spn-E to the perinuclear through the Lotus domain and novel identified SRS motif, respectively. The proper organized nuage guaranteed the further processing of precursors. The Tej also serves the molecular kinetic of the nuage components *in vitro* by utilizing its disordered region, indicating the possibility of Tej contributing to the dynamic of nuage granules.

5. REFERENCE

Abad, José P., Beatriz de Pablos, Kazutoyo Osoegawa, Pieter J. de Jong, Antonia Martín-Gallardo, and Alfredo Villasante. 2004. 'Genomic Analysis of Drosophila melanogaster Telomeres: Full-length Copies of HeT-A and TART Elements at Telomeres', *Molecular Biology and Evolution*, 21: 1613-19.

Anand, A., and T. Kai. 2012. 'The tudor domain protein kumo is required to assemble the nuage and to generate germline piRNAs in Drosophila', *EMBO J*, 31: 870-82.

Anantharaman, Vivek, Dapeng Zhang, and L. Aravind. 2010. 'OST-HTH: a novel predicted RNA-binding domain', *Biology Direct*, 5: 13.

Andersen, P. R., L. Tirian, M. Vunjak, and J. Brennecke. 2017. 'A heterochromatin-dependent transcription machinery drives piRNA expression', *Nature*, 549: 54-59.

Aravin, A. A., M. S. Klenov, V. V. Vagin, F. Bantignies, G. Cavalli, and V. A. Gvozdev. 2004. 'Dissection of a natural RNA silencing process in the Drosophila melanogaster germ line', *Mol Cell Biol*, 24: 6742-50.

Aravin, A. A., N. M. Naumova, A. V. Tulin, V. V. Vagin, Y. M. Rozovsky, and V. A. Gvozdev. 2001. 'Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the D. melanogaster germline', *Curr Biol*, 11: 1017-27.

Aravin, A., D. Gaidatzis, S. Pfeffer, M. Lagos-Quintana, P. Landgraf, N. Iovino, P. Morris, M. J. Brownstein, S. Kuramochi-Miyagawa, T. Nakano, M. Chien, J. J. Russo, J. Ju, R. Sheridan, C. Sander, M. Zavolan, and T. Tuschl. 2006. 'A novel class of small RNAs bind to MILI protein in mouse testes', *Nature*, 442: 203-7.

Aravin, Alexei A., Ravi Sachidanandam, Angelique Girard, Katalin Fejes-Toth, and Gregory J. Hannon. 2007. 'Developmentally Regulated piRNA Clusters Implicate MILI in Transposon Control', *Science*, 316: 744-47.

Blackburn, E. H. 1992. 'TELOMERASES', *Annual Review of Biochemistry*, 61: 113-29.

Boeynaems, S., S. Alberti, N. L. Fawzi, T. Mittag, M. Polymenidou, F. Rousseau, J. Schymkowitz, J. Shorter, B. Wolozin, L. Van Den Bosch, P. Tompa, and M. Fuxreiter. 2018. 'Protein Phase Separation: A New Phase in Cell Biology', *Trends Cell Biol*, 28: 420-35.

Boland, Andreas, Eric Huntzinger, Steffen Schmidt, Elisa Izaurralde, and Oliver Weichenrieder. 2011. 'Crystal structure of the MID-PIWI lobe of a eukaryotic Argonaute protein', *Proceedings of the National Academy of Sciences*, 108: 10466-71.

Boswell, Robert E, and Anthony P Mahowald. 1985. 'tudor, a gene required for assembly of the germ plasm in Drosophila melanogaster', *Cell*, 43: 97-104.

Brahms, Hero, Lydie Meheus, Veronique De Brabandere, U. T. Z. Fischer, and Reinhard LÜHrmann. 2002. 'Symmetrical dimethylation of arginine residues in spliceosomal Sm protein B/B′ and the Sm-like protein LSm4, and their interaction with the SMN protein', *RNA (New York, N.Y.)*, 7: 1531-42.

Brangwynne, Clifford P, Christian R Eckmann, David S Courson, Agata Rybarska, Carsten Hoege, Jöbin Gharakhani, Frank Jülicher, and Anthony A Hyman. 2009. 'Germline P granules are liquid droplets that localize by controlled dissolution/condensation', *Science*, 324: 1729-32.

Brangwynne, Clifford P, Timothy J Mitchison, and Anthony A Hyman. 2011. 'Active liquid-like behavior of nucleoli determines their size and shape in Xenopus laevis oocytes', *Proceedings of the National Academy of Sciences*, 108: 4334-39.

Brasset, E., A. R. Taddei, F. Arnaud, B. Faye, A. M. Fausto, M. Mazzini, F. Giorgi, and C. Vaury. 2006. 'Viral particles of the endogenous retrovirus ZAM from Drosophila melanogasteruse a pre-existing endosome/exosome pathway for transfer to the oocyte', *Retrovirology*, 3: 25.

Brennecke, J., A. A. Aravin, A. Stark, M. Dus, M. Kellis, R. Sachidanandam, and G. J. Hannon. 2007. 'Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila', *Cell*, 128: 1089-103.

Brennecke, J., C. D. Malone, A. A. Aravin, R. Sachidanandam, A. Stark, and G. J. Hannon. 2008. 'An epigenetic role for maternally inherited piRNAs in transposon silencing', *Science*, 322: 1387-92.

Brown, P. O., B. Bowerman, H. E. Varmus, and J. M. Bishop. 1987. 'Correct integration of retroviral DNA in vitro', *Cell*, 49: 347-56.

Carmell, M. A., Z. Xuan, M. Q. Zhang, and G. J. Hannon. 2002. 'The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis', *Genes Dev*, 16: 2733-42.

Cenik, E. S., and P. D. Zamore. 2011. 'Argonaute proteins', *Curr Biol*, 21: R446-9.

Chalvet, F., L. Teysset, C. Terzian, N. Prud'homme, P. Santamaria, A. Bucheton, and A. Pélisson. 1999. 'Proviral amplification of the Gypsy endogenous retrovirus of Drosophila melanogaster involves env-independent invasion of the female germline', *EMBO J*, 18: 2659-69.

Chen, Y. A., E. Stuwe, Y. Luo, M. Ninova, A. Le Thomas, E. Rozhavskaya, S. Li, S. Vempati, J. D. Laver, D. J. Patel, C. A. Smibert, H. D. Lipshitz, K. F. Toth, and A. A. Aravin. 2016. 'Cutoff Suppresses RNA Polymerase II Termination to Ensure Expression of piRNA Precursors', *Mol Cell*, 63: 97-109.

