

REVIEW

Open Access



Roles of piRNAs in transposon and pseudogene regulation of germline mRNAs and lncRNAs

Chen Wang¹ and Haifan Lin^{2*}

* Correspondence: haifan.lin@yale.edu

²Yale Stem Cell Center and Department of Cell Biology, Yale University School of Medicine, New Haven, CT 06519, USA
Full list of author information is available at the end of the article

Abstract

PIWI proteins, a subfamily of PAZ/PIWI Domain family RNA-binding proteins, are best known for their function in silencing transposons and germline development by partnering with small noncoding RNAs called PIWI-interacting RNAs (piRNAs). However, recent studies have revealed multifaceted roles of the PIWI-piRNA pathway in regulating the expression of other major classes of RNAs in germ cells. In this review, we summarize how PIWI proteins and piRNAs regulate the expression of many disparate RNAs, describing a highly complex global genomic regulatory relationship at the RNA level through which piRNAs functionally connect all major constituents of the genome in the germline.

Introduction

PIWI proteins represent one of the two subfamilies of the PAZ/PIWI Domain (PPD) protein family, with the other subfamily termed Ago proteins. All PPD family proteins contain a variable N-terminal domain followed by a highly conserved PAZ domain, which together with the MID domain binds a small RNA [1, 2] (Fig. 1a, b). The C-terminal PIWI domain resembles RNase H and in some PPD family proteins are capable of cleaving target RNAs [26–28]. Ago subfamily proteins are present in most types of cells. They bind to microRNAs (miRNAs) and small interfering RNAs (siRNAs), both of which are 21-nucleotide in length and are produced from double-stranded precursors via a Dicer-dependent process [29]. In contrast, PIWI subfamily proteins are mostly expressed in the germline, even though in most arthropod species, PIWI proteins are expressed in the soma as well [30–32]. PIWI proteins bind to piRNAs that are generally 24–32 nucleotides in length and are also enriched in the germline [33–36].

PiRNAs in most organisms are processed from long single-stranded precursors in a Dicer-independent manner (Fig. 1c), except for *C. elegans*, in which piRNAs (21 U-RNAs) are produced from short, single-stranded precursors (~ 26 nt capped transcripts) in a Dicer-independent manner [15, 37] (Fig. 1d). The long single-stranded precursors RNAs were transcribed from loci termed piRNA clusters in the genome



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

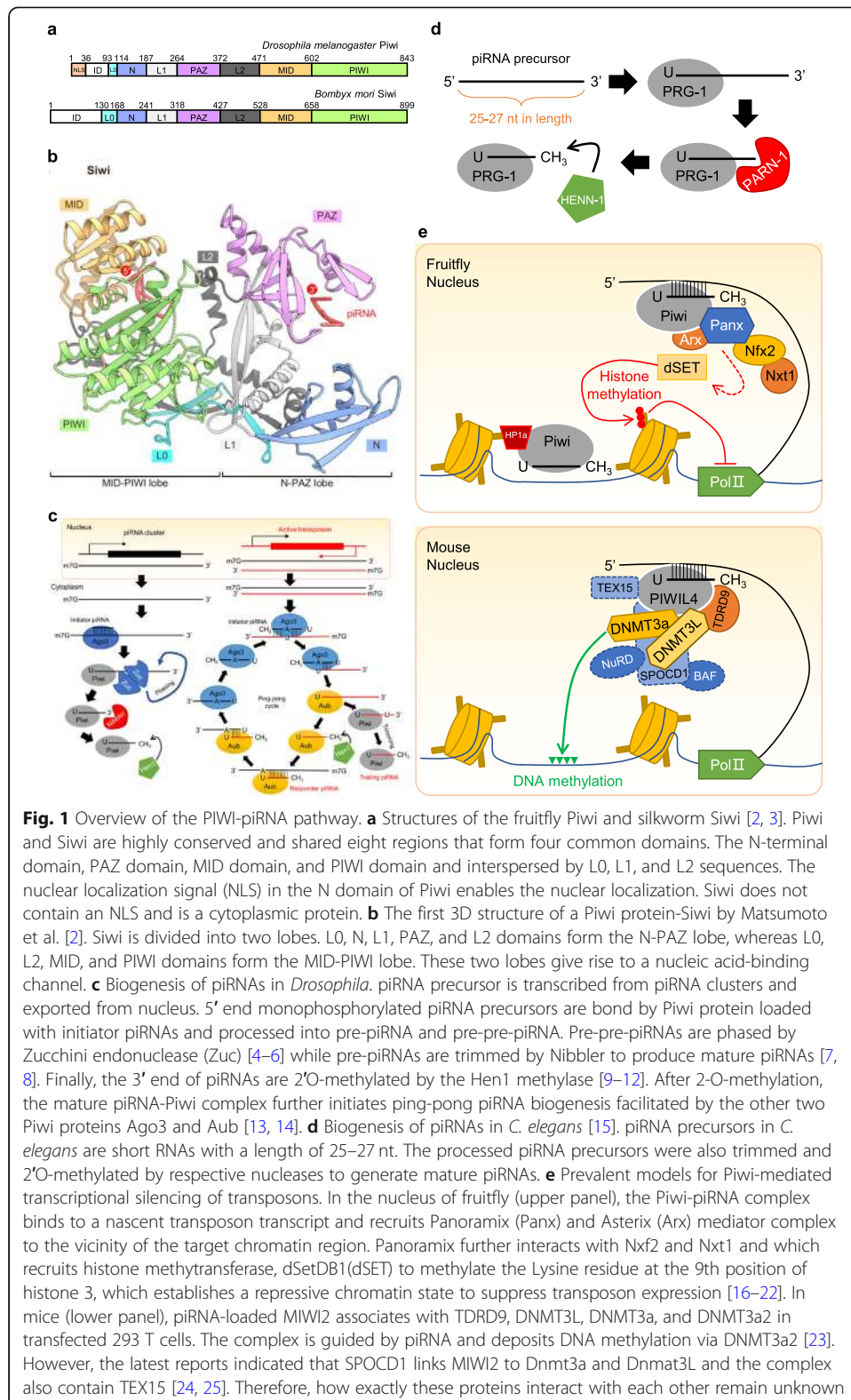


Fig. 1 Overview of the PIWI-piRNA pathway. **a** Structures of the fruitfly Piwi and silkworm Siwi [2, 3]. Piwi and Siwi are highly conserved and shared eight domains. The N-terminal domain, PAZ domain, MID domain, and PIWI domain and interspersed by L0, L1, and L2 sequences. The nuclear localization signal (NLS) in the N domain of Piwi enables the nuclear localization. Siwi does not contain an NLS and is a cytoplasmic protein. **b** The first 3D structure of a Piwi protein-Siwi by Matsumoto et al. [2]. Siwi is divided into two lobes. L0, N, L1, PAZ, and L2 domains form the N-PAZ lobe, whereas L0, L2, MID, and PIWI domains form the MID-PIWI lobe. These two lobes give rise to a nucleic acid-binding channel. **c** Biogenesis of piRNAs in *Drosophila*. piRNA precursor is transcribed from piRNA clusters and exported from nucleus. 5' end monophosphorylated piRNA precursors are bound by Piwi protein loaded with initiator piRNAs and processed into pre-piRNA and pre-pre-piRNA. Pre-pre-piRNAs are phased by Zucchini endonuclease (Zuc) [4–6] while pre-piRNAs are trimmed by Nibbler to produce mature piRNAs [7, 8]. Finally, the 3' end of piRNAs are 2'O-methylated by the Hen1 methylase [9–12]. After 2'O-methylation, the mature piRNA-Piwi complex further initiates ping-pong piRNA biogenesis facilitated by the other two Piwi proteins Ago3 and Aub [13, 14]. **d** Biogenesis of piRNAs in *C. elegans* [15]. piRNA precursors in *C. elegans* are short RNAs with a length of 25–27 nt. The processed piRNA precursors were also trimmed and 2'O-methylated by respective nucleases to generate mature piRNAs. **e** Prevalent models for Piwi-mediated transcriptional silencing of transposons. In the nucleus of fruitfly (upper panel), the Piwi-piRNA complex binds to a nascent transposon transcript and recruits Panorax (Panx) and Asterix (Arx) mediator complex to the vicinity of the target chromatin region. Panorax further interacts with Nxf2 and Nxt1 and which recruits histone methyltransferase, dSetDB1(dSET) to methylate the Lysine residue at the 9th position of histone 3, which establishes a repressive chromatin state to suppress transposon expression [16–22]. In mice (lower panel), piRNA-loaded MIWI2 associates with TDRD9, DNMT3L, DNMT3a, and DNMT3a2 in transfected 293 T cells. The complex is guided by piRNA and deposits DNA methylation via DNMT3a2 [23]. However, the latest reports indicated that SPOCD1 links MIWI2 to Dnmt3a and Dnmt3L and the complex also contain TEX15 [24, 25]. Therefore, how exactly these proteins interact with each other remain unknown

and processed through ping-pong cycles that post-transcriptionally amplifies piRNAs with overlapping complementarity by accelerating their production from precursors by the alternative action of their associated PIWI proteins as follows [13, 26, 33–36]

(Fig. 1c). The transcribed long precursors are initially processed into 5' end monophosphorylated RNAs by PIWI protein loaded with initiator piRNAs. The processed precursors (pre-pre-piRNAs) are further cleaved by PIWI proteins to generate responder pre-piRNAs from the 5' ends of pre-pre-piRNAs. These pre-piRNAs are trimmed by exonucleases to generate functional piRNAs [38–41] and give rise to mature responder piRNAs that recognize the complementary strands of long precursors and initiated the ping-pong cycles. Meanwhile, 5' end monophosphorylated precursor RNAs are fragmented into a string of phased trailing pre-RNAs by PIWI and PIWI-coupled-proteins on the mitochondrial outer membrane [4–6]. Importantly, piRNAs are symbolically 2'-O-methylated at the 3' end by S-adenosylmethionine-dependent methyltransferases [9–12]. Perhaps the two best-known functions of PIWI proteins and piRNAs are transposon silencing and fertility [42]. However, many recent studies have started to reveal a much broader role of the PIWI-piRNA pathway in mediating the regulation of major constituents of the genome, which is the focus of this review.

The function of PIWI proteins and piRNAs in transposon silencing and fertility: an update

Before reviewing the new functions of the PIWI-piRNA pathway, here we provide an update on its function in transposon silencing and fertility, which also serves as needed background information of the rest of the review. Transposon silencing has been widely regarded as a requirement for fertility. However, this relationship has not been supported by definitive evidence. In contrast, these two functions are separable at least in *Drosophila* [43] and mice [44].

Transposon silencing is achieved by repression the expression of retrotransposon RNAs at both transcriptional and posttranscriptional levels [45]. Transcriptional silencing is mediated by nuclear PIWI proteins such as PIWI in *Drosophila* and MIWI2 (a.k.a. PIWIL4) in mice, whereas posttranscriptional silencing is mediated by cytoplasmic PIWI proteins such as Aubergine (Aub) and Ago3 in *Drosophila* or MIWI (a.k.a. PIWIL1) and MILI (a.k.a. PIWIL2) in mice. A prevalent model of transcriptional silencing in *Drosophila* is that the PIWI-piRNA complex binds to a nascent transposon transcript (Fig. 1e upper panel). The complex interacts with mediator proteins Asterix and Panoramic (a.k.a. Silencio). Furthermore, Panoramic interacts with Nxf2 (Nuclear Export Factor 2) and Nxt1 (Nuclear Transport Factor 2 Like Export Factor 1) which recruit a histone methyltransferase, dSetDB1 (a.k.a. Eggless), to methylate Lysine 9 residue in histone 3 (H3K9). This promotes the repressive chromatin state [16–22, 46–48]. Alternative models such as the *Drosophila* Piwi protein directly recruiting Heterochromatin Protein 1a to initiate heterochromatinization or recruiting the linker histone H1 to induce repressed chromatin state have also been proposed [49–51]. The above three models do not mutually exclude each other and might indeed co-exist in the cell. They share two common features: (1) transcriptional repression requires the specific binding of a PIWI-piRNA complex to a nascent RNA at the target site to recruit epigenetic/chromatin factors but does not require the PIWI slicer activity [52, 53]; (2) transcriptional repression occurs by modifying chromatin structure.

