Small RNAs just got bigger: Piwi-interacting RNAs (piRNAs) in mammalian testes

V. Narry Kim

School of Biological Sciences and Institute of Molecular Biology and Genetics, Seoul National University, Seoul, 151-742, Korea

Small RNAs constitute a large family of regulatory molecules with diverse functions in eukaryotes. Hallmarks of small RNAs are their dependence on double-stranded RNAs (dsRNA)-specific RNase III-type enzymes for biogenesis and their association with Argonaute family proteins for the silencing process. At least two classes of small RNAs have previously been described: microRNAs (miRNAs) derived from hairpin-shaped precursors and small interfering RNAs (siRNAs) generated from long dsRNAs. Recent articles reported a novel class of small RNAs that are expressed specifically and abundantly in the spermatogenic cells of mice. These RNAs are bigger (26–31 nucleotides [nt]) than most previously described small RNAs (21–23 nt) and are associated with Piwi-subclade members of the Argonaute protein family. Although the biogenesis and function of these RNAs are yet to be determined, these findings may add new dimensions in small RNA biology and germline cell biology.

Different classes of small RNAs

Eukaryotes produce various types of small RNAs of 19–30 nucleotides (nt) in length that function in diverse pathways (Bartel 2004; Kim 2005; Brodersen and Voinnet 2006; Vaucheret 2006). Small RNAs act as guides to direct mRNA degradation, translational repression, heterochromatin formation, and DNA elimination. Because the active forms of small RNAs are often indistinguishable biochemically or functionally in some species, they are conventionally grouped based on their origins into two classes: microRNAs (miRNAs) and small interfering RNAs (siRNAs) [Fig. 1]. miRNAs are generated from the dsRNA region of the hairpin-shaped precursors, whereas siRNAs are derived from long double-stranded RNAs (dsRNAs). Both miRNAs and siRNAs bind to mRNA and induce mRNA cleavage, translational repression, and cleavage-independent mRNA decay.

Several groups of natural siRNAs [Nat-siRNAs] have been described in various eukaryotic species. Nat-siRNAs in plants and endogenous siRNAs [endo-siRNAs] in nematodes are generated from dsRNAs derived from endogenous sense and antisense transcripts [Borsani et al. 2005; Lee et al. 2006]. Trans-acting siRNAs [tasiRNAs] in plants are processed from dsRNAs that are synthesized by an RNA-dependent RNA polymerase using endogenous transcripts as the templates [Peragine et al. 2004]. Endogenous tasiRNAs direct cleavage of endogenous cognate mRNAs in trans [the target genes are different from the gene that the siRNA originates]. Tiny noncoding RNAs [tncRNAs] are tasiRNA-like RNAs found in nematode worms [Ambros et al. 2003; Lee et al. 2006]. Small scan RNAs [scnRNAs] are thought to scan for regions in DNA for genome rearrangement in Tetrahymena thermophila [Mochizuki et al. 2002; Taverna et al. 2002; Liu et al. 2004]. The Argonaute family protein, Twi1p, mediates the process and a Dicer homolog is required for the accumulation of scnRNAs [Mochizuki et al. 2002]. Repeat-associated siRNAs [rasiRNAs], on the other hand, are roughly defined as small RNAs that match repetitive sequence elements in sense and antisense orientations [Aravin et al. 2001, 2003, 2004; Djikeng et al. 2001; Hall et al. 2002; Hamilton et al. 2002; Llave et al. 2002; Mette et al. 2002; Pal-Bhadra et al. 2002; Reinhart and Bartel 2002; Volpe et al. 2002]. RasiRNAs are presumably derived from long dsRNAs, but their biogenesis mechanism remains unclear. RasiRNAs may function in the establishment of heterochromatin in repetitive elements.