Choi, H. M. T., M. Schwarzkopf, M. E. Fornace, A. Acharya, G. Artavanis, J. Stegmaier, A. Cunha, and N. A. Pierce. 2018. 'Third-generation in situ hybridization chain reaction: multiplexed, quantitative, sensitive, versatile, robust', *Development*, 145.

Cipriani, P. G., O. Bay, J. Zinno, M. Gutwein, H. H. Gan, V. K. Mayya, G. Chung, J. X. Chen, H. Fahs, Y. Guan, T. F. Duchaine, M. Selbach, F. Piano, and K. C. Gunsalus. 2021. 'Novel LOTUS-domain proteins are organizational hubs that recruit C. elegans Vasa to germ granules', *Elife*, 10.

Côté, J., and S. Richard. 2005. 'Tudor domains bind symmetrical dimethylated arginines', *J Biol Chem*, 280: 28476-83.

Courchaine, Edward M., Andrew E. S. Barentine, Korinna Straube, Dong-Ryoung Lee, Joerg Bewersdorf, and Karla M. Neugebauer. 2021. 'DMA-tudor interaction modules control the specificity of in vivo condensates', *Cell*, 184: 3612-25.e17.

Cox, D. N., A. Chao, J. Baker, L. Chang, D. Qiao, and H. Lin. 1998. 'A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal', *Genes Dev*, 12: 3715-27.

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

Czech, B., J. B. Preall, J. McGinn, and G. J. Hannon. 2013. 'A transcriptome-wide RNAi screen in the Drosophila ovary reveals factors of the germline piRNA pathway', *Mol Cell*, 50: 749-61.

Czech, Benjamin, and Gregory J Hannon. 2016. 'A Happy 3′ Ending to the piRNA Maturation Story', *Cell*, 164: 838-40.

Dennis, Cynthia, Emilie Brasset, Arpita Sarkar, and Chantal Vaury. 2016. 'Export of piRNA precursors by EJC triggers assembly of cytoplasmic Yb-body in Drosophila', *Nature Communications*, 7: 13739.

Desset, S., C. Meignin, B. Dastugue, and C. Vaury. 2003. 'COM, a heterochromatic locus governing the control of independent endogenous retroviruses from Drosophila melanogaster', *Genetics*, 164: 501-9.

Ding, D., J. Liu, U. Midic, Y. Wu, K. Dong, A. Melnick, K. E. Latham, and C. Chen. 2018. 'TDRD5 binds piRNA precursors and selectively enhances pachytene piRNA processing in mice', *Nat Commun*, 9: 127.

Ding, Deqiang, Jiali Liu, Kunzhe Dong, Uros Midic, Rex A. Hess, Huirong Xie, Elena Y. Demireva, and Chen Chen. 2017. 'PNLDC1 is essential for piRNA 3′ end trimming and transposon silencing during spermatogenesis in mice', *Nature Communications*, 8: 819.

Ding, Deqiang, Chao Wei, Kunzhe Dong, Jiali Liu, Alexander Stanton, Chao Xu, Jinrong Min, Jian Hu, and Chen Chen. 2020. 'LOTUS domain is a novel class of G-rich and G-quadruplex RNA binding domain', *Nucleic Acids Research*, 48: 9262-72.

Dombroski, B. A., Q. Feng, S. L. Mathias, D. M. Sassaman, A. F. Scott, H. H. Kazazian, Jr., and J. D. Boeke. 1994. 'An in vivo assay for the reverse transcriptase of human retrotransposon L1 in Saccharomyces cerevisiae', *Mol Cell Biol*, 14: 4485-92.

Dosztányi, Z., V. Csizmok, P. Tompa, and I. Simon. 2005. 'IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content', *Bioinformatics*, 21: 3433-4.

Dunker, A. Keith, J. David Lawson, Celeste J. Brown, Ryan M. Williams, Pedro Romero, Jeong S. Oh, Christopher J. Oldfield, Andrew M. Campen, Catherine M. Ratliff, Kerry W. Hipps, Juan Ausio, Mark S. Nissen, Raymond Reeves, ChulHee Kang, Charles R. Kissinger, Robert W. Bailey, Michael D. Griswold, Wah Chiu, Ethan C. Garner, and Zoran Obradovic. 2001. 'Intrinsically disordered protein', *Journal of Molecular Graphics and Modelling*, 19: 26-59.

Elkayam, E., C. D. Kuhn, A. Tocilj, A. D. Haase, E. M. Greene, G. J. Hannon, and L. Joshua-Tor. 2012. 'The structure of human argonaute-2 in complex with miR-20a', *Cell*, 150: 100-10.

ElMaghraby, M. F., P. R. Andersen, F. Pühringer, U. Hohmann, K. Meixner, T. Lendl, L. Tirian, and J. Brennecke. 2019. 'A Heterochromatin-Specific RNA Export Pathway Facilitates piRNA Production', *Cell*, 178: 964-79.e20.

Feltzin, V. L., M. Khaladkar, M. Abe, M. Parisi, G. J. Hendriks, J. Kim, and N. M. Bonini. 2015. 'The exonuclease Nibbler regulates age-associated traits and modulates piRNA length in Drosophila', *Aging Cell*, 14: 443-52.

Friesen, Westley J., Severine Massenet, Sergey Paushkin, Anastasia Wyce, and Gideon Dreyfuss. 2001. 'SMN, the Product of the Spinal Muscular Atrophy Gene, Binds Preferentially to Dimethylarginine-Containing Protein Targets', *Mol Cell*, 7: 1111-17.

Gillespie, D. E., and C. A. Berg. 1995. 'Homeless is required for RNA localization in Drosophila oogenesis and encodes a new member of the DE-H family of RNA-dependent ATPases', *Genes Dev*, 9: 2495-508.

Girard, Angélique, Ravi Sachidanandam, Gregory J. Hannon, and Michelle A. Carmell. 2006. 'A germline-specific class of small RNAs binds mammalian Piwi proteins', *Nature*, 442: 199-202.

Gleason, Ryan J., Amit Anand, Toshie Kai, and Xin Chen. 2018. 'Protecting and Diversifying the Germline', *Genetics*, 208: 435-71.

Gokcezade, J., G. Sienski, and P. Duchek. 2014. 'Efficient CRISPR/Cas9 plasmids for rapid and versatile genome editing in Drosophila', *G3 (Bethesda)*, 4: 2279-82.

Golumbeski, G. S., A. Bardsley, F. Tax, and R. E. Boswell. 1991. 'tudor, a posterior-group gene of Drosophila melanogaster, encodes a novel protein and an mRNA localized during mid-oogenesis', *Genes Dev*, 5: 2060-70.