In mice, piRNA-dependent transcriptional silencing in the germline is also achieved by methylation of DNA through a recruitment scheme that is less known but similar to that in *Drosophila* [23, 54–60] (Fig. 1e, lower panel). For example, Tudor Domain-

Containing 9 (TDRD9) complexes with MIWI2 and suppresses transposons in a piRNA-dependent manner [61]. MIWI2 is guided by piRNA to the nascent transcript of specific genomic region to mediate the methylation of the target DNA [60]. MIWI2 in turn appears to directly recruit with DNMT3A, DNMT3A2, AND DNMT3L, and TDRD9 to achieve methylation, since MIWI2 seems to directly interact with these proteins in co-expression and immunoprecipitation assay in a 293 T cell [23]. However, a most recent report indicates that a nuclear protein, SPOCD1, interacts with MIWI2 as well as with DNMT3A, DNMT3L, and components of NURD and BAF chromatin remodeling complexes during de novo DNA methylation [24]. Among the DNA methylases, DNMT3C functions as one of the key methyltransferases that protects male germ cells from transposon activity [62]. Another nuclear protein, TEX15, also interacts with MIWI2 in the process [25]. A separate report indicates that TEX15 interacts with MILI, and the genome of *Tex15* mutant is hypomethylated similar to *Mili* and *Dnmt3c* but not *Miwi2* mutants [63]. Therefore, the above have revealed some of the key players in the PIWI-piRNA-mediated mechanism in the de novo DNA methylation, but substantial work is still needed to delineate their exact interactions and relative contributions to DNA methylation.

In addition, a recent study showed that the Ubiquitin-like, Containing PHD and RING finger Domains 1 (UHRF1) protein may interact with MIWI and MILI to deposit DNA methylation or to further recruit PRMT5 to methylate arginine of histones [64]. This study demonstrated the requirement of UHRF1 in histone and DNA methylation. However, the co-immunoprecipitation data did not convincingly show the UHRF1-MIWI/MILI interaction. In addition, MIWI and MILI are in the cytoplasm yet UHRF1 is in the nucleus. Thus, the nature of MIWI/MILI interaction with UHRF1 remains to be further explored. Finally, the zinc finger protein MORC1 promotes DNA methylation of transposons but not protein-coding genes in male germ cells during the period of global de novo methylation [65]. Whether/how these proteins work together to methylate the genome of male germ cells awaits further investigation.

Post-transcriptional silencing occurs in the cytoplasm, where a PIWI-piRNA complex binds to a piRNA-complementary transposon RNAs and cleaves the RNA by the slicer activity of the PIWI protein (Fig. 1c). This silencing meanwhile fulfills piRNA biogenesis. Both transcriptional and post-transcriptional silencing of transposons by the piRNA pathway has been well-reviewed (e.g., [66]); thus, this topic will not be further covered in our review.

In contrast to transposon silencing, little is known about whether/how PIWI proteins and piRNAs regulate gene expression and other major types of RNAs in the germline. In this review, we summarize recent research on this less-investigated topic. Recent progress has led to surprising discoveries that transposons actively regulate the expression of mRNAs and long noncoding RNAs (lncRNAs), mediated by the PIWI-piRNA pathway. Moreover, pseudogenes also regulate the expression of their cognate mRNAs, also mediated by the PIWI-piRNA pathway. Finally, the PIWI-piRNA pathway even regulates the expression of (sub-) telomeric and (peri-) centromeric transcripts. These surprising findings have begun to reveal genome-wide functions of the PIWI-piRNA pathway in functionally links the major classes of genomic sequences and protein-coding genes at the RNA level.

Regulation of mRNAs by transposons

A recently discovered function of the piRNA pathway in the cytoplasm is to mediate the regulation of mRNAs by transposons (Fig. 2). This discovery came as a surprise at multiple levels—from realizing that many mRNAs contain transposon sequences to learning that transposons can use these sequences as Trojan's horses to regulate target mRNAs, and to revealing the essential role of PIWI proteins and piRNAs in this regulation.

Many mRNAs in mammals contain transposon sequences in the 3' untranslated region

The first hint of transposons regulation towards mRNAs came from observations that some expressed retrotransposon sequences, especially short interspersed nuclear element (SINE) and long interspersed nuclear element (LINE) families, mapped to the 3' untranslated regions (UTRs) of mRNAs in the mammalian genome [67, 68] (Fig. 2). For example, 27.7% of mouse and 28.5% of human mRNAs contain at least one retrotransposon fragment, often a SINE, that is predominantly in the 3' UTR of mRNAs (Table 1), while a few others are in 5'UTR and even fewer are in protein-coding sequences (CDS) [68]. In *Drosophila*, densities of putative piRNA target sites in both the 5' UTR and especially the 3' UTR regions are also higher than that in coding sequences [69], even though this bias is not as pronounced as in mammalian genomes. The paucity of the transposon insertions in CDS likely results from evolutionary selection, because an insertion in CDS disrupts a protein's sequence and often its function, and thus, it is eliminated from the gene pool. However, there are frequent instances in which retrotransposon insertions reduce host mRNA expression [67, 68], indicating a potential role of the inserted transposon sequence in regulating mRNA expression.

Transposon sequences in mRNAs are Trojan's horses that lead to mRNA degradation

Recent studies indicate that transposon sequences in mRNAs are the targeting sites of transposon-derived piRNAs that degrade these mRNAs [68] (Fig. 2a). MIWI promotes the degradation of the mRNA from a mCherry reporter containing a B1 or B2 SINE sequence at its 3' UTR [68]. This RNA-degradation function depends on its slicer function. RNA immunoprecipitation (RIP) against MIWI and MILI followed by RT-PCR, as well as high-throughput sequencing of RNA isolated by crosslinking immunoprecipitation (HITS-CLIP) in the adult mouse testis showed that MIWI and MILI directly bind to mRNAs [76, 77]. Additionally, MIWI binds to mRNAs as a mRNA ribonucleoprotein complex in the testis [77, 78]. MIWI CLIP-seq indicates that piRNAs preferably bind to the 3' UTR of mRNAs in mouse elongating spermatids [72, 76]. Moreover, a group of meiotic mRNAs were significantly upregulated in *Miwi*^{-/-} mice round spermatids, and the degradation of meiotic mRNA exhibited piRNA generation signatures [68, 79]. All these findings indicate that PIWI proteins and their associated piRNAs might be directly involved in regulating the stability of target mRNAs.

The involvement of transposon-derived piRNAs in regulating their target mRNAs is further supported by experiments that investigate the effect of deleting piRNAs or their target sequences on mRNA stability. In mouse late spermatocytes deficient in *Miwi* or

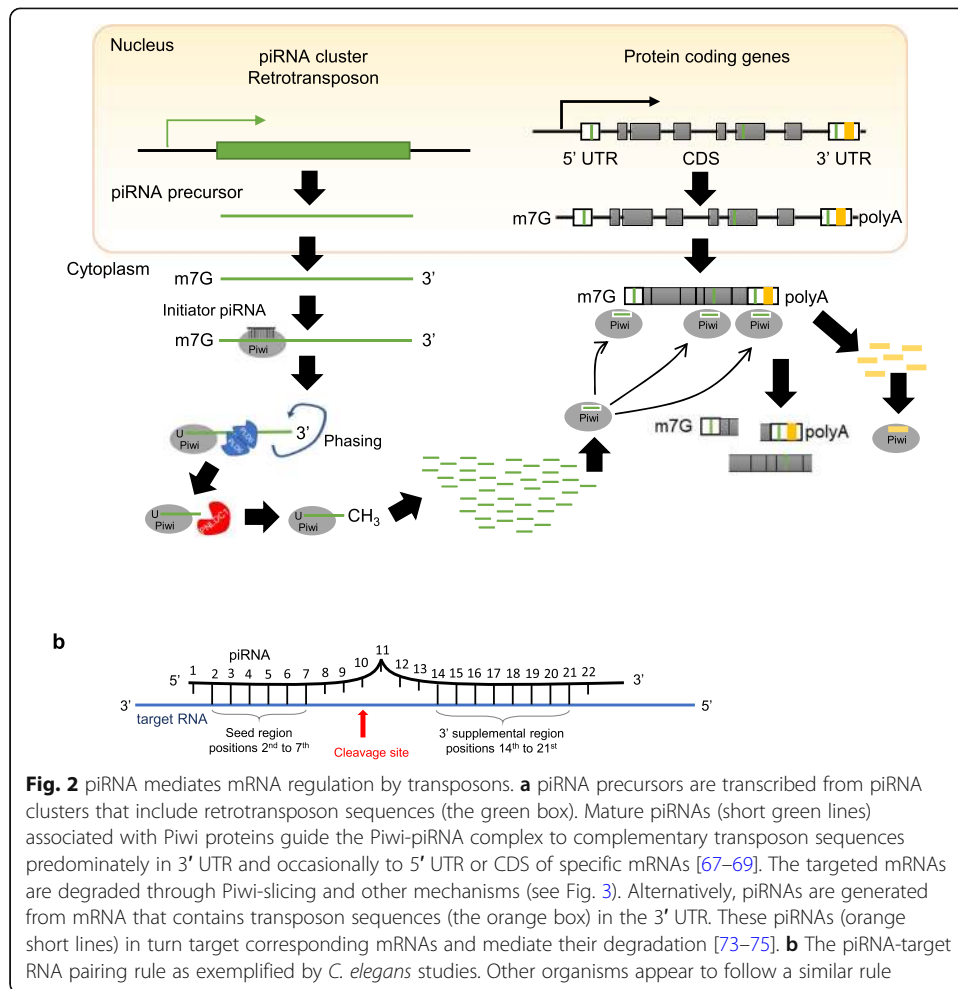


Table 1 Classification of retrotransposon sequence in mRNA 3' UTR

| Retrotransposon [67] | Human 3' UTR | Mouse 3' UTR |
|----------------------|--------------|--------------|
| DNA | 6.6% | 2.9% |
| LINE.L1 | 4.7% | 2.8% |
| LINE.L2 | 4.2% | 2.5% |
| LTR.ERV1 | 6.4% | 2.9% |
| LTR.ERVK | 0.7% | 2.6% |
| LTR.ERVL | 4.5% | 4.3% |
| LTR.MALR | 5.2% | 3.8% |
| SINE.Alu | 5.7% | 4.7% |
| SINE.B2 | N.A. | 5.0% |
| SINE.B4 | N.A. | 4.8% |
| SINE.MIR | 2.7% | 6.3% |

Mov10l1, another gene involved in piRNA biogenesis, piRNA biogenesis is deficient, and the expression levels of a large number of protein-coding genes are altered [68]. In addition, depletion of the retrotransposon sequences in either the 3' UTR of *prelid1* or the piRNA cluster that produces the targeting piRNAs results in a six-fold up-regulation of the *prelid1* mRNA level in late spermatocytes. Furthermore, this MIWI- and MOV10L1-mediated regulation of mRNAs requires the slicer activity of MIWI [68, 80]. Most recently, it was reported that deleting a specific piRNA cluster, *pi6*, led to up-regulating of mRNAs related to sperm function, in addition to affecting piRNA-piRNA precursor interactions [81]. This provides an additional example for piRNA function in regulating mRNAs and in spermatogenesis.