In mammals, only the miRNA class has been identified, and there have been no descriptions of any other class of endogenous small RNAs. This review will summarize recent findings of a previously unknown class of small RNAs specifically expressed in mammalian germine cells [Aravin et al. 2006; Girard et al. 2006; Grivna et al. 2006; Watanabe et al. 2006].

Piwi subfamily proteins and spermatogenesis

Argonaute family proteins are the prime components of small RNA complexes that are highly conserved among
Spermatogenesis can be divided into three phases. The first phase is mitotic self-renewing of germline stem cells, producing primary spermatogonia. The second phase is the meiosis of primary spermatocytes, which generates haploid round spermatids. The prophase of the first meiotic division progresses from the leptotene to the pachytene stage to early round spermatids.

Their expression overlaps in the mid–pachytene stage. Spermatogenetic arrest points are also different in MILI-null and MIWI-null mice (Fig. 2). MILI-null mice are arrested at the early pachytene spermatocyte phase, whereas MIWI-null mice are arrested at the round spermatid stage.

It is unclear whether *Drosophila piwi* is the ortholog of mammalian Piwi proteins. Fly PIWI is expressed in the nuclei of germine cells as well as of the supporting somatic cells, whereas MIWI and MILI are found in the cytoplasm of germ cells (Cox et al. 1998). In addition, fly *piwi* mutant shows defects not only in spermatogenesis but also in the maintenance of germine stem cells (Cox et al. 1998). *Drosophila aub*, another member of the *piwi* subfamily, may be a closer homolog to *Miwi* and *Mili* because *AUB* is found in the cytoplasm of spermatogonia and spermatocytes. Loss of *aub* function leads to abnormal development of spermatocytes and round spermatids (Schmidt et al. 1999).

**Discovery of piRNAs**

Piwi-interacting RNAs (piRNAs) were recently isolated from mouse testes independently by four groups (Aravin et al. 2006; Girard et al. 2006; Grivna et al. 2006; Watanabe et al. 2006). In three laboratories, the initial recognition of ∼30-nt piRNAs was made by simply looking at the total testis RNA on ethidium bromide (or SYBR green)-stained gel because piRNAs are much more abundant in testes than other small RNA species (Aravin et al. 2006; Girard et al. 2006; Grivna et al. 2006; Watanabe et al. 2006). In fact, it was later estimated that approximately one million piRNA molecules exist per spermatocyte or round spermatid (Aravin et al. 2006). Subsequent cloning led to the discovery of numerous small RNAs that range from 26 to ∼31 nt.

With an aim of identifying MIWI-associated small RNA, Hannon and colleagues (Girard et al. 2006) first visualized testes RNA by pCp labeling at the 3′ end (Grivna et al. 2006). Intrigued by the abundant RNAs of unexpected size [29–30 nt], they went on to sequence the RNAs. Apart from conventional cDNA cloning method, Hannon and colleagues also employed highly sensitive piRNA libraries.

**Figure 1.** Classification of small RNAs. Definition and classification of small RNAs conventionally relies on their biogenesis mechanism. Two relatively well-defined classes of small RNAs include micro RNAs (miRNAs) and small interfering RNAs (siRNAs). The biogenesis mechanism for piRNAs is currently unknown.

**Figure 2.** Expression of piRNAs and Piwi subfamily members during mouse spermatogenesis. Mammalian spermatogenesis occurs within the seminiferous epithelium inside the seminiferous tubule and can be divided into three phases.
parallel pyrosequencing technology to obtain >50,000 RNA sequences for which they coined the name “piwi-interacting RNAs” (piRNAs). Despite intensive sequencing, many sequences were found only once, indicating that the piRNA population is extremely diverse.

Similarly, Tuschl and colleagues (Aravin et al. 2006) investigated MILI-associated RNAs. By immunoprecipitation from testis lysate using anti-MILI antibody, they found that 26- to 28-nt RNAs are associated with MILI and that there are additional 29- to 31-nt RNA species that do not interact with MILI. Subsequent cloning resulted in >15,000 sequences.