Goriaux, C., S. Desset, Y. Renaud, C. Vaury, and E. Brasset. 2014. 'Transcriptional properties and splicing of the flamenco piRNA cluster', *EMBO Rep*, 15: 411-8.

Grimson, Andrew, Mansi Srivastava, Bryony Fahey, Ben J. Woodcroft, H. Rosaria Chiang, Nicole King, Bernard M. Degnan, Daniel S. Rokhsar, and David P. Bartel. 2008. 'Early origins and evolution of microRNAs and Piwi-interacting RNAs in animals', *Nature*, 455: 1193-97.

Grivna, S. T., E. Beyret, Z. Wang, and H. Lin. 2006. 'A novel class of small RNAs in mouse spermatogenic cells', *Genes Dev*, 20: 1709-14.

Gunawardane, L. S., K. Saito, K. M. Nishida, K. Miyoshi, Y. Kawamura, T. Nagami, H. Siomi, and M. C. Siomi. 2007. 'A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in Drosophila', *Science*, 315: 1587-90.

Han, B. W., J. H. Hung, Z. Weng, P. D. Zamore, and S. L. Ameres. 2011. 'The 3'-to-5' exoribonuclease Nibbler shapes the 3' ends of microRNAs bound to Drosophila Argonaute1', *Curr Biol*, 21: 1878-87.

Han, B. W., W. Wang, C. Li, Z. Weng, and P. D. Zamore. 2015. 'Noncoding RNA. piRNA-guided transposon cleavage initiates Zucchini-dependent, phased piRNA production', *Science*, 348: 817-21.

Handler, D., K. Meixner, M. Pizka, K. Lauss, C. Schmied, F. S. Gruber, and J. Brennecke. 2013. 'The genetic makeup of the Drosophila piRNA pathway', *Mol Cell*, 50: 762-77.

Handler, D., D. Olivieri, M. Novatchkova, F. S. Gruber, K. Meixner, K. Mechtler, A. Stark, R. Sachidanandam, and J. Brennecke. 2011. 'A systematic analysis of Drosophila TUDOR domain-containing proteins identifies Vreteno and the Tdrd12 family as essential primary piRNA pathway factors', *EMBO J*, 30: 3977-93.

Harris, Adam N., and Paul M. Macdonald. 2001. 'Aubergine encodes a Drosophila polar granule component required for pole cell formation and related to eIF2C', *Development*, 128 14: 2823-32.

Hayashi, R., J. Schnabl, D. Handler, F. Mohn, S. L. Ameres, and J. Brennecke. 2016. 'Genetic and mechanistic diversity of piRNA 3'-end formation', *Nature*, 539: 588-92.

Hedges, D. J., and V. P. Belancio. 2011. 'Restless genomes humans as a model organism for understanding host-retrotransposable element dynamics', *Adv Genet*, 73: 219-62.

Hedges, D. J., and P. L. Deininger. 2007. 'Inviting instability: Transposable elements, double-strand breaks, and the maintenance of genome integrity', *Mutat Res*, 616: 46-59.

Homolka, David, Radha Raman Pandey, Coline Goriaux, Emilie Brasset, Chantal Vaury, Ravi Sachidanandam, Marie-Odile Fauvarque, and Ramesh S Pillai. 2015. 'PIWI Slicing and RNA Elements in Precursors Instruct Directional Primary piRNA Biogenesis', *Cell Reports*, 12: 418-28.

Hondele, Maria, Ruchika Sachdev, Stephanie Heinrich, Juan Wang, Pascal Vallotton, Beatriz MA Fontoura, and Karsten Weis. 2019. 'DEAD-box ATPases are global regulators of phase-separated organelles', *Nature*, 573: 144-48.

Horwich, M. D., C. Li, C. Matranga, V. Vagin, G. Farley, P. Wang, and P. D. Zamore. 2007. 'The Drosophila RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC', *Curr Biol*, 17: 1265-72.

Houwing, Saskia, Leonie M. Kamminga, Eugene Berezikov, Daniela Cronembold, Angélique Girard, Hans van den Elst, Dmitri V. Filippov, Heiko Blaser, Erez Raz, Cecilia B. Moens, Ronald H. A. Plasterk, Gregory J. Hannon, Bruce W. Draper, and René F. Ketting. 2007. 'A Role for Piwi and piRNAs in Germ Cell Maintenance and Transposon Silencing in Zebrafish', *Cell*, 129: 69-82.

Huang, H., Y. Li, K. E. Szulwach, G. Zhang, P. Jin, and D. Chen. 2014. 'AGO3 Slicer activity regulates mitochondria-nuage localization of Armitage and piRNA amplification', *J Cell Biol*, 206: 217-30.

Huang, Xiawei, Katalin Fejes Tóth, and Alexei A. Aravin. 2017. 'piRNA Biogenesis in Drosophila melanogaster', *Trends in Genetics*, 33: 882-94.

Huang, Xiawei, Hongmiao Hu, Alexandre Webster, Fan Zou, Jiamu Du, Dinshaw J. Patel, Ravi Sachidanandam, Katalin Fejes Toth, Alexei A. Aravin, and Sisi Li. 2021. 'Binding of guide piRNA triggers methylation of the unstructured N-terminal region of Aub leading to assembly of the piRNA amplification complex', *Nature Communications*, 12: 4061.

Hur, J. K., Y. Luo, S. Moon, M. Ninova, G. K. Marinov, Y. D. Chung, and A. A. Aravin. 2016. 'Splicing-independent loading of TREX on nascent RNA is required for efficient expression of dual-strand piRNA clusters in Drosophila', *Genes Dev*, 30: 840-55.

Iki, Taichiro, Moe Takami, and Toshie Kai. 2020. 'Modulation of Ago2 Loading by Cyclophilin 40 Endows a Unique Repertoire of Functional miRNAs during Sperm Maturation in Drosophila', *Cell Reports*, 33: 108380.

Ipsaro, J. J., A. D. Haase, S. R. Knott, L. Joshua-Tor, and G. J. Hannon. 2012. 'The structural biochemistry of Zucchini implicates it as a nuclease in piRNA biogenesis', *Nature*, 491: 279-83.

Ishizu, H., Y. W. Iwasaki, S. Hirakata, H. Ozaki, W. Iwasaki, H. Siomi, and M. C. Siomi. 2015. 'Somatic Primary piRNA Biogenesis Driven by cis-Acting RNA Elements and trans-Acting Yb', *Cell Rep*, 12: 429-40.