The suppression of retrotransposon-containing mRNAs in mouse primordial ovarian follicles also requires MILI and other piRNA pathway components such as mouse Vasa homolog (MVH) and TDRD9 [82]. In *Drosophila*, piRNAs suppress the *Stellate* repeat mRNA to ensure spermatogenesis, even though this regulation was initially reported as a siRNA-mediated mechanism [83]. These piRNAs, however, cannot repress the *Stellate* mRNA from closely related *Drosophila* species, thus introducing reproductive isolation [84]. All these observations demonstrate a direct role of PIWI proteins and their partner piRNAs in regulating mRNAs.

PIWI-piRNAs targeting rules

The slicing of piRNA-targeted mRNA mostly occurs at the 10th position from the 5' end of the targeting piRNAs [13, 68]. In-depth analysis of the cleavage site in the 3' UTR of *Tdrd1* mRNA with respect to its targeting piRNAs reveal that most piRNAs are aligned antisense to the mRNA sequence with their 5' nucleotide being 10 nucleotides away from the cleavage site, but several percent of piRNAs are aligned with their 5' nucleotide 11–20 nucleotides away [68]. This could indicate that the cleavage site for these piRNA-complementary mRNAs is downstream of the 10th position of piRNAs.

Despite the relatively clear information on the cleavage site as supported by in vitro, in vivo, and structural analyses, the piRNA targeting rule is still elusive and increasing effort has been put into revealing this puzzle. In *C. elegans* and mice, piRNA binding to target RNA is sequence-dependent but not completely sequence-specific, with position 2–8 as the seed sequence and 14–22 nucleotides are also important for piRNA targeting [69, 72, 85, 86], reminiscent of the microRNA seed sequence for target recognition (Fig. 2b). Further attempts in ectopic expression of human piRNAs in mouse testes show that perfect matching in nucleotides 2–11 of piRNA is required for mRNA targeting [79]. In a different study with deep RNA sequencing that allowed up to four mismatches at random positions in a piRNA, all of the 172 mRNAs that were significantly upregulated in *Piwi1*^{-/-} testes contain transposon piRNA target sites in their 3' UTR, with such a transposon sequence experimentally demonstrated to be responsible for MIWI-piRNA-mediated degradation [68]. Notably, these 172 mRNAs are quite different from the MIWI-target RNAs identified by Vourekas et al. (2016) using iCLIP. This likely reflects that iCLIP approach revealed stable binding of PIWI to target mRNAs that do not lead to mRNA degradation but possibly other regulatory effects, in contrast to RNA sequencing that revealed the degraded products of MIWI binding and cleavage of another set of target mRNAs (i.e., the mRNA degradome).

In *Drosophila*, mutating 1, 2, 3 nucleotides of a given piRNA in different positions did not abolish the binding ability of Piwi to its target nascent RNA at the genomic site, but reduce its binding affinity proportional to the number of mutations [50]. These observations also support the notion that piRNA targeting is sequence-dependent but not completely sequence-specific.

mRNAs also produce piRNAs

The source of piRNAs is sometimes mRNAs. In *Drosophila* ovaries, transposons that are integrated in the 3' UTR of actively transcribed genes induce piRNA production, which in-turn suppresses corresponding gene expression [73, 74] (Fig. 2a, right part). For example, in the *Drosophila* ovarian germline, proto-oncogene *c-Fos* is repressed post-transcriptionally in germline stem cells (GSCs) by Piwi protein. As a by-product of this repression, piRNAs are produced from the transposon elements in the 3' UTR of *c-Fos* mRNA [74]. This repression is critical for GSC self-renewal and differentiation. In the *Drosophila* ovarian somatic cell (OSC) line, the 3' UTR of *c-Fos* mRNA alone is sufficient to introduce gene suppression in a Piwi-dependent manner [74]. Similarly, in this OSC line, the 3' UTR of a protein-coding gene *traffic jam* produces piRNAs to repress the expression of another protein-coding gene, *fasciclin III* [75]. However, in this case, the *traffic jam* piRNAs are not transposon-originated piRNAs.

piRNAs can target the protein-coding sequence of mRNA for its degradation

A few publications have reported that piRNAs can also target the protein-coding sequence (CDS) of a mRNA to degrade the mRNA. For instance, in the *Drosophila* embryo, Aub directly binds to the CDS and the 3' UTR of maternal mRNAs in a piRNA-dependent manner to facilitate the decay of these maternal mRNAs [87]. The Aub-piRNA complex is also associated with CCR4 in *Drosophila* GSCs and represses the expression of *Casitas B-cell lymphoma (Cbl)* mRNA [88]. This mRNA contains target sites for transposon-derived piRNA in both 5' UTR and 3' UTR, and these target sites overlap with Aub CLIP hits. In *Caenorhabditis*, a sex-chromosome-derived piRNA suppresses the expression of *xol-1 (XO Lethal)*, the master sex determination gene, via targeting the CDS region of the *xol-1* mRNA [89]. In the silkworm, a *Feminizer*-derived piRNA targets the CDS of the *Masculinizer* mRNA and determines the sex of the silkworm [90]. In human breast cancer cell lines, a piRNA called *piR-FTH1* shows an inverse correlation with the expression of *Fth1 (ferritin heavy chain 1)* [91]. MILI and MIWI2 bind to *piR-FTH1* to repress the *Fth1* mRNA via targeting the CDS region of *Fth1* [91]. These observations from diverse and independent processes, each of which with biological significance, indicate that transposon sequences in CDS are equally effective in regulating mRNA stability. Thus, the paucity of piRNA target sites in the CDS is likely only due to their disruptive effect on protein-coding capacity but not a lower efficiency for piRNA-targeted degradation.

Finally, the insertion of transposon sequences into mRNAs as piRNA target sequences appears to be a dynamic process at least for certain cell types. The transposon landscape in cultured *Drosophila* ovarian somatic sheet (OSS) cells differs from that of cultured *Drosophila* OSCs [92], indicating that factors even as culture conditions may

select for transposon insertions that are in favor of the regulation of mRNA expression that meets the needs of the given condition.

Protein machinery involved in mRNA degradation by transposon-derived piRNAs

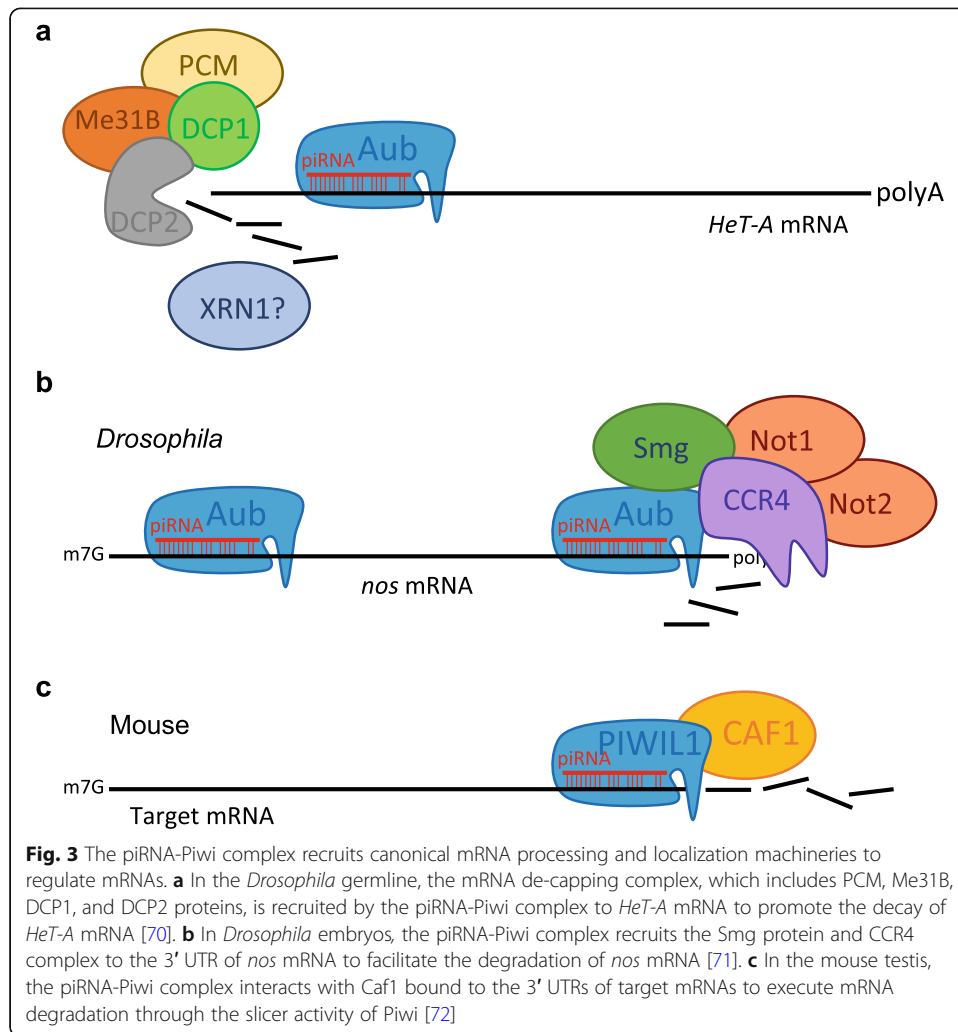
A few studies have further suggested that the PIWI-piRNA regulation of mRNAs involves canonical mRNA decay machineries. In the *Drosophila* germline, Aub and Ago3 colocalize with the mRNA degradation proteins Decapping Protein 1/2 (DCP1/2), Maternal expression at 31B (Me31B), and Pacman (PCM) in the nuage, a perinuclear riboprotein assembly, forming a complex (the pi-body) in a piRNA-dependent manner that facilitates mRNA degradation [70] (Fig. 3a). In early *Drosophila* embryos, the key piRNA pathway components Piwi, Aub, Ago3, Armitage (Armi), Spindle-E (Spn-E), and Squash (Squ) are all involved in the deadenylation and decay of maternal *Nanos* (*Nos*) mRNA through a piRNA targeting region in the *nos* 3' UTR [71]. Aub is known to achieve so by forming a complex with RNA-binding protein Smaug (Smg) and the deadenylase CCR4 at the sites of piRNA-mRNA interaction (Fig. 3b). The deadenylation machinery further recognizes the retrotransposon sequences in the *nos* 3' UTR and executes deadenylation within 4 h [71]. In mouse spermatogenic cells, the piRNA-induced silencing complex (pi-RISC) contains MIWI and CAF1 (Chromatin Assembly Factor 1), a key catalytic subunit of the CCR4-NOT deadenylase complex [72] (Fig. 3c). Thus, the piRNA mediated mRNA degradation is not only achieved by MIWI slicing of the target mRNAs but also by concurrent deadenylation, and possibly decapping, of mRNAs.

piRNAs activate the translation of mRNAs

The homology between PIWI proteins and eIF2C implies the possibility that they might promote translation, as first proposed for Aub in promoting translation of *nanos* mRNA in *Drosophila* embryos [93]. This function of Aub has been clearly demonstrated by a most recent study [94]. In the *Drosophila* early embryo, Aub physically interacts with eIF3d and the poly(A)-binding protein (PABP) to activate translation of *nanos* mRNA in the germ plasm (Fig. 4a).

PIWI proteins and piRNAs have also been implicated in promoting translation in mammalian systems. MILI is required for germline stem cell self-renewal in the mouse testis [96]. It forms a stable RNA-independent complex with eIF3a and associates with the m7G cap-binding complex that contains eIF4E and eIF4G in spermatogonia and spermatocytes [96]. In 7 dpp spermatogonia (most of which are germline stem cells), the *Mili* mutation has no significant effect on the cellular mRNA level but significantly reduces the rate of protein synthesis. Thus, MILI appears to positively regulate translation for germline stem cell self-renewal.