Imai and colleagues (Watanabe et al. 2006) cloned small RNAs not only from mouse testes but also from mouse oocytes. Cloning of testis small RNA yielded 381 clones of ~27-nt RNAs. From oocytes, they isolated seven miRNAs and 40 potential rasiRNAs that are ~22 nt in length. This suggests that piRNAs may have a specific role in male germline, although it is possible that oocyte-specific RNAs were underrepresented in the sample due to the limited number of oocytes in mouse ovaries.

Lin and colleagues (Grivna et al. 2006) noticed that the ~30-nt RNA band was not detectable in their Miwi-null mice (Girard et al. 2006). The investigators then cloned 29- to 34-nt RNA isolated from total testicular RNA and discovered 40 different RNAs ranging from 25 to 35 nt [mainly 29–31 nt]. Among these, two overlapped with the Imai set (Watanabe et al. 2006). These RNAs were precipitated specifically with anti-MIWI antibody, indicating that MIWI is associated with this class of RNA and is required for their accumulation.

Size distribution of cloned RNAs indicated that there are two distinct populations of piRNAs (Aravin et al. 2006; Girard et al. 2006; Grivna et al. 2006; Watanabe et al. 2006). piRNAs of the first group are 29–31 nt in length and associated with MIWI protein. Consistent with their dependence on Miwi, their expression coincides with that of MIWI, which is from the mid–pachytene to the early round spermatid stage (Fig. 2). This group of piRNAs can be detected from day 15 mice but does not reach the maximal level until day 18. The second group includes slightly shorter piRNAs (26–28 nt in length) that are preferentially associated with MILI (Fig. 2; Aravin et al. 2006). Detected abundantly by day 14, MILI-associated RNAs are expressed as early as from spermatogonia and begin to diminish after the round spermatid stage. In adult mouse testes, the first group (29- to 31-nt piRNA) is more abundant than the second group (26–28 nt). These two groups may be generated from the same loci via differential processing because the probes for some piRNAs can detect both ~30-nt band and ~26-nt band (Aravin et al. 2006; Watanabe et al. 2006). Determination of the termini by rapid amplification of cDNA ends (RACE) also revealed that the 5’ ends of piRNA clones were invariable while the 3’ ends were shorter by 2 nt in MILI-associated RNA clones (Aravin et al. 2006). Whether these two groups of RNAs play distinct roles remains to be determined.

Genomic distribution and characteristics of piRNAs

Cloned piRNAs show rather uneven distribution among chromosomes (Aravin et al. 2006; Girard et al. 2006; Grivna et al. 2006; Watanabe et al. 2006). They are enriched in chromosomes 17, 5, 4, and 2 but seem to be largely excluded from sex chromosomes. They are slightly enriched in intergenic regions that are poor in genes or repeats. Although ~17% of piRNAs were mapped to repeat sequences [mostly corresponding to retrotransposons], this proportion is lower than expected by random sampling of 30-nt sequences from the genome (~40%). Imai and colleagues (Watanabe et al. 2006) classified these repeat-matching RNAs in testes into rasiRNAs. However, judging from their size distribution, at least some of these repeat-matching RNAs may need to be recategorized into the piRNA class.

Because small RNAs are usually defined and classified based on their dsRNA origin, one may argue that these RNAs may be the degradation products from large transcripts or belong to an unrelated RNA family. However, the following characteristics strongly advocate that piRNAs are authentic members of the small RNA family. Firstly, their lengths distribute in a significantly narrow range. Northern blotting also showed discrete bands representing either 26- to 28-nt or 29- to 31-nt RNAs. Secondly, as mentioned earlier, they are physically associated with Argonaute family proteins, which are the key factors in RNA silencing pathways. Thirdly, their 5’ ends have a strong preference for uracil (~86%) that is characteristic of small RNAs. For instance, ~76% of miRNAs have U at the 5’-most position. The reason for this sequence bias remains unclear, but it has been proposed that RNA processing factors such as Dicer and Drosha cleave preferentially at the 5’ side of U.