Ishizu, Hirotsugu, Tatsuki Kinoshita, Shigeki Hirakata, Chihiro Komatsuzaki, and Mikiko C. Siomi. 2019. 'Distinct and Collaborative Functions of Yb and Armitage in Transposon-Targeting piRNA Biogenesis', *Cell Reports*, 27: 1822-35.e8.

Jeske, M., M. Bordi, S. Glatt, S. Müller, V. Rybin, C. W. Müller, and A. Ephrussi. 2015. 'The Crystal Structure of the Drosophila Germline Inducer Oskar Identifies Two Domains with Distinct Vasa Helicase- and RNA-Binding Activities', *Cell Rep*, 12: 587-98.

Jeske, M., C. W. Müller, and A. Ephrussi. 2017. 'The LOTUS domain is a conserved DEAD-box RNA helicase regulator essential for the recruitment of Vasa to the germ plasm and nuage', *Genes Dev*, 31: 939-52.

Kaminker, J. S., C. M. Bergman, B. Kronmiller, J. Carlson, R. Svirskas, S. Patel, E. Frise, D. A. Wheeler, S. E. Lewis, G. M. Rubin, M. Ashburner, and S. E. Celniker. 2002. 'The transposable elements of the Drosophila melanogaster euchromatin: a genomics perspective', *Genome Biol*, 3: Research0084.

Kamminga, L. M., M. J. Luteijn, M. J. den Broeder, S. Redl, L. J. Kaaij, E. F. Roovers, P. Ladurner, E. Berezikov, and R. F. Ketting. 2010. 'Hen1 is required for oocyte development and piRNA stability in zebrafish', *EMBO J*, 29: 3688-700.

Katahira, Jun. 2012. 'mRNA export and the TREX complex', *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, 1819: 507-13.

Kato, M., T. W. Han, S. Xie, K. Shi, X. Du, L. C. Wu, H. Mirzaei, E. J. Goldsmith, J. Longgood, J. Pei, N. V. Grishin, D. E. Frantz, J. W. Schneider, S. Chen, L. Li, M. R. Sawaya, D. Eisenberg, R. Tycko, and S. L. McKnight. 2012. 'Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels', *Cell*, 149: 753-67.

Kawaguchi, Shinichi, Mizuki Ueki, and Toshie Kai. 2020. 'Drosophila MARF1 ensures proper oocyte maturation by regulating nanos expression', *PLOS ONE*, 15: e0231114.

Khurana, J. S., J. Wang, J. Xu, B. S. Koppetsch, T. C. Thomson, A. Nowosielska, C. Li, P. D. Zamore, Z. Weng, and W. E. Theurkauf. 2011. 'Adaptation to P element transposon invasion in Drosophila melanogaster', *Cell*, 147: 1551-63.

Kidwell, M. G., J. F. Kidwell, and J. A. Sved. 1977. 'Hybrid Dysgenesis in DROSOPHILA MELANOGASTER: A Syndrome of Aberrant Traits Including Mutation, Sterility and Male Recombination', *Genetics*, 86: 813-33.

Kim, Jeesun, Jeremy Daniel, Alexsandra Espejo, Aimee Lake, Murli Krishna, Li Xia, Yi Zhang, and Mark T Bedford. 2006. 'Tudor, MBT and chromo domains gauge the degree of lysine methylation', *EMBO Rep*, 7: 397-403.

Kina, H., T. Yoshitani, K. Hanyu-Nakamura, and A. Nakamura. 2019. 'Rapid and efficient generation of GFP-knocked-in Drosophila by the CRISPR-Cas9-mediated genome editing', *Dev Growth Differ*, 61: 265-75.

King, O. D., A. D. Gitler, and J. Shorter. 2012. 'The tip of the iceberg: RNA-binding proteins with prion-like domains in neurodegenerative disease', *Brain Res*, 1462: 61-80.

Kirino, Y., and Z. Mourelatos. 2007. 'The mouse homolog of HEN1 is a potential methylase for Piwi-interacting RNAs', *RNA (New York, N.Y.)*, 13: 1397-401.

Kirino, Yohei, Namwoo Kim, Mariàngels de Planell-Saguer, Eugene Khandros, Stephanie Chiorean, Peter S Klein, Isidore Rigoutsos, Thomas A Jongens, and Zissimos Mourelatos. 2009. 'Arginine methylation of Piwi proteins catalysed by dPRMT5 is required for Ago3 and Aub stability', *Nature cell biology*, 11: 652-58.

Kirino, Yohei, Anastassios Vourekas, Nabil Sayed, Flavia de Lima Alves, Travis Thomson, Paul Lasko, Juri Rappsilber, Thomas A. Jongens, and Zissimos Mourelatos. 2010. 'Arginine methylation of Aubergine mediates Tudor binding and germ plasm localization', *RNA (New York, N.Y.)*, 16: 70-78.

Kistler, K. E., T. Trcek, T. R. Hurd, R. Chen, F. X. Liang, J. Sall, M. Kato, and R. Lehmann. 2018. 'Phase transitioned nuclear Oskar promotes cell division of Drosophila primordial germ cells', *Elife*, 7.

Klattenhoff, C., H. Xi, C. Li, S. Lee, J. Xu, J. S. Khurana, F. Zhang, N. Schultz, B. S. Koppetsch, A. Nowosielska, H. Seitz, P. D. Zamore, Z. Weng, and W. E. Theurkauf. 2009. 'The Drosophila HP1 homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters', *Cell*, 138: 1137-49.

Klenov, Mikhail S, Olesya A Sokolova, Evgeny Y Yakushev, Anastasia D Stolyarenko, Elena A Mikhaleva, Sergey A Lavrov, and Vladimir A Gvozdev. 2011. 'Separation of stem cell maintenance and transposon silencing functions of Piwi protein', *Proceedings of the National Academy of Sciences*, 108: 18760-65.

Kneuss, Emma, Marzia Munafò, Evelyn L. Eastwood, Undine-Sophie Deumer, Jonathan B. Preall, Gregory J. Hannon, and Benjamin Czech. 2019. 'Specialization of the Drosophila nuclear export family protein Nxf3 for piRNA precursor export', *Genes & development*, 33: 1208-20.

Kofler, Robert, Tom Hill, Viola Nolte, Andrea J. Betancourt, and Christian Schlötterer. 2015. 'The recent invasion of natural Drosophila simulans populations by the P-element', *Proceedings of the National Academy of Sciences*, 112: 6659.

Lambowitz, A. M., and S. Zimmerly. 2011. 'Group II introns: mobile ribozymes that invade DNA', *Cold Spring Harb Perspect Biol*, 3: a003616.