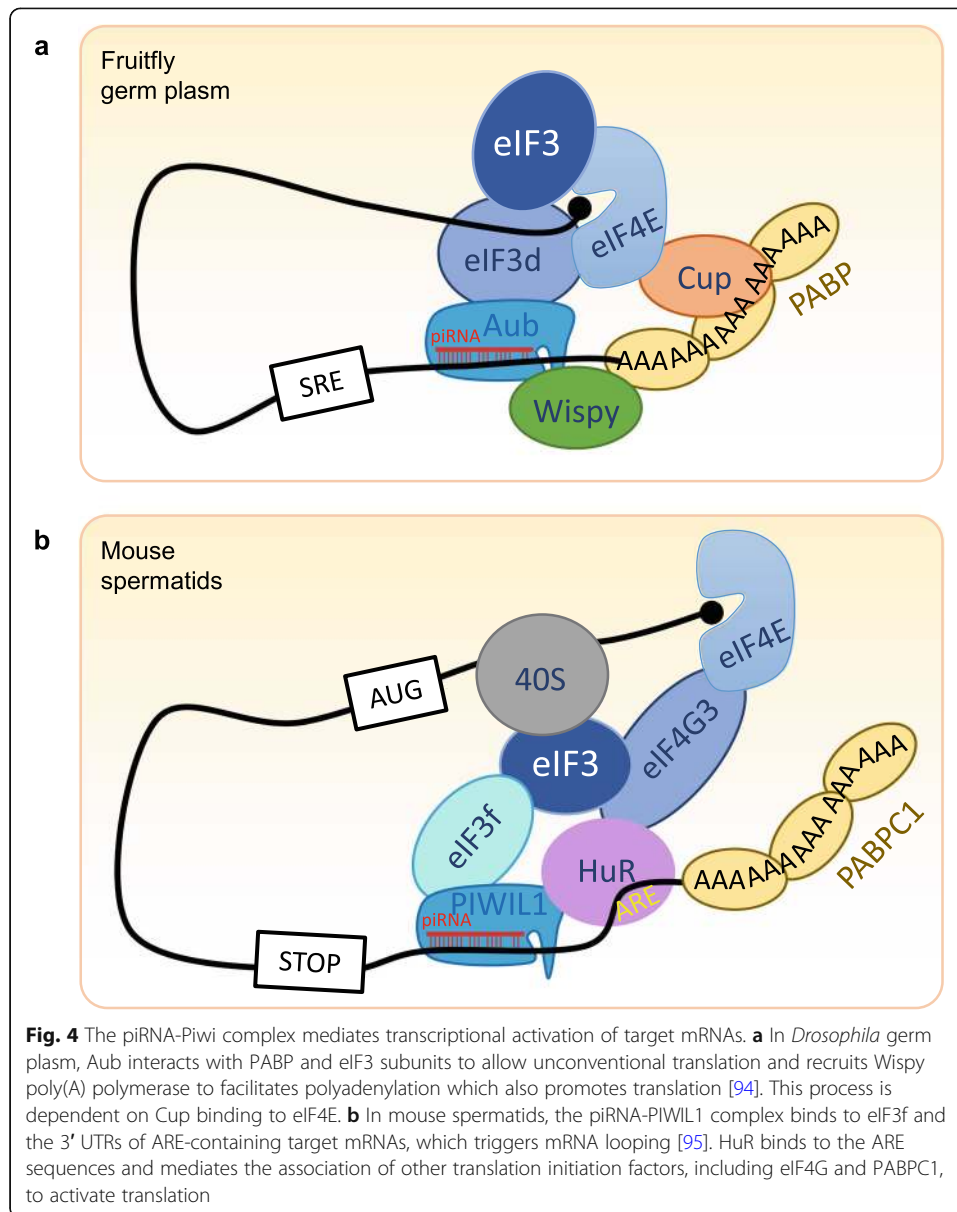
The MIWI-piRNA complex also associates with the capping complex, including cap-binding protein eIF4E, and regulates translation in the mouse spermatocytes and spermatids [97]. Recently, it has been further reported that, in mouse spermatids, a fraction of piRNAs associate with MIWI and recognize ARE (AU-rich element)-containing target mRNAs through imperfect base pairing in the 3' UTR of target mRNAs [95]. ARE-binding protein HuR loaded with target mRNA and the



translation initiation factor eIF3f form a complex with piRNA and MIWI to further activate the translation of mRNAs required for acrosome formation during spermatid development (Fig. 4b). These lines of evidence illustrate a sophisticated regulatory nexus of PIWI-piRNA complex in modulating the transcriptome at posttranscriptional level. Despite the above findings, the function of PIWI proteins in translational regulation has not been extensively studied and deserves more attention.

piRNAs regulate the localization of mRNAs

In addition to working with the mRNA decay machinery, the PIWI-piRNA complex has also been reported to be involved in the subcellular localization of mRNAs. For example, in *Drosophila*, the Aub-piRNA complex facilitates the localization of critical germ cell mRNAs to the germ plasm [98]. This function is achieved by working with Wispy, a germline-specific non-canonical poly(A) polymerase. During *Drosophila* oogenesis, Aub participates in the formation of the polar granule that localizes germline-specific mRNAs.



Regulation of lncRNAs by transposons

Recent studies also indicate that transposon sequences in lncRNAs are the targeting sites of transposon-derived piRNAs that identify these lncRNAs for degradation. Global characterization of transposon distribution in the human genome revealed that the vast majority of lncRNAs (83.4%) overlap with at least one transposon, in contrast to protein-coding sequences—only 6.2% of which overlap with transposons [99]. In the human genome, 75% of lncRNA transcripts contain an exon that originates at least partially from a transposon. Different types of transposons are differentially represented in lncRNA loci, and transposon landscapes differ among various tissue types [99]. lncRNAs that overlap with transposons are less expressed as compared to those that do not overlap with a transposon, indicating a suppressive role of transposon insertion [99].

Transposons repress lncRNA expression via the PIWI-piRNA pathway

The involvement of piRNAs in transposon regulation of lncRNAs has been demonstrated in multiple organisms. In mouse spermatocytes, retrotransposon sequences, mostly SINE, are distributed across the entire length of approximately 1500 lncRNAs that are upregulated in *Miwi*^{-/-} testis, which represents ~25% of the expressed lncRNAs in these cells [68]. This distribution is in contrast to their predominant localization to the 3' UTR of mRNAs. The transposon-containing lncRNAs become overexpressed in *Miwi* and *Mov10l1* mutants, implying that the transposon regulation of lncRNA expression is via the piRNA pathway [68] similar to its role in mediating transposon regulation of mRNAs (Fig. 5a). Presumably, the PIWI-piRNA complex slices the target lncRNAs in a way similar to their slicing of mRNAs. However, this remains to be directly demonstrated.

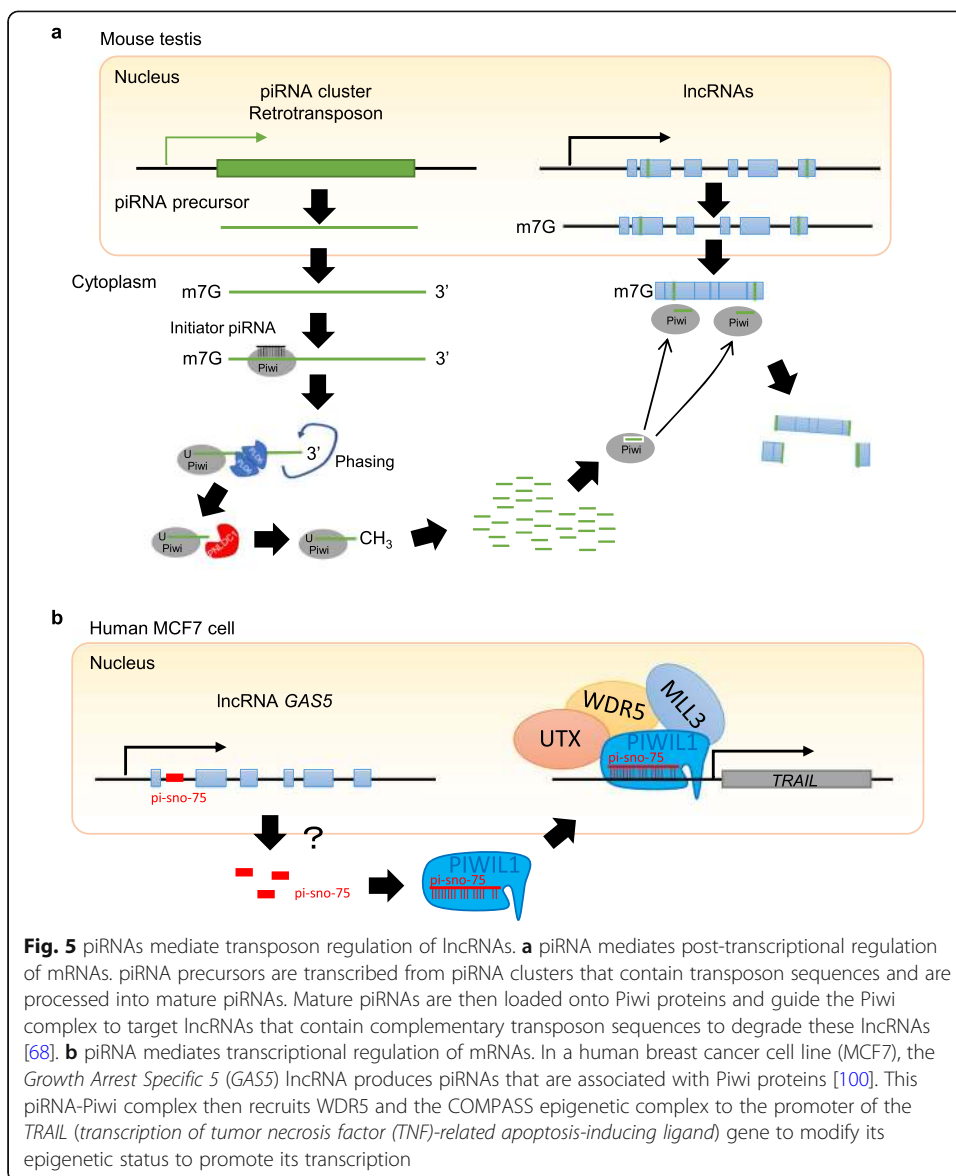
Other regulatory relationships between transposons, lncRNAs, and piRNAs

The piRNA regulation of lncRNA expression can be manifested in different ways. Some piRNA clusters that are expressed during the pachytene stage originate from testis-specific lncRNAs, indicating that lncRNAs themselves can be precursors of piRNAs [14, 101]. Although not as much is known about the relationship among transposons, lncRNAs, and piRNAs in the primordial germ cells, in their precursor cells in humans, i.e., human embryonic stem cells and induced pluripotent stem cells (iPSCs), the presence of transposon families such as HERVH (human endogenous retrovirus H) elements in lncRNAs is correlated with high expression levels of host lncRNAs, implicating that HERVH sequences positively regulate the expression of lncRNAs in a cell-type-specific manner [99], even though it is not known whether such regulation occurs at the transcriptional or post-transcriptional level. In *Drosophila* OSS and OSC cells, transposon insertions around lncRNA stimulate the expression of corresponding lncRNAs in a Piwi-dependent manner [92], implying a possible role of piRNA in transcriptional regulation of lncRNA expression.