Figure 3. Strand bias of piRNAs. The first and second tracks indicate piRNAs mapping to the negative and positive strands, respectively. Monodirectional clusters encode piRNAs mainly on only one strand, whereas bidirectional clusters encode piRNAs on two strands that are segregated into two separate parts.
Unanswered questions and future directions

It remains elusive how these 26- to 35-nt RNAs are generated. So far, there is no evidence that these RNAs are generated from dsRNA precursor. The strong strand bias of clustered piRNAs suggests that they may be generated from the single-stranded precursor. Supporting this notion, both EST analysis and RT-PCR experiments revealed putative primary transcripts but failed to detect the antisense transcripts [Watanabe et al. 2006]. Unlike miRNA precursors, the regions covering piRNA clones do not fold into stem–loop structures. Therefore, the biogenesis pathway of piRNAs appears to be distinct from that of miRNAs and siRNAs. Several possible mechanisms have been proposed for piRNA biogenesis. First, long-range dsRNA structures may be formed in primary transcripts, which may explain the strand bias of piRNAs [Aravin et al. 2006]. Second, the antisense transcript may be expressed at such a low level that they escape detection [Watanabe et al. 2006]. Third, primary transcripts may be digested by a ssRNA-directed unidentified enzyme. Fourth, Imai and colleagues [Watanabe et al. 2006] proposed that piRNAs may be derived from RNA-DNA duplexes that are produced by the reverse transcriptase of retrotransposons.

Although it was previously shown that human Piwi proteins can interact with Dicer [Sasaki et al. 2003], it has yet to be determined whether RNase III-type enzymes such as Dicer and Drosha are indeed involved in piRNA biogenesis. Conditional knockout of Dicer and Drosha will help answer this question. Biochemical analysis of Piwi complex will also be required to identify the biogenesis factors for piRNA.

Perhaps the most crucial question is as to the function of piRNAs. Because null mice of Miwi and Mili show clear defects in spermatogenic cells, piRNAs may be involved in spermatogenesis possibly by regulating meiosis and/or suppressing retrotransposons. One possibility is that piRNAs act like miRNAs by guiding the Piwi complex to their target miRNAs whose activity is required for spermatogenesis regulation. However, most piRNAs are not complementary to miRNA, suggesting that they are unlikely to function as post-translational regulators of protein synthesis. In fact, it is conceivable that the target molecules may not be RNA molecules. Interestingly, individual piRNA sequences are not conserved, whereas piRNA loci are conserved in the syntenic regions of other mammalian species [Girard et al. 2006]. This led to the suggestion that piRNA generation process itself from these loci may be significant for the function [Girard et al. 2006]. Identification of the target molecules of the piRNA–Piwi complex will be a crucial step toward understanding the functions of piRNAs. Genetic analyses on these loci will be necessary to elucidate the function of piRNAs. It will also be imperative to reveal the other protein components of the piRNA–Piwi complex. Biochemical and cell biological analysis of Piwi proteins will provide insights into these enigmatic and intriguing RNA species.

Acknowledgments

I thank Inha Heo and Dong-Hyeok Kim for their assistance with manuscript preparation. This work was supported by the Basic Research Program [R02-2004-000-10173-0], the National Research Laboratory Program [M1050000010905000010910], the Korea Research Foundation Grant (KRF-2005-005-J16001), and the BK21 program.

Note added in proof

Recent work by Kingston and colleagues [Lau et al. 2006] also described piRNAs from rat testes. Rat piRNAs are associated with the piRNA complex [piRC], which contains Riwi, the rat homolog to Piwi, as well as RRecQ1, a homolog to QDE-3 from Neurospora that has been implicated in RNA silencing pathways.

References


Girard, A., Sachidanandam, R., Hannon, G.J., and Carmell, M.A.