Langmead, Ben, Cole Trapnell, Mihai Pop, and Steven L. Salzberg. 2009. 'Ultrafast and memory-efficient alignment of short DNA sequences to the human genome', *Genome Biol*, 10: R25.

Lau, N. C., A. G. Seto, J. Kim, S. Kuramochi-Miyagawa, T. Nakano, D. P. Bartel, and R. E. Kingston. 2006. 'Characterization of the piRNA complex from rat testes', *Science*, 313: 363-7.

Leblanc, P., S. Desset, B. Dastugue, and C. Vaury. 1997. 'Invertebrate retroviruses: ZAM a new candidate in D.melanogaster', *EMBO J*, 16: 7521-31.

Leblanc, P., S. Desset, F. Giorgi, A. R. Taddei, A. M. Fausto, M. Mazzini, B. Dastugue, and C. Vaury. 2000. 'Life cycle of an endogenous retrovirus, ZAM, in Drosophila melanogaster', *J Virol*, 74: 10658-69.

Li, C., V. V. Vagin, S. Lee, J. Xu, S. Ma, H. Xi, H. Seitz, M. D. Horwich, M. Syrzycka, B. M. Honda, E. L. Kittler, M. L. Zapp, C. Klattenhoff, N. Schulz, W. E. Theurkauf, Z. Weng, and P. D. Zamore. 2009. 'Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies', *Cell*, 137: 509-21.

Li, P., S. Banjade, H. C. Cheng, S. Kim, B. Chen, L. Guo, M. Llaguno, J. V. Hollingsworth, D. S. King, S. F. Banani, P. S. Russo, Q. X. Jiang, B. T. Nixon, and M. K. Rosen. 2012. 'Phase transitions in the assembly of multivalent signalling proteins', *Nature*, 483: 336-40.

Liang, L., W. Diehl-Jones, and P. Lasko. 1994. 'Localization of vasa protein to the Drosophila pole plasm is independent of its RNA-binding and helicase activities', *Development*, 120: 1201-11.

Lim, Ai Khim, and Toshie Kai. 2007. 'Unique germ-line organelle, nuage, functions to repress selfish genetic elements in Drosophila melanogaster', *Proceedings of the National Academy of Sciences*, 104: 6714-19.

Lim, Robyn S. M., Amit Anand, Chiemi Nishimiya-Fujisawa, Satoru Kobayashi, and Toshie Kai. 2014. 'Analysis of Hydra PIWI proteins and piRNAs uncover early evolutionary origins of the piRNA pathway', *Developmental Biology*, 386: 237-51.

Lim, S. L., Z. P. Qu, R. D. Kortschak, D. M. Lawrence, J. Geoghegan, A. L. Hempfling, M. Bergmann, C. C. Goodnow, C. J. Ormandy, L. Wong, J. Mann, H. S. Scott, D. Jamsai, D. L. Adelson, and M. K. O'Bryan. 2015. 'HENMT1 and piRNA Stability Are Required for Adult Male Germ Cell Transposon Repression and to Define the Spermatogenic Program in the Mouse', *PLoS Genet*, 11: e1005620.

Lin, M. D., X. Jiao, D. Grima, S. F. Newbury, M. Kiledjian, and T. B. Chou. 2008. 'Drosophila processing bodies in oogenesis', *Dev Biol*, 322: 276-88.

Liu, N., M. Abe, L. R. Sabin, G. J. Hendriks, A. S. Naqvi, Z. Yu, S. Cherry, and N. M. Bonini. 2011. 'The exoribonuclease Nibbler controls 3' end processing of microRNAs in Drosophila', *Curr Biol*, 21: 1888-93.

Luan, D. D., M. H. Korman, J. L. Jakubczak, and T. H. Eickbush. 1993. 'Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition', *Cell*, 72: 595-605.

Ma, J. B., K. Ye, and D. J. Patel. 2004. 'Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain', *Nature*, 429: 318-22.

Malone, C. D., J. Brennecke, M. Dus, A. Stark, W. R. McCombie, R. Sachidanandam, and G. J. Hannon. 2009. 'Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary', *Cell*, 137: 522-35.

Malone, C. D., and G. J. Hannon. 2009. 'Small RNAs as guardians of the genome', *Cell*, 136: 656-68.

Marnik, Elisabeth A., Miguel Vasconcelos Almeida, Patricia G. Cipriani, George Chung, Edoardo Caspani, Emil Karaulanov, Falk Butter, Catherine S. Sharp, John Zinno, Hin Hark Gan, Fabio Piano, René F. Ketting, Kristin C. Gunsalus, and Dustin L. Updike. 2021. 'The Caenorhabditis elegans TDRD5/7-like protein, LOTR-1, interacts with the helicase ZNFX-1 to balance epigenetic signals in the germline', *bioRxiv*.

Masuda, S., R. Das, H. Cheng, E. Hurt, N. Dorman, and R. Reed. 2005. 'Recruitment of the human TREX complex to mRNA during splicing', *Genes Dev*, 19: 1512-7.

Matsumoto, N., H. Nishimasu, K. Sakakibara, K. M. Nishida, T. Hirano, R. Ishitani, H. Siomi, M. C. Siomi, and O. Nureki. 2016. 'Crystal Structure of Silkworm PIWI-Clade Argonaute Siwi Bound to piRNA', *Cell*, 167: 484-97.e9.

Meikar, O., M. Da Ros, H. Korhonen, and N. Kotaja. 2011. 'Chromatoid body and small RNAs in male germ cells', *Reproduction*, 142: 195-209.

Mejlumian, L., A. Pélisson, A. Bucheton, and C. Terzian. 2002. 'Comparative and functional studies of Drosophila species invasion by the gypsy endogenous retrovirus', *Genetics*, 160: 201-9.

Mendel, M., and R. S. Pillai. 2019. 'Nxf3: a middleman with the right connections for unspliced piRNA precursor export', *Genes Dev*, 33: 1095-97.

Mohn, F., D. Handler, and J. Brennecke. 2015. 'Noncoding RNA. piRNA-guided slicing specifies transcripts for Zucchini-dependent, phased piRNA biogenesis', *Science*, 348: 812-17.

Mohn, F., G. Sienski, D. Handler, and J. Brennecke. 2014. 'The rhino-deadlock-cutoff complex licenses noncanonical transcription of dual-strand piRNA clusters in Drosophila', *Cell*, 157: 1364-79.

Moon, S., M. Cassani, Y. A. Lin, L. Wang, K. Dou, and Z. Z. Zhang. 2018. 'A Robust Transposon-Endogenizing Response from Germline Stem Cells', *Dev Cell*, 47: 660-71.e3.