In addition, some lncRNAs can produce piRNAs to promote gene expression at the transcriptional level. In a human breast cancer cell line (MCF7), a lncRNA called *Growth Arrest Specific 5 (GAS5)* facilitates the transcription of *TRAIL* [*transcription of tumor necrosis factor (TNF)-related apoptosis-inducing ligand*] mRNA. This is through GAS5-derived piRNAs that are associated with PIWI proteins [100]. These piRNA-PIWI complexes then recruit WDR5 and the COMPASS complex, a key complex for epigenetic modification, to the promoter of the *TRAIL* gene to modify the epigenetic status of the gene, which facilitates *TRAIL* transcription. This indicates an important role of PIWI-piRNA complex in epigenetic regulation of gene expressions (Fig. 5b). In *Tetrahymena*, piRNAs that are derived from noncoding RNA loci bind to *Tetrahymena* PIWI protein Twi8p and its target lncRNAs, which leads to a decrease in the level of the targeted lncRNAs [102]. These roles are similar to what has been found in mammalian systems.

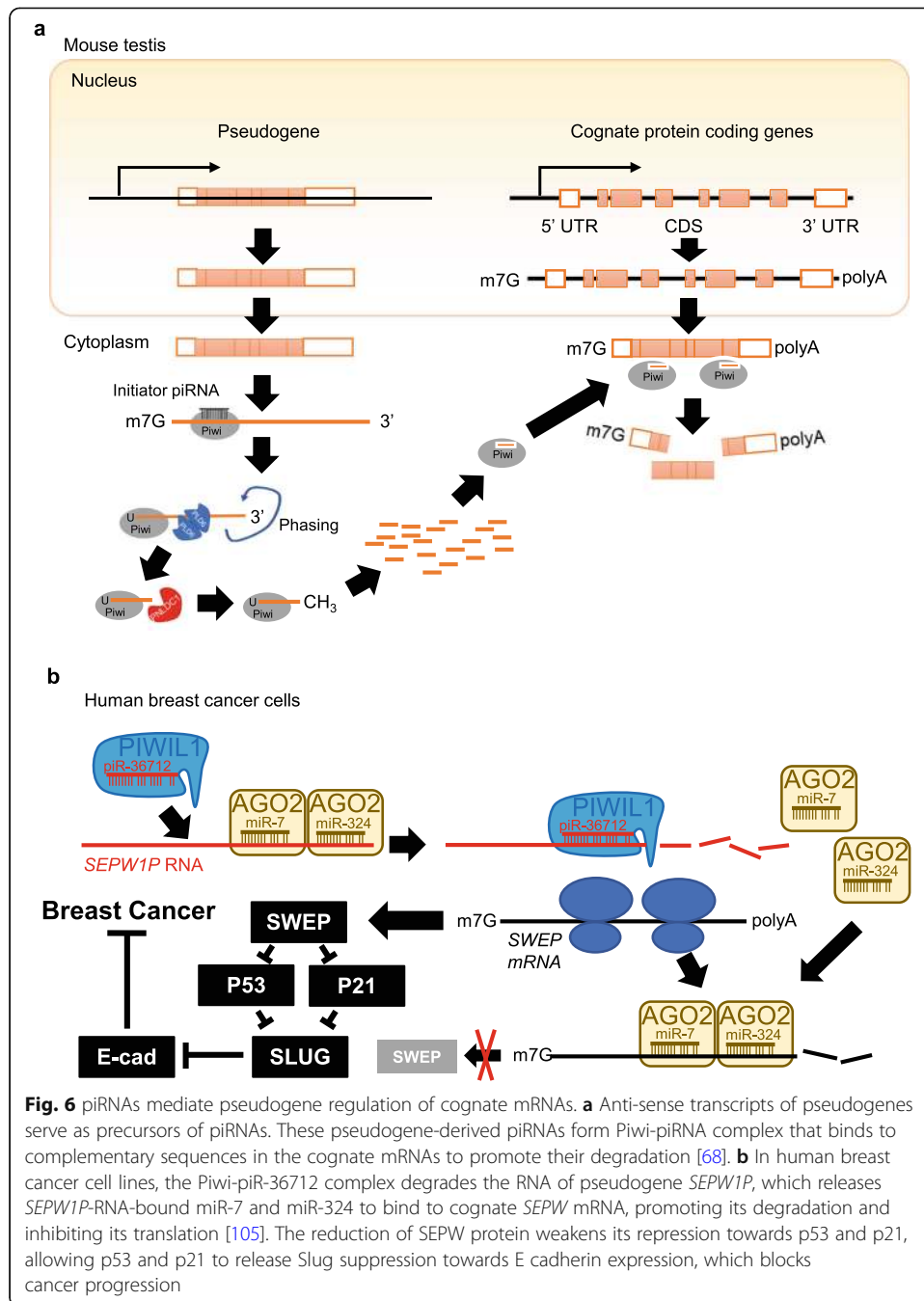
The regulation of mRNAs by pseudogenes

There are 14,000 pseudogenes in the human genome [103]. Analysis of human sperm small RNAs reveals that some piRNA clusters are located within pseudogenes [104].



piRNAs derived from pseudogenes are predicted to target protein-coding cognate genes. In mouse late spermatocytes, at least 14 genes are significantly regulated by their pseudogenes [68]. Thus, pseudogene regulation of mRNAs via piRNAs may be a significant mechanism in the germline that has been largely ignored.

Recent studies indicate that piRNA derived from pseudogenes can regulate the expression of their cognate mRNAs. The first clear demonstration of such a function comes from a study of the mouse late spermatocyte. The mRNA level of the *Stambp* gene is dramatically increased in *Miwi*^{-/-} and *Mov10l1*^{-/-} late spermatocytes, indicating that the piRNA pathway may be involved in preventing the over-expression of *Stambp* mRNA [68] (Fig. 6a). However, the *Stambp* mRNA complements to PIWIL1-associated piRNAs that do not map to transposons but uniquely to its pseudogene, *Stambp-ps1* [68]. Indeed, knocking out *Stambp-ps1* expression abolishes the *Stambp-ps1*-derived piRNAs and causes drastic increases in



the *Stambp* mRNA level, illustrating the importance of a pseudogene in the regulation of its cognate active gene via the piRNA [68].

Conversely, pseudogenes can also be regulated by piRNAs. In human breast cancer cell lines MCF7 and ZR75-1, a piRNA called piR-36,712 directly binds to pseudogene *SEPW1P* RNA and reduces its level [105] (Fig. 6b). This releases its bound miRNAs miR-7 miR-324, allowing more miR-7 miR-324 to target the mRNA of its cognate gene *SEPW* to reduce its stability and translation. The reduced *SEPW* level in turn weakens its inhibition towards p53 and p21, allowing p53 and p21 to reduce the suppression of E cadherin expression by a zinc finger protein called Slug, which blocks cancer

development [105]. This type of regulation may be specific to cancer cells because piRNAs are not detected in normal somatic cells, such as breast epithelial cells,

Pseudogene-derived piRNAs have been identified across the evolutionary tree. In *Tetrahymena*, pseudogene-derived small RNAs have been identified that have binding preference to Twi2 (*Tetrahymena* PIWI2) but not to Twi7 or Twi8 [106]. In the marmoset testis, piRNA clusters contain processed pseudogenes that are antisense-oriented relative to their cognate genes so they can produce piRNAs that target their cognate genes (cf. Fig. 6aa) [107]. In addition, pig piRNA clusters are also highly enriched for exon sequences of pseudogenes [108]. In vertebrate genomes, endogenous bornavirus-like nucleoprotein elements (EBLNs) represent a special type of genomic sequences derived from ancient bornaviral nucleoprotein mRNA via retrotransposition. Human EBLNs are actively transcribed pseudogenes that give rise to piRNAs with a possible role in interfering with ancient bornaviral infection [109]. Thus, regulation by pseudogene-derived piRNAs is likely a conserved regulatory pathway.

Diverse functions of satellite-repeat derived piRNAs

Satellite-repeated derived piRNAs negatively regulates mRNA stability

In *Drosophila*, it has been shown that pericentromeric piRNAs, specifically, *AT-chX* piRNAs and *Su(Ste)* piRNAs, also suppresses the *vasa* mRNA of a different species of *Drosophila* during interspecies mating to ensure reproductive isolation [84]. Most recently, it was shown that a large number of piRNAs were derived from a satellite repeat in both the soma tissues and germline of the mosquito *Aedes aegypti* [110]. Two of them, tapiR1 and tapiR2 (tandem repeat-associated piRNA1 and 2), are associated with *Aedes aegypti* PIWI protein Piwi4. The resulting Piwi4-piRNA complex silence target RNAs (both mRNA and lncRNA) via a sequence-specific recognition rule reminiscent of microRNA seed sequence. Notably, the piRNA-generating satellite repeats were highly conserved across mosquito species for approximately 200 million years and are very similar to that of other higher organisms. Hence, there might be an evolutionarily conserved mechanism for post-transcriptional regulation mediated by PIWI-satellite piRNA complex.

Subtelomeric piRNAs regulate telomeric function

Remarkably, the piRNAs are even involved in regulating telomeres and centromeres. Telomeric and centromeric piRNAs have been reported in various organisms. Shortly after piRNAs were characterized, piRNAs were mapped to subtelomeric regions of the *Drosophila* genome. One of these piRNAs, the 3R-TAS1 piRNA, regulates the epigenetic state of the target subtelomeric sequence [111] and is essential for germline stem cell maintenance [112]. Telomeric piRNAs in *Drosophila* are required for the deposition of Heterochromatin Protein 1 (HP1) and its homolog Rhino as well as epigenetic suppression of telomeric retrotransposon [113, 114]. Loss of piRNA pathway proteins results in the significant upregulation of telomeric retroelement transcripts and downregulation of telomeric piRNAs [115–118]. HP1a, a partner of Piwi protein [49, 111], is required for the generation of piRNAs that map to telomeres and peri-centromeres in the *Drosophila* germline [119]. The nuclear CCR-NOT complex is required for the

degradation of telomeric transcripts in a Piwi-dependent manner [120]. In addition, *Aub*, *Armi*, and *Ago3* mutations reduce telomeric piRNAs production and disrupt the binding between telomere and telomere protection complex [117, 121]. All these findings together indicate an essential role of the piRNA pathway in subtelomeric and telomeric function in *Drosophila*.

Interestingly, the prevalence of subtelomeric piRNA clusters is significantly higher in the genomes of *Drosophila* in the wilderness than in standardized laboratory *Drosophila* stocks [122]. Since a well-established telomeric function in *Drosophila* is to suppress telomeric retrotransposons, this difference may reflect another function of telomeric piRNAs in the suppression of telomeric retrotransposons, which might be more needed by flies in the wilderness that are possibly more ridden with retrotransposons.

Because mammalian telomeres are composed of simple repeats instead of telomeric retroelement, the telomeric function of piRNA in *Drosophila* is unlikely to be conserved. However, subtelomeric sequence and structure is highly conserved from *Drosophila* to mammals. Hence, some other important function of the subtelomere, such as repressing and regulating nearby euchromatic gene expression (position effect), might be conserved in higher organisms.

In *Tetrahymena*, telomeric repeat-derived small RNAs selectively bind to Twi10 [106]. In *Caenorhabditis*, perfect telomeric small RNAs are extremely rare among small RNAs that are immunoprecipitated with Ago proteins; however, there are significant numbers of telomeric small RNAs with mismatches were associated with Ago proteins [123]. These small RNAs might be involved in telomere protection in a way similar to that in *Drosophila*.

Pericentromeric and centromeric piRNAs regulate centromere function

In *Drosophila*, pericentromeric satellite repeats are known to be crucial for centromere function and give rise to piRNAs [13, 28, 124] that are critical for the suppression of *satellite* RNA expression [28, 125, 126]. The involvement of piRNAs in telomeres and centromeres is likely to be crucial for the genome integrity and segregation during germ cell division. A clear requirement of the Piwi-piRNA pathway for chromosome segregation is illustrated by a most recent study in mice, where MIWI prevents aneuploidy during meiosis by piRNA-guided cleavage of excess major and minor satellite RNAs [127]. This allows normal assembly of homologous kinetochore pairs. Overexpression of these satellite RNAs in wildtype meiotic cells also causes aneuploidy. Dicer facilitates the degradation of MIWI cleavage products and other double-strand satellite RNAs. These findings, supported by another latest study on Dicer [128], start to reveal a direct role of PIWI proteins and satellite RNAs in chromosome segregation during meiosis.