Munafò, M., V. Manelli, F. A. Falconio, A. Sawle, E. Kneuss, E. L. Eastwood, J. W. E. Seah, B. Czech, and G. J. Hannon. 2019. 'Daedalus and Gasz recruit Armitage to mitochondria, bringing piRNA precursors to the biogenesis machinery', *Genes Dev*, 33: 844-56.

Murota, Yukiko, Hirotsugu Ishizu, Shinichi Nakagawa, Yuka W Iwasaki, Shinsuke Shibata, Miharu K Kamatani, Kuniaki Saito, Hideyuki Okano, Haruhiko Siomi, and Mikiko C Siomi. 2014. 'Yb Integrates piRNA Intermediates and Processing Factors into Perinuclear Bodies to Enhance piRISC Assembly', *Cell Reports*, 8: 103-13.

Nishida, Kazumichi M, Yuka W Iwasaki, Yukiko Murota, Akihiro Nagao, Taro Mannen, Yumiko Kato, Haruhiko Siomi, and Mikiko C Siomi. 2015. 'Respective Functions of Two Distinct Siwi Complexes Assembled during PIWI-Interacting RNA Biogenesis in Bombyx Germ Cells', *Cell Reports*, 10: 193-203.

Nishimasu, H., H. Ishizu, K. Saito, S. Fukuhara, M. K. Kamatani, L. Bonnefond, N. Matsumoto, T. Nishizawa, K. Nakanaga, J. Aoki, R. Ishitani, H. Siomi, M. C. Siomi, and O. Nureki. 2012. 'Structure and function of Zucchini endoribonuclease in piRNA biogenesis', *Nature*, 491: 284-7.

Nishimura, T., I. Nagamori, T. Nakatani, N. Izumi, Y. Tomari, S. Kuramochi-Miyagawa, and T. Nakano. 2018. 'PNLDC1, mouse pre-piRNA Trimmer, is required for meiotic and post-meiotic male germ cell development', *EMBO Rep*, 19.

Nott, T. J., E. Petsalaki, P. Farber, D. Jervis, E. Fussner, A. Plochowietz, T. D. Craggs, D. P. Bazett-Jones, T. Pawson, J. D. Forman-Kay, and A. J. Baldwin. 2015. 'Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles', *Mol Cell*, 57: 936-47.

Olivieri, D., K. A. Senti, S. Subramanian, R. Sachidanandam, and J. Brennecke. 2012. 'The cochaperone shutdown defines a group of biogenesis factors essential for all piRNA populations in Drosophila', *Mol Cell*, 47: 954-69.

Olivieri, D., M. M. Sykora, R. Sachidanandam, K. Mechtler, and J. Brennecke. 2010. 'An in vivo RNAi assay identifies major genetic and cellular requirements for primary piRNA biogenesis in Drosophila', *EMBO J*, 29: 3301-17.

Orgel, L. E., and F. H. C. Crick. 1980. 'Selfish DNA: the ultimate parasite', *Nature*, 284: 604-07.

Ott, K. M., T. Nguyen, and C. Navarro. 2014. 'The DExH box helicase domain of spindle-E is necessary for retrotransposon silencing and axial patterning during Drosophila oogenesis', *G3 (Bethesda)*, 4: 2247-57.

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

Pak, C. W., M. Kosno, A. S. Holehouse, S. B. Padrick, A. Mittal, R. Ali, A. A. Yunus, D. R. Liu, R. V. Pappu, and M. K. Rosen. 2016. 'Sequence Determinants of Intracellular Phase Separation by Complex Coacervation of a Disordered Protein', *Mol Cell*, 63: 72-85.

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

Pane, A., K. Wehr, and T. Schüpbach. 2007. 'zucchini and squash encode two putative nucleases required for rasiRNA production in the Drosophila germline', *Dev Cell*, 12: 851-62.

Pardue, M. L., and P. Debaryshe. 2011. 'Adapting to life at the end of the line: How Drosophila telomeric retrotransposons cope with their job', *Mob Genet Elements*, 1: 128-34.

Pardue, M. L., and P. G. DeBaryshe. 2003. 'Retrotransposons provide an evolutionarily robust non-telomerase mechanism to maintain telomeres', *Annu Rev Genet*, 37: 485-511.

Parker, J. S., S. M. Roe, and D. Barford. 2004. 'Crystal structure of a PIWI protein suggests mechanisms for siRNA recognition and slicer activity', *EMBO J*, 23: 4727-37.

———. 2005. 'Structural insights into mRNA recognition from a PIWI domain-siRNA guide complex', *Nature*, 434: 663-6.

Patil, V. S., A. Anand, A. Chakrabarti, and T. Kai. 2014. 'The Tudor domain protein Tapas, a homolog of the vertebrate Tdrd7, functions in the piRNA pathway to regulate retrotransposons in germline of Drosophila melanogaster', *BMC Biol*, 12: 61.

Patil, Veena S., and Toshie Kai. 2010. 'Repression of Retroelements in Drosophila Germline via piRNA Pathway by the Tudor Domain Protein Tejas', *Current Biology*, 20: 724-30.

Pek, J. W., V. S. Patil, and T. Kai. 2012. 'piRNA pathway and the potential processing site, the nuage, in the Drosophila germline', *Dev Growth Differ*, 54: 66-77.

Pepling, Melissa E., Margaret de Cuevas, and Allan C. Spradling. 1999. 'Germline cysts: a conserved phase of germ cell development?', *Trends in Cell Biology*, 9: 257-62.

Piñeyro, David, Elisenda López-Panadès, María Lucena-Pérez, and Elena Casacuberta. 2011. 'Transcriptional analysis of the HeT-A retrotransposon in mutant and wild type stocks reveals high sequence variability at Drosophila telomeres and other unusual features', *BMC genomics*, 12: 1-14.

Ponting, Christopher P. 1997. 'Tudor domains in proteins that interact with RNA', *Trends in biochemical sciences*, 22: 51-52.

Preall, J. B., B. Czech, P. M. Guzzardo, F. Muerdter, and G. J. Hannon. 2012. 'shutdown is a component of the Drosophila piRNA biogenesis machinery', *RNA (New York, N.Y.)*, 18: 1446-57.

Price, I. F., H. L. Hertz, B. Pastore, J. Wagner, and W. Tang. 2021. 'Proximity labeling identifies LOTUS domain proteins that promote the formation of perinuclear germ granules in C. elegans', *Elife*, 10.

Prud'homme, N., M. Gans, M. Masson, C. Terzian, and A. Bucheton. 1995. 'Flamenco, a gene controlling the gypsy retrovirus of Drosophila melanogaster', *Genetics*, 139: 697-711.

Ramírez, Fidel, Vivek Bhardwaj, Laura Arrigoni, Kin Chung Lam, Björn A. Grüning, José Villaveces, Bianca Habermann, Asifa Akhtar, and Thomas Manke. 2018. 'High-resolution TADs reveal DNA sequences underlying genome organization in flies', *Nature Communications*, 9: 189.

Rangan, P., C. D. Malone, C. Navarro, S. P. Newbold, P. S. Hayes, R. Sachidanandam, G. J. Hannon, and R. Lehmann. 2011. 'piRNA production requires heterochromatin formation in Drosophila', *Curr Biol*, 21: 1373-9.

Rapsomaniki, Maria Anna, Panagiotis Kotsantis, Ioanna-Eleni Symeonidou, Nickolaos-Nikiforos Giakoumakis, Stavros Taraviras, and Zoi Lygerou. 2012. 'easyFRAP: an interactive, easy-to-use tool for qualitative and quantitative analysis of FRAP data', *Bioinformatics*, 28: 1800-01.

Robine, N., N. C. Lau, S. Balla, Z. Jin, K. Okamura, S. Kuramochi-Miyagawa, M. D. Blower, and E. C. Lai. 2009. 'A broadly conserved pathway generates 3'UTR-directed primary piRNAs', *Curr Biol*, 19: 2066-76.

Rozhkov, N. V., N. G. Schostak, E. S. Zelentsova, I. A. Yushenova, O. G. Zatsepina, and M. B. Evgen'ev. 2013. 'Evolution and dynamics of small RNA response to a retroelement invasion in Drosophila', *Mol Biol Evol*, 30: 397-408.

Rubin, G. M., M. G. Kidwell, and P. M. Bingham. 1982. 'The molecular basis of P-M hybrid dysgenesis: the nature of induced mutations', *Cell*, 29: 987-94.

Saito, K., H. Ishizu, M. Komai, H. Kotani, Y. Kawamura, K. M. Nishida, H. Siomi, and M. C. Siomi. 2010. 'Roles for the Yb body components Armitage and Yb in primary piRNA biogenesis in Drosophila', *Genes Dev*, 24: 2493-8.

Saito, K., K. M. Nishida, T. Mori, Y. Kawamura, K. Miyoshi, T. Nagami, H. Siomi, and M. C. Siomi. 2006. 'Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the Drosophila genome', *Genes Dev*, 20: 2214-22.

Saito, K., Y. Sakaguchi, T. Suzuki, T. Suzuki, H. Siomi, and M. C. Siomi. 2007. 'Pimet, the Drosophila homolog of HEN1, mediates 2'-O-methylation of Piwi- interacting RNAs at their 3' ends', *Genes Dev*, 21: 1603-8.

Sankaranarayanan, M., and T. T. Weil. 2020. 'Granule regulation by phase separation during Drosophila oogenesis', *Emerg Top Life Sci*, 4: 343-52.

Sato, Kaoru, Yuka W Iwasaki, Aoi Shibuya, Piero Carninci, Yuuta Tsuchizawa, Hirotsugu Ishizu, Mikiko C Siomi, and Haruhiko Siomi. 2015. 'Krimper Enforces an Antisense Bias on piRNA Pools by Binding AGO3 in the Drosophila Germline', *Mol Cell*, 59: 553-63.

Savitsky, Mikhail, Dmitry Kwon, Pavel Georgiev, Alla Kalmykova, and Vladimir Gvozdev. 2006. 'Telomere elongation is under the control of the RNAi-based mechanism in the Drosophila germline', *Genes & development*, 20: 345-54.

Schirle, N. T., and I. J. MacRae. 2012. 'The crystal structure of human Argonaute2', *Science*, 336: 1037-40.

Schwarz, D. S., Y. Tomari, and P. D. Zamore. 2004. 'The RNA-induced silencing complex is a Mg2+-dependent endonuclease', *Curr Biol*, 14: 787-91.

Sengoku, Toru, Osamu Nureki, Akira Nakamura, Satoru Kobayashi, and Shigeyuki Yokoyama. 2006. 'Structural Basis for RNA Unwinding by the DEAD-Box Protein Drosophila Vasa', *Cell*, 125: 287-300.

Shoji, M., T. Tanaka, M. Hosokawa, M. Reuter, A. Stark, Y. Kato, G. Kondoh, K. Okawa, T. Chujo, T. Suzuki, K. Hata, S. L. Martin, T. Noce, S. Kuramochi-Miyagawa, T. Nakano, H. Sasaki, R. S. Pillai, N. Nakatsuji, and S. Chuma. 2009. 'The TDRD9-MIWI2 complex is essential for piRNA-mediated retrotransposon silencing in the mouse male germline', *Dev Cell*, 17: 775-87.

Slaidina, M., T. U. Banisch, S. Gupta, and R. Lehmann. 2020. 'A single-cell atlas of the developing Drosophila ovary identifies follicle stem cell progenitors', *Genes Dev*, 34: 239-49.

Snee, M. J., and P. M. Macdonald. 2004. 'Live imaging of nuage and polar granules: evidence against a precursor-product relationship and a novel role for Oskar in stabilization of polar granule components', *J Cell Sci*, 117: 2109-20.

———. 2009. 'Dynamic organization and plasticity of sponge bodies', *Dev Dyn*, 238: 918-30.

Song, J. J., S. K. Smith, G. J. Hannon, and L. Joshua-Tor. 2004. 'Crystal structure of Argonaute and its implications for RISC slicer activity', *Science*, 305: 1434-7.

Strässer, K., S. Masuda, P. Mason, J. Pfannstiel, M. Oppizzi, S. Rodriguez-Navarro, A. G. Rondón, A. Aguilera, K. Struhl, R. Reed, and E. Hurt. 2002. 'TREX is a conserved complex coupling transcription with messenger RNA export', *Nature*, 417: 304-8.

Styhler, S., A. Nakamura, A. Swan, B. Suter, and P. Lasko. 1998. 'vasa is required for GURKEN accumulation in the oocyte, and is involved in oocyte differentiation and germline cyst development', *Development*, 125: 1569-78.

Tang, W., S. Tu, H. C. Lee, Z. Weng, and C. C. Mello. 2016. 'The RNase PARN-1 Trims piRNA 3' Ends to Promote Transcriptome Surveillance in C. elegans', *Cell*, 164: 974-84.