Conclusion

A eukaryotic genome is composed of protein-coding genes, transposons, pseudogenes, centromeres, telomeres, and other repeat sequences. Recent studies, mostly in the germline, start to reveal important roles of piRNAs in functionally linking these major constituents of the genome. piRNAs can be derived from transposon RNAs, pseudogene RNAs, mRNAs, lncRNAs, and telomeric/centromeric repeat RNAs and

subsequently partner with PIWI proteins to recognize complementary sequences in these RNAs. These observations together start to reveal a complex network of regulation mediated by piRNAs that unifies the genome at the post-transcriptional level. This is perhaps the first regulatory network of any kind that executes genome-wide regulation in any cell type. Future investigations on this complex network at the mechanistic level will further reveal how this network interacts with other canonical RNA regulation mechanisms to define the genomic function in germ cells and shed light on such function in other cell types.

Acknowledgements

We thank Christopher Antos, Ting Lu, Shou Shi, Yuanyuan Gong, and Sanhong Liu for critical reading of the manuscript.

Authors' contributions

H.L. conceived the topic and the outline of the review. C.W. and H.L. wrote the review. Both authors read and approved the final manuscript.

Funding

C.W. is supported by ShanghaiTech University, and H.L. is funded by NIH R37HD42012.

Competing interests

The authors declare no competing interests.

Glossary

Transposon:

A DNA segment that can undergo transposition into a different site in the genome.

Retrotransposons:

Transposons that transcribe itself into a RNA intermediate for reverse transcription to produce double-stranded DNA for transposition (insertion).

Pseudogene:

A DNA segment that is an imperfect copy of a functional gene, often including a (sometime altered) promoter. Most pseudogenes are transcribed but not translated due to mutations in protein-coding sequences.

lncRNA:

Long non-coding RNAs, which are conventionally defined as RNAs that are at least 200-nucleotide in length but do not contain a protein-coding sequence.

Dicer:

A type of RNase III family endoribonucleases that cleave double-stranded RNA into short double-stranded RNA fragments of mostly 21–22 nucleotides.

Centromere:

The region of a chromosome, containing highly repetitive DNA sequence, that serves as the attachment site for microtubules of the spindle during cell division.

Pericentromere:

The chromosomal regions flanking the centromere that contains highly repetitive DNA and are heterochromatic.

Telomere:

The two end-regions of a chromosome that protect the chromosome from shortening after DNA replication or from fusion to other chromosomes.

Subtelomere:

The two regions of a chromosome that is right next to the telomere, containing of highly repetitive DNA sequences and are heterochromatic.

Pachytene:

A mid-stage of the prophase of meiosis I, following zygotene, characterized by the presence of shortened and thickened paired chromosomes, with the two chromatids clearly separate.

Author details

¹Shanghai Institute for Advanced Immunochemical Studies, ShanghaiTech University, Shanghai 201210, China. ²Yale Stem Cell Center and Department of Cell Biology, Yale University School of Medicine, New Haven, CT 06519, USA.

Received: 5 August 2020 Accepted: 7 December 2020

Published online: 08 January 2021

References

1. Song JJ, Smith SK, Hannon GJ, Joshua-Tor L. Crystal structure of Argonaute and its implications for RISC slicer activity. *Science*. 2004;305:1434–7.
2. Matsumoto N, Nishimasu H, Sakakibara K, Nishida KM, Hirano T, Ishitani R, Siomi H, Siomi MC, Nureki O. Crystal structure of silkworm PIWI-clade Argonaute Siwi bound to piRNA. *Cell*. 2016;167:484–97 e489.
3. Yamaguchi S, Oe A, Nishida KM, Yamashita K, Kajiya A, Hirano S, Matsumoto N, Dohmae N, Ishitani R, Saito K, et al. Crystal structure of *Drosophila* Piwi. *Nat Commun*. 2020;11:858.

4. Homolka D, Pandey RR, Goriaux C, Brassat E, Vauray C, Sachidanandam R, Fauvarque MO, Pillai RS. PIWI slicing and RNA elements in precursors instruct directional primary piRNA biogenesis. *Cell Rep.* 2015;12:418–28.
5. Mohn F, Handler D, Brennecke J. Noncoding RNA. piRNA-guided slicing specifies transcripts for Zucchini-dependent, phased piRNA biogenesis. *Science.* 2015;348:812–7.
6. Han BW, Wang W, Li C, Weng Z, Zamore PD. Noncoding RNA. piRNA-guided transposon cleavage initiates Zucchini-dependent, phased piRNA production. *Science.* 2015;348:817–21.
7. Hayashi R, Schnabl J, Handler D, Mohn F, Ameres SL, Brennecke J. Genetic and mechanistic diversity of piRNA 3'-end formation. *Nature.* 2016;539:588–92.
8. Wang H, Ma Z, Niu K, Xiao Y, Wu X, Pan C, Zhao Y, Wang K, Zhang Y, Liu N. Antagonistic roles of nibbler and Hen1 in modulating piRNA 3' ends in *Drosophila*. *Development.* 2016;143:530–9.
9. Horwich MD, Li C, Matranga C, Vagin V, Farley G, Wang P, Zamore PD. The *Drosophila* RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC. *Curr Biol.* 2007;17:1265–72.
10. Saito K, Sakaguchi Y, Suzuki T, Suzuki T, Siomi H, Siomi MC. Pimet, the *Drosophila* homolog of HEN1, mediates 2'-O-methylation of Piwi-interacting RNAs at their 3' ends. *Genes Dev.* 2007;21:1603–8.
11. Ohara T, Sakaguchi Y, Suzuki T, Ueda H, Miyauchi K, Suzuki T. The 3' termini of mouse Piwi-interacting RNAs are 2'-O-methylated. *Nat Struct Mol Biol.* 2007;14:349–50.
12. Kirino Y, Mourelatos Z. Mouse Piwi-interacting RNAs are 2'-O-methylated at their 3' termini. *Nat Struct Mol Biol.* 2007;14:347–8.
13. Brennecke J, Aravin AA, Stark A, Dus M, Kellis M, Sachidanandam R, Hannon GJ. Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell.* 2007;128:1089–103.
14. Gainetdinov I, Skvortsova Y, Kondratieva S, Funikov S, Azhikina T. Two modes of targeting transposable elements by piRNA pathway in human testis. *RNA.* 2017;23:1614–25.
15. Gu W, Lee HC, Chaves D, Youngman EM, Pazour GJ, Conte D Jr, Mello CC. CapSeq and CIP-TAP identify Pol II start sites and reveal capped small RNAs as *C. elegans* piRNA precursors. *Cell.* 2012;151:1488–500.
16. Wang SH, Elgin SC. *Drosophila* Piwi functions downstream of piRNA production mediating a chromatin-based transposon silencing mechanism in female germ line. *Proc Natl Acad Sci U S A.* 2011;108:21164–9.
17. Muerdter F, Guzzardo PM, Gillis J, Luo Y, Yu Y, Chen C, Fekete R, Hannon GJ. A genome-wide RNAi screen draws a genetic framework for transposon control and primary piRNA biogenesis in *Drosophila*. *Mol Cell.* 2013;50:736–48.
18. Sienski G, Donertas D, Brennecke J. Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression. *Cell.* 2012;151:964–80.
19. Le Thomas A, Rogers AK, Webster A, Marinov GK, Liao SE, Perkins EM, Hur JK, Aravin AA, Toth KF. Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. *Genes Dev.* 2013;27:390–9.
20. Ohtani H, Iwasaki YW, Shibuya A, Siomi H, Siomi MC, Saito K. DmGTSF1 is necessary for Piwi-piRISC-mediated transcriptional transposon silencing in the *Drosophila* ovary. *Genes Dev.* 2013;27:1656–61.
21. Sienski G, Batki J, Senti KA, Donertas D, Tirian L, Meixner K, Brennecke J. Silencio/CG9754 connects the Piwi-piRNA complex to the cellular heterochromatin machinery. *Genes Dev.* 2015;29:2258–71.
22. Yu Y, Gu J, Jin Y, Luo Y, Preall JB, Ma J, Czech B, Hannon GJ. Panoramix enforces piRNA-dependent cotranscriptional silencing. *Science.* 2015;350:339–42.
23. Kojima-Kita K, Kuramochi-Miyagawa S, Nagamori I, Ogonuki N, Ogura A, Hasuwa H, Akazawa T, Inoue N, Nakano T. MIWI2 as an effector of DNA methylation and gene silencing in embryonic male germ cells. *Cell Rep.* 2016;16:2819–28.
24. Zoch A, Auchynnikava T, Berrens RV, Kabayama Y, Schopp T, Heep M, Vasiliauskaite L, Perez-Rico YA, Cook AG, Shkumatava A, et al. SPOCD1 is an essential executor of piRNA-directed de novo DNA methylation. *Nature.* 2020;584:635–9.
25. Schopp T, Zoch A, Berrens RV, Auchynnikava T, Kabayama Y, Vasiliauskaite L, Rappsilber J, Allshire RC, O'Carroll D. TEX15 is an essential executor of MIWI2-directed transposon DNA methylation and silencing. *Nat Commun.* 2020;11:3739.
26. Gunawardane LS, Saito K, Nishida KM, Miyoshi K, Kawamura Y, Nagami T, Siomi H, Siomi MC. A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in *Drosophila*. *Science.* 2007;315:1587–90.
27. Okamura K, Ishizuka A, Siomi H, Siomi MC. Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways. *Genes Dev.* 2004;18:1655–66.
28. Vagin VV, Sigova A, Li C, Seitz H, Gvozdev V, Zamore PD. A distinct small RNA pathway silences selfish genetic elements in the germline. *Science.* 2006;313:320–4.
29. Kim VN, Han J, Siomi MC. Biogenesis of small RNAs in animals. *Nat Rev Mol Cell Biol.* 2009;10:126–39.
30. Teixeira FK, Okuniewska M, Malone CD, Coux RX, Rio DC, Lehmann R. piRNA-mediated regulation of transposon alternative splicing in the soma and germ line. *Nature.* 2017;552:268–72.
31. Lewis SH, Quarles KA, Yang Y, Tanguy M, Frezal L, Smith SA, Sharma PP, Cordaux R, Gilbert C, Giraud I, et al. Pan-arthropod analysis reveals somatic piRNAs as an ancestral defence against transposable elements. *Nat Ecol Evol.* 2018;2:174–81.
32. Ross RJ, Weiner MM, Lin H. PIWI proteins and PIWI-interacting RNAs in the soma. *Nature.* 2014;505:353–9.
33. Aravin A, Gaidatzis D, Pfeffer S, Lagos-Quintana M, Landgraf P, Iovino N, Morris P, Brownstein MJ, Kuramochi-Miyagawa S, Nakano T, et al. A novel class of small RNAs bind to MILI protein in mouse testes. *Nature.* 2006;442:203–7.
34. Girard A, Sachidanandam R, Hannon GJ, Carmell MA. A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature.* 2006;442:199–202.
35. Grivna ST, Beyret E, Wang Z, Lin H. A novel class of small RNAs in mouse spermatogenic cells. *Genes Dev.* 2006;20:1709–14.
36. Lau NC, Seto AG, Kim J, Kuramochi-Miyagawa S, Nakano T, Bartel DP, Kingston RE. Characterization of the piRNA complex from rat testes. *Science.* 2006;313:363–7.
37. Ruby JG, Jan C, Player C, Axtell MJ, Lee W, Nusbaum C, Ge H, Bartel DP. Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in *C. elegans*. *Cell.* 2006;127:1193–207.
38. Saxe JP, Chen M, Zhao H, Lin H. Tdrkh is essential for spermatogenesis and participates in primary piRNA biogenesis in the germline. *EMBO J.* 2013;32:1869–85.
39. Kawaoka S, Izumi N, Katsuma S, Tomari Y. 3' end formation of PIWI-interacting RNAs in vitro. *Mol Cell.* 2011;43:1015–22.
40. Tang W, Tu S, Lee HC, Weng Z, Mello CC. The RNase PARN-1 trims piRNA 3' ends to promote transcriptome surveillance in *C. elegans*. *Cell.* 2016;164:974–84.

41. Izumi N, Shoji K, Sakaguchi Y, Honda S, Kirino Y, Suzuki T, Katsuma S, Tomari Y. Identification and functional analysis of the pre-piRNA 3' trimmer in silkworms. *Cell*. 2016;164:962–73.
42. Juliano C, Wang J, Lin H. Uniting germline and stem cells: the function of Piwi proteins and the piRNA pathway in diverse organisms. *Annu Rev Genet*. 2011;45:447–69.
43. Klenov MS, Sokolova OA, Yakushev EY, Stolyarenko AD, Mikhaleva EA, Lavrov SA, Gvozdev VA. Separation of stem cell maintenance and transposon silencing functions of Piwi protein. *Proc Natl Acad Sci U S A*. 2011;108:18760–5.
44. Xu M, You Y, Hunsicker P, Hori T, Small C, Griswold MD, Hecht NB. Mice deficient for a small cluster of Piwi-interacting RNAs implicate Piwi-interacting RNAs in transposon control. *Biol Reprod*. 2008;79:51–7.
45. Zhang F, Wang J, Xu J, Zhang Z, Koppetsch BS, Schultz N, Vreven T, Meignin C, Davis I, Zamore PD, et al. UAP56 couples piRNA clusters to the perinuclear transposon silencing machinery. *Cell*. 2012;151:871–84.
46. Fabry MH, Ciabrelli F, Munafo M, Eastwood EL, Kneuss E, Falciatori I, Falconio FA, Hannon GJ, Czech B. piRNA-guided co-transcriptional silencing coopts nuclear export factors. *eLife* 2019;8:e47999.
47. Batki J, Schnabl J, Wang J, Handler D, Andreev VI, Stieger CE, Novatchkova M, Lampersberger L, Kauneckaitė K, Xie W, et al. The nascent RNA binding complex SFINX licenses piRNA-guided heterochromatin formation. *Nat Struct Mol Biol*. 2019;26:720–31.
48. Murano K, Iwasaki YW, Ishizu H, Mashiko A, Shibuya A, Kondo S, Adachi S, Suzuki S, Saito K, Natsume T, et al. Nuclear RNA export factor variant initiates piRNA-guided co-transcriptional silencing. *EMBO J*. 2019;38:e102870.
49. Brower-Toland B, Findley SD, Jiang L, Liu L, Yin H, Dus M, Zhou P, Elgin SC, Lin H. Drosophila PIWI associates with chromatin and interacts directly with HP1a. *Genes Dev*. 2007;21:2300–11.
50. Huang XA, Yin H, Sweeney S, Raha D, Snyder M, Lin H. A major epigenetic programming mechanism guided by piRNAs. *Dev Cell*. 2013;24:502–16.
51. Iwasaki YW, Murano K, Ishizu H, Shibuya A, Iyoda Y, Siomi MC, Siomi H, Saito K. Piwi modulates chromatin accessibility by regulating multiple factors including histone H1 to repress transposons. *Mol Cell*. 2016;63:408–19.
52. De Fazio S, Bartonicek N, Di Giacomo M, Abreu-Goodger C, Sankar A, Funaya C, Antony C, Moreira PN, Enright AJ, O'Carroll D. The endonuclease activity of Mili fuels piRNA amplification that silences LINE1 elements. *Nature*. 2011;480:259–63.
53. Darricarrere N, Liu N, Watanabe T, Lin H. Function of Piwi, a nuclear Piwi/Argonaute protein, is independent of its slicer activity. *Proc Natl Acad Sci U S A*. 2013;110:1297–302.
54. Aravin AA, Sachidanandam R, Bourc'his D, Schaefer C, Pezic D, Toth KF, Bestor T, Hannon GJ. A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. *Mol Cell*. 2008;31:785–99.
55. Aravin AA, Bourc'his D. Small RNA guides for de novo DNA methylation in mammalian germ cells. *Genes Dev*. 2008;22:970–5.
56. Molaro A, Falciatori I, Hodges E, Aravin AA, Marran K, Rafii S, McCombie WR, Smith AD, Hannon GJ. Two waves of de novo methylation during mouse germ cell development. *Genes Dev*. 2014;28:1544–9.
57. Ishiuchi T, Torres-Padilla ME. LINEing germ and embryonic stem cells' silencing of retrotransposons. *Genes Dev*. 2014;28:1381–3.
58. Manakov SA, Pezic D, Marinov GK, Pastor WA, Sachidanandam R, Aravin AA. MIWI2 and MILI have differential effects on piRNA biogenesis and DNA methylation. *Cell Rep*. 2015;12:1234–43.
59. Nagamori I, Kobayashi H, Shiromoto Y, Nishimura T, Kuramochi-Miyagawa S, Kono T, Nakano T. Comprehensive DNA methylation analysis of retrotransposons in male germ cells. *Cell Rep*. 2015;12:1541–7.
60. Watanabe T, Cui X, Yuan Z, Qi H, Lin H. MIWI2 targets RNAs transcribed from piRNA-dependent regions to drive DNA methylation in mouse prospermatogonia. *EMBO J*. 2018;37:e95329.
61. Shoji M, Tanaka T, Hosokawa M, Reuter M, Stark A, Kato Y, Kondoh G, Okawa K, Chujo T, Suzuki T, et al. The TDRD9-MIWI2 complex is essential for piRNA-mediated retrotransposon silencing in the mouse male germline. *Dev Cell*. 2009;17:775–87.
62. Barau J, Teissandier A, Zamudio N, Roy S, Nalesso V, Herault Y, Guillou F, Bourc'his D. The DNA methyltransferase DNMT3C protects male germ cells from transposon activity. *Science*. 2016;354:909–12.
63. Yang F, Lan Y, Pandey RR, Homolka D, Berger SL, Pillai RS, Bartolomei MS, Wang PJ. TEX15 associates with MILI and silences transposable elements in male germ cells. *Genes Dev*. 2020;34:745–50.
64. Dong J, Wang X, Cao C, Wen Y, Sakashita A, Chen S, Zhang J, Zhang Y, Zhou L, Luo M, et al. UHRF1 suppresses retrotransposons and cooperates with PRMT5 and PIWI proteins in male germ cells. *Nat Commun*. 2019;10:4705.
65. Pastor WA, Stroud H, Nee K, Liu W, Pezic D, Manakov S, Lee SA, Moissiard G, Zamudio N, Bourc'his D, et al. MORC1 represses transposable elements in the mouse male germline. *Nat Commun*. 2014;5:5795.
66. Ozata DM, Gainetdinov I, Zoch A, O'Carroll D, Zamore PD. PIWI-interacting RNAs: small RNAs with big functions. *Nat Rev Genet*. 2019;20:89–108.
67. Faulkner GJ, Kimura Y, Daub CO, Wani S, Plessy C, Irvine KM, Schroder K, Cloonan N, Steptoe AL, Lassmann T, et al. The regulated retrotransposon transcriptome of mammalian cells. *Nat Genet*. 2009;41:563–71.
68. Watanabe T, Cheng EC, Zhong M, Lin H. Retrotransposons and pseudogenes regulate mRNAs and lncRNAs via the piRNA pathway in the germline. *Genome Res*. 2015;25:368–80.
69. Vourekas A, Alexiou P, Vrettos N, Maragkakis M, Mourelatos Z. Sequence-dependent but not sequence-specific piRNA adhesion traps mRNAs to the germ plasm. *Nature*. 2016;531:390–4.
70. Lim AK, Tao L, Kai T. piRNAs mediate posttranscriptional retroelement silencing and localization to pi-bodies in the *Drosophila* germline. *J Cell Biol*. 2009;186:333–42.
71. Rouget C, Papin C, Boureux A, Meunier AC, Franco B, Robine N, Lai EC, Pelisson A, Simonelig M. Maternal mRNA deadenylation and decay by the piRNA pathway in the early *Drosophila* embryo. *Nature*. 2010;467:1128–32.
72. Gou LT, Dai P, Yang JH, Xue Y, Hu YP, Zhou Y, Kang JY, Wang X, Li H, Hua MM, et al. Pachytene piRNAs instruct massive mRNA elimination during late spermiogenesis. *Cell Res*. 2014;24:680–700.
73. Shpiz S, Ryazansky S, Olovnikov I, Abramov Y, Kalmykova A. Euchromatic transposon insertions trigger production of novel Pi- and endo-siRNAs at the target sites in the *drosophila* germline. *PLoS Genet*. 2014;10:e1004138.
74. Klein JD, Qu C, Yang X, Fan Y, Tang C, Peng JC. c-Fos repression by Piwi regulates *Drosophila* ovarian germline formation and tissue morphogenesis. *PLoS Genet*. 2016;12:e1006281.