Tian, Y., D. K. Simanshu, J. B. Ma, and D. J. Patel. 2011. 'Structural basis for piRNA 2'-O-methylated 3'-end recognition by Piwi PAZ (Piwi/Argonaute/Zwille) domains', *Proc Natl Acad Sci U S A*, 108: 903-10.

Vagin, V. V., A. Sigova, C. Li, H. Seitz, V. Gvozdev, and P. D. Zamore. 2006. 'A distinct small RNA pathway silences selfish genetic elements in the germline', *Science*, 313: 320-4.

Viphakone, Nicolas, Guillaume M. Hautbergue, Matthew Walsh, Chung-Te Chang, Arthur Holland, Eric G. Folco, Robin Reed, and Stuart A. Wilson. 2012. 'TREX exposes the RNA-binding domain of Nxf1 to enable mRNA export', *Nature Communications*, 3: 1006.

Wang, W., B. W. Han, C. Tipping, D. T. Ge, Z. Zhang, Z. Weng, and P. D. Zamore. 2015. 'Slicing and Binding by Ago3 or Aub Trigger Piwi-Bound piRNA Production by Distinct Mechanisms', *Mol Cell*, 59: 819-30.

Wang, W., M. Yoshikawa, B. W. Han, N. Izumi, Y. Tomari, Z. Weng, and P. D. Zamore. 2014. 'The initial uridine of primary piRNAs does not create the tenth adenine that Is the hallmark of secondary piRNAs', *Mol Cell*, 56: 708-16.

Webster, A., S. Li, J. K. Hur, M. Wachsmuth, J. S. Bois, E. M. Perkins, D. J. Patel, and A. A. Aravin. 2015. 'Aub and Ago3 Are Recruited to Nuage through Two Mechanisms to Form a Ping-Pong Complex Assembled by Krimper', *Mol Cell*, 59: 564-75.

Wei, K. F., L. J. Wu, J. Chen, Y. F. Chen, and D. X. Xie. 2012. 'Structural evolution and functional diversification analyses of argonaute protein', *J Cell Biochem*, 113: 2576-85.

Weil, T. T., R. M. Parton, B. Herpers, J. Soetaert, T. Veenendaal, D. Xanthakis, I. M. Dobbie, J. M. Halstead, R. Hayashi, C. Rabouille, and I. Davis. 2012. 'Drosophila patterning is established by differential association of mRNAs with P bodies', *Nat Cell Biol*, 14: 1305-13.

Wells, Jonathan N., and Cédric Feschotte. 2020. 'A Field Guide to Eukaryotic Transposable Elements', *Annual Review of Genetics*, 54: 539-61.

Wenda, J. M., D. Homolka, Z. Yang, P. Spinelli, R. Sachidanandam, R. R. Pandey, and R. S. Pillai. 2017. 'Distinct Roles of RNA Helicases MVH and TDRD9 in PIWI Slicing-Triggered Mammalian piRNA Biogenesis and Function', *Dev Cell*, 41: 623-37.e9.

Wilsch-Bräuninger, M., H. Schwarz, and C. Nüsslein-Volhard. 1997. 'A sponge-like structure involved in the association and transport of maternal products during Drosophila oogenesis', *J Cell Biol*, 139: 817-29.

Xiol, Jordi, Pietro Spinelli, Maike A Laussmann, David Homolka, Zhaolin Yang, Elisa Cora, Yohann Couté, Simon Conn, Jan Kadlec, Ravi Sachidanandam, Marko Kaksonen, Stephen Cusack,

Anne Ephrussi, and Ramesh S Pillai. 2014. 'RNA Clamping by Vasa Assembles a piRNA Amplifier Complex on Transposon Transcripts', *Cell*, 157: 1698-711.

Yamashiro, H., M. Negishi, T. Kinoshita, H. Ishizu, H. Ohtani, and M. C. Siomi. 2020. 'Armitage determines Piwi-piRISC processing from precursor formation and quality control to inter-organelle translocation', *EMBO Rep*, 21: e48769.

Yamashiro, H., and M. C. Siomi. 2018. 'PIWI-Interacting RNA in Drosophila: Biogenesis, Transposon Regulation, and Beyond', *Chem Rev*, 118: 4404-21.

Yang, Jin, Hal P Bogerd, P Jeremy Wang, David C Page, and Bryan R Cullen. 2001. 'Two closely related human nuclear export factors utilize entirely distinct export pathways', *Mol Cell*, 8: 397-406.

Yuan, Y. R., Y. Pei, J. B. Ma, V. Kuryavyi, M. Zhadina, G. Meister, H. Y. Chen, Z. Dauter, T. Tuschl, and D. J. Patel. 2005. 'Crystal structure of A. aeolicus argonaute, a site-specific DNA-guided endoribonuclease, provides insights into RISC-mediated mRNA cleavage', *Mol Cell*, 19: 405-19.

Zamparini, A. L., M. Y. Davis, C. D. Malone, E. Vieira, J. Zavadil, R. Sachidanandam, G. J. Hannon, and R. Lehmann. 2011. 'Vreteno, a gonad-specific protein, is essential for germline development and primary piRNA biogenesis in Drosophila', *Development*, 138: 4039-50.

Zhang, F., J. Wang, J. Xu, Z. Zhang, B. S. Koppetsch, N. Schultz, T. Vreven, C. Meignin, I. Davis, P. D. Zamore, Z. Weng, and W. E. Theurkauf. 2012. 'UAP56 couples piRNA clusters to the perinuclear transposon silencing machinery', *Cell*, 151: 871-84.

Zhang, Z., J. Wang, N. Schultz, F. Zhang, S. S. Parhad, S. Tu, T. Vreven, P. D. Zamore, Z. Weng, and W. E. Theurkauf. 2014. 'The HP1 homolog rhino anchors a nuclear complex that suppresses piRNA precursor splicing', *Cell*, 157: 1353-63.

Zhu, Li, Suresh K. Kandasamy, Susan E. Liao, and Ryuya Fukunaga. 2018. 'LOTUS domain protein MARF1 binds CCR4-NOT deadenylase complex to post-transcriptionally regulate gene expression in oocytes', *Nature Communications*, 9: 4031.

PUBLICATION AND CONTRIBUTION

Publication

Lin Yuxuan, Kawaguchi Shinichi, Suyama Ritsuko, Iki Taichiro, Kai Toshie. Tejas Functions in piRNA Biogenesis Via Nuage Assembly in *Drosophila*. (Preparing for Submission)

Conference Presentation

Molecular basis of Nuage organization by Tejas in Drosophila piRNA biogenesis. (Poster)
The 44th Annual Meeting of the Molecular Biology Society of Japan. December 1 - 3, 2021.