75. Saito K, Inagaki S, Mituyama T, Kawamura Y, Ono Y, Sakota E, Kotani H, Asai K, Siomi H, Siomi MC. A regulatory circuit for piwi by the large Maf gene traffic jam in *Drosophila*. *Nature*. 2009;461:1296–9.
76. Vourekas A, Zheng Q, Alexiou P, Maragkakis M, Kirino Y, Gregory BD, Mourelatos Z. Mili and Miwi target RNA repertoire reveals piRNA biogenesis and function of Miwi in spermiogenesis. *Nat Struct Mol Biol*. 2012;19:773–81.
77. Deng W, Lin H. miwi, a murine homolog of piwi, encodes a cytoplasmic protein essential for spermatogenesis. *Dev Cell*. 2002;2:819–30.
78. Kotaja N, Bhattacharyya SN, Jaskiewicz L, Kimmins S, Parvinen M, Filipowicz W, Sassone-Corsi P. The chromatoid body of male germ cells: similarity with processing bodies and presence of Dicer and microRNA pathway components. *Proc Natl Acad Sci U S A*. 2006;103:2647–52.
79. Goh WS, Falcatori I, Tam OH, Burgess R, Meikar O, Kotaja N, Hammell M, Hannon GJ. piRNA-directed cleavage of meiotic transcripts regulates spermatogenesis. *Genes Dev*. 2015;29:1032–44.
80. Zhang P, Kang JY, Gou LT, Wang J, Xue Y, Skogerboe G, Dai P, Huang DW, Chen R, Fu XD, et al. MIWI and piRNA-mediated cleavage of messenger RNAs in mouse testes. *Cell Res*. 2015;25:193–207.
81. Wu PH, Fu Y, Cecchini K, Ozata DM, Arif A, Yu T, Colpan C, Gainetdinov I, Weng Z, Zamore PD. The evolutionarily conserved piRNA-producing locus pi6 is required for male mouse fertility. *Nat Genet*. 2020;52:728–39.
82. Lim AK, Lorthongpanich C, Chew TG, Tan CW, Shue YT, Balu S, Gounko N, Kuramochi-Miyagawa S, Matzuk MM, Chuma S, et al. The nuage mediates retrotransposon silencing in mouse primordial ovarian follicles. *Development*. 2013;140:3819–25.
83. Aravin AA, Naumova NM, Tulin AV, Vagin VV, Rozovsky YM, Gvozdev VA. Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the *D. melanogaster* germline. *Curr Biol*. 2001;11:1017–27.
84. Kotov AA, Adashev VE, Godneeva BK, Ninova M, Shatskikh AS, Bazylev SS, Aravin AA, Olenina LV. piRNA silencing contributes to interspecies hybrid sterility and reproductive isolation in *Drosophila melanogaster*. *Nucleic Acids Res*. 2019;47:4255–71.
85. Shen EZ, Chen H, Ozturk AR, Tu S, Shirayama M, Tang W, Ding YH, Dai SY, Weng Z, Mello CC. Identification of piRNA binding sites reveals the Argonaute regulatory landscape of the *C. elegans* Germline. *Cell*. 2018;172:937–51 e918.
86. Zhang D, Tu S, Stubna M, Wu WS, Huang WC, Weng Z, Lee HC. The piRNA targeting rules and the resistance to piRNA silencing in endogenous genes. *Science*. 2018;359:587–92.
87. Barckmann B, Pierson S, Dufourt J, Papin C, Armenise C, Port F, Grentzinger T, Chambeyron S, Baronian G, Desvignes JP, et al. Aubergine iCLIP reveals piRNA-dependent decay of mRNAs involved in germ cell development in the early embryo. *Cell Rep*. 2015;12:1205–16.
88. Rojas-Rios P, Chartier A, Pierson S, Simonelig M. Aubergine and piRNAs promote germline stem cell self-renewal by repressing the proto-oncogene *Cbl*. *EMBO J*. 2017;36:3194–211.
89. Tang W, Seth M, Tu S, Shen EZ, Li Q, Shirayama M, Weng Z, Mello CC. A sex chromosome piRNA promotes robust dosage compensation and sex determination in *C. elegans*. *Dev Cell*. 2018;44:762–70 e763.
90. Kiuchi T, Koga H, Kawamoto M, Shoji K, Sakai H, Arai Y, Ishihara G, Kawaoka S, Sugano S, Shimada T, et al. A single female-specific piRNA is the primary determiner of sex in the silkworm. *Nature*. 2014;509:633–6.
91. Balaratnam S, West N, Basu S. A piRNA utilizes HILI and HIWI2 mediated pathway to down-regulate ferritin heavy chain 1 mRNA in human somatic cells. *Nucleic Acids Res*. 2018;46:10635–48.
92. Sytnikova YA, Rahman R, Chirn GW, Clark JP, Lau NC. Transposable element dynamics and PIWI regulation impacts lncRNA and gene expression diversity in *Drosophila* ovarian cell cultures. *Genome Res*. 2014;24:1977–90.
93. Harris AN, Macdonald PM. Aubergine encodes a *Drosophila* polar granule component required for pole cell formation and related to eIF2C. *Development*. 2001;128:2823–32.
94. Ramat A, Garcia-Silva MR, Jahan C, Nait-Saidi R, Dufourt J, Garret C, Chartier A, Cremaschi J, Patel V, Decourcelle M, et al. The PIWI protein Aubergine recruits eIF3 to activate translation in the germ plasm. *Cell Res*. 2020;30:421–35.
95. Dai P, Wang X, Gou LT, Li ZT, Wen Z, Chen ZG, Hua MM, Zhong A, Wang L, Su H, et al. A translation-activating function of MIWI/piRNA during mouse spermiogenesis. *Cell*. 2019;179:1566–81 e1516.
96. Unhavaithaya Y, Hao Y, Beyret E, Yin H, Kuramochi-Miyagawa S, Nakano T, Lin H. MILI, a PIWI-interacting RNA-binding protein, is required for germ line stem cell self-renewal and appears to positively regulate translation. *J Biol Chem*. 2009;284:6507–19.
97. Grivna ST, Pyhtila B, Lin H. MIWI associates with translational machinery and PIWI-interacting RNAs (piRNAs) in regulating spermatogenesis. *Proc Natl Acad Sci U S A*. 2006;103:13415–20.
98. Dufourt J, Bontonou G, Chartier A, Jahan C, Meunier AC, Pierson S, Harrison PF, Papin C, Beilharz TH, Simonelig M. piRNAs and Aubergine cooperate with Wispoly poly(A) polymerase to stabilize mRNAs in the germ plasm. *Nat Commun*. 2017;8:1305.
99. Kelley D, Rinn J. Transposable elements reveal a stem cell-specific class of long noncoding RNAs. *Genome Biol*. 2012;13:R107.
100. He X, Chen X, Zhang X, Duan X, Pan T, Hu Q, Zhang Y, Zhong F, Liu J, Zhang H, et al. An lnc RNA (GAS5)/SnoRNA-derived piRNA induces activation of TRAIL gene by site-specifically recruiting MLL/COMPASS-like complexes. *Nucleic Acids Res*. 2015;43:3712–25.
101. Li XZ, Roy CK, Dong X, Bolcun-Filas E, Wang J, Han BW, Xu J, Moore MJ, Schimenti JC, Weng Z, Zamore PD. An ancient transcription factor initiates the burst of piRNA production during early meiosis in mouse testes. *Mol Cell*. 2013;50:67–81.
102. Farley BM, Collins K. Transgenerational function of *Tetrahymena* Piwi protein Twi8p at distinctive noncoding RNA loci. *RNA*. 2017;23:530–45.
103. Zhang Z, Harrison PM, Liu Y, Gerstein M. Millions of years of evolution preserved: a comprehensive catalog of the processed pseudogenes in the human genome. *Genome Res*. 2003;13:2541–58.
104. Pantano L, Jodar M, Bak M, Balleca JL, Tommerup N, Oliva R, Vavouri T. The small RNA content of human sperm reveals pseudogene-derived piRNAs complementary to protein-coding genes. *RNA*. 2015;21:1085–95.
105. Tan L, Mai D, Zhang B, Jiang X, Zhang J, Bai R, Ye Y, Li M, Pan L, Su J, et al. PIWI-interacting RNA-36712 restrains breast cancer progression and chemoresistance by interaction with SEPW1 pseudogene SEPW1P RNA. *Mol Cancer*. 2019;18:9.
106. Couvillion MT, Lee SR, Hogstad B, Malone CD, Tonkin LA, Sachidanandam R, Hannon GJ, Collins K. Sequence, biogenesis, and function of diverse small RNA classes bound to the Piwi family proteins of *Tetrahymena thermophila*. *Genes Dev*. 2009;23:2016–32.
107. Hirano T, Iwasaki YW, Lin ZY, Imamura M, Seki NM, Sasaki E, Saito K, Okano H, Siomi MC, Siomi H. Small RNA profiling and characterization of piRNA clusters in the adult testes of the common marmoset, a model primate. *RNA*. 2014;20:1223–37.

108. Gebert D, Ketting RF, Zischler H, Rosenkranz D. piRNAs from pig testis provide evidence for a conserved role of the Piwi pathway in post-transcriptional gene regulation in mammals. *PLoS One*. 2015;10:e0124860.
109. Parrish NF, Fujino K, Shiromoto Y, Iwasaki YW, Ha H, Xing J, Makino A, Kuramochi-Miyagawa S, Nakano T, Siomi H, et al. piRNAs derived from ancient viral processed pseudogenes as transgenerational sequence-specific immune memory in mammals. *RNA*. 2015;21:1691–703.
110. Halbach R, Miesen P, Joosten J, Taskopru E, Rondeel I, Pennings B, Vogels CBF, Merklings SH, Koenraadt CJ, Lambrechts L, van Rij RP. A satellite repeat-derived piRNA controls embryonic development of *Aedes*. *Nature*. 2020;580:274–7.
111. Yin H, Lin H. An epigenetic activation role of Piwi and a Piwi-associated piRNA in *Drosophila melanogaster*. *Nature*. 2007;450:304–8.
112. Smulders-Srinivasan TK, Szakmary A, Lin H. A *Drosophila* chromatin factor interacts with the Piwi-interacting RNA mechanism in niche cells to regulate germline stem cell self-renewal. *Genetics*. 2010;186:573–83.
113. Radion E, Morgunova V, Ryazansky S, Akulenko N, Lavrov S, Abramov Y, Komarov PA, Glukhov SI, Olovnikov I, Kalmykova A. Key role of piRNAs in telomeric chromatin maintenance and telomere nuclear positioning in *Drosophila* germline. *Epigenetics Chromatin*. 2018;11:40.
114. Shpiz S, Olovnikov I, Sergeeva A, Lavrov S, Abramov Y, Savitsky M, Kalmykova A. Mechanism of the piRNA-mediated silencing of *Drosophila* telomeric retrotransposons. *Nucleic Acids Res*. 2011;39:8703–11.
115. Savitsky M, Kwon D, Georgiev P, Kalmykova A, Gvozdev V. Telomere elongation is under the control of the RNAi-based mechanism in the *Drosophila* germline. *Genes Dev*. 2006;20:345–54.
116. Pane A, Wehr K, Schupbach T. zucchini and squash encode two putative nucleases required for rasiRNA production in the *Drosophila* germline. *Dev Cell*. 2007;12:851–62.
117. Simmons MJ, Ryzek DF, Lamour C, Goodman JW, Kummer NE, Merriman PJ. Cytotype regulation by telomeric P elements in *Drosophila melanogaster*: evidence for involvement of an RNA interference gene. *Genetics*. 2007;176:1945–55.
118. Josse T, Teyssset L, Todeschini AL, Sidor CM, Anxolabehere D, Ronsseray S. Telomeric trans-silencing: an epigenetic repression combining RNA silencing and heterochromatin formation. *PLoS Genet*. 2007;3:1633–43.
119. Teo RYW, Anand A, Sridhar V, Okamura K, Kai T. Heterochromatin protein 1a functions for piRNA biogenesis predominantly from pericentric and telomeric regions in *Drosophila*. *Nat Commun*. 2018;9:1735.
120. Kordyukova M, Sokolova O, Morgunova V, Ryazansky S, Akulenko N, Glukhov S, Kalmykova A. Nuclear Ccr4-Not mediates the degradation of telomeric and transposon transcripts at chromatin in the *Drosophila* germline. *Nucleic Acids Res*. 2020;48(1):141–56.
121. Khurana JS, Xu J, Weng Z, Theurkauf WE. Distinct functions for the *Drosophila* piRNA pathway in genome maintenance and telomere protection. *PLoS Genet*. 2010;6:e1001246.
122. Asif-Laidin A, Delmarre V, Laurentie J, Miller WJ, Ronsseray S, Teyssset L. Short and long-term evolutionary dynamics of subtelomeric piRNA clusters in *Drosophila*. *DNA Res*. 2017;24:459–72.
123. Frenk S, Lister-Shimauchi EH, Ahmed S. Telomeric small RNAs in the genus *Caenorhabditis*. *RNA*. 2019;25:1061–77.
124. Story B, Ma X, Ishihara K, Li H, Hall K, Peak A, Anoja P, Park J, Haug J, Blanchette M, Xie T. Defining the expression of piRNA and transposable elements in *Drosophila* ovarian germline stem cells and somatic support cells. *Life Sci Alliance*. 2019;2(5) e201800211.
125. Nishida KM, Saito K, Mori T, Kawamura Y, Nagami-Okada T, Inagaki S, Siomi H, Siomi MC. Gene silencing mechanisms mediated by Aubergine piRNA complexes in *Drosophila* male gonad. *RNA*. 2007;13:1911–22.
126. Nagao A, Mituyama T, Huang H, Chen D, Siomi MC, Siomi H. Biogenesis pathways of piRNAs loaded onto AGO3 in the *Drosophila* testis. *RNA*. 2010;16:2503–15.
127. Hsieh CL, Xia J, Lin H. MIWI prevents aneuploidy during meiosis by cleaving excess satellite RNA. *EMBO J*. 2020;39:e103614.
128. Yadav RP, Makela JA, Hyssala H, Cisneros-Montalvo S, Kotaja N. DICER regulates the expression of major satellite repeat transcripts and meiotic chromosome segregation during spermatogenesis. *Nucleic Acids Res*. 2020;48:7135–53.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

