A Novel Class of Somatic Small RNAs Similar to Germ Cell Pachytene PIWI-interacting Small RNAs*‡

Nicole Ortogero‡, Andrew S. Schuster‡, Daniel K. Oliver‡, Connor R. Riordan‡, Annie S. Hong‡, Grant W. Hennig‡, Dickson Luong‡, Jianqiang Bao‡, Bhupal P. Bhetwal‡, Seungil Ro‡, John R. McCarrey‡, and Wei Yan§

From the ⁄Department of Physiology and Cell Biology, University of Nevada School of Medicine, Reno, Nevada 89557 and the §Department of Biology, University of Texas at San Antonio, San Antonio, Texas 78249

Background: Germ cells exclusively express PIWI-interacting small RNAs for transposon and gene regulation.

Results: Somatic cells express similar small RNAs that do not require known small RNA proteins and that partially complement mRNAs.

Conclusion: These somatic small RNAs represent a novel small RNA population, which potentially regulates mRNA translation.

Significance: Defining novel small RNAs is essential for elucidating the mechanisms that control gene expression.

PIWI-interacting RNAs (piRNAs) are small noncoding RNAs that bind PIWI family proteins exclusively expressed in the germ cells of mammalian gonads. MIWI2-associated piRNAs are essential for silencing transposons during primordial germ cell development, and MIWI-bound piRNAs are required for normal spermatogenesis during adulthood in mice. Although piRNAs have long been regarded as germ cell-specific, increasing lines of evidence suggest that somatic cells also express piRNA-like RNAs (pilRNAs). Here, we report the detection of abundant pilRNAs in somatic cells, which are similar to MIWI-associated piRNAs mainly expressed in pachytene spermatocytes and round spermatids in the testis. Based on small RNA deep sequencing and quantitative PCR analyses, pilRNA expression is dynamic and displays tissue specificity. Although pilRNAs are similar to pachytene piRNAs in both size and genomic origins, they have a distinct ping-pong signature. Furthermore, pilRNA biogenesis appears to utilize a yet to be identified pathway, which is different from all currently known small RNA biogenetic pathways. In addition, pilRNAs appear to preferentially target the 3′-UTRs of mRNAs in a partially complementary manner. Our data suggest that pilRNAs, as an integral component of the small RNA transcriptome in somatic cell lineages, represent a distinct population of small RNAs that may have functions similar to germ cell piRNAs.
interstitial cells of Cajal (ICC) from the murine gastrointestinal tract, and murine small intestine, in addition to millions of murine stomach small RNAs that align to newly defined germ cell piRNA clusters. Our in-depth analyses revealed a high degree of similarities between these somatic piRNAs and pachytene piRNAs in size, nucleotide composition, clustering, genomic origin, and potential targets. However, unlike pachytene piRNAs, piRNAs have a distinct ping-pong signature. Finally, biogenesis of piRNAs is mediated by an unidentified mechanism different from any of the currently known small RNA biogenetic pathways indicative of an entirely novel population of small RNAs.

EXPERIMENTAL PROCEDURES

Knock-out Mouse Lines Used—All animal work was conducted following the animal use protocol approved by the Institutional Animal Care and Use Committee of the University of Nevada, Reno. For FACS-based Sertoli cell purification, control (Amh-Cre; Rosa26-mTmG<sup>+/−</sup>), Sertoli cell-specific Dicer conditional knock-out (cKO) (Amh-Cre; Dicer<sup>lox/lox</sup>; Rosa26-mTmG<sup>+/−</sup>), and Sertoli cell-specific Drosha cKO (Amh-Cre; Drosha<sup>lox/lox</sup>; Rosa26-mTmG<sup>+/−</sup>) mice were generated as described (16–18). Miwi universal KO mouse line was obtained from the Mutant Mouse Regional Resource Centers. Mov10I universal knock-out mice were purchased from the Jackson Laboratory. Miwi2 knock-out mice were generated as described (19).

Tissue Collection and Cell Purification—Tissue samples (brain, colon, heart, kidney, liver, lung, small intestine, spleen, stomach, and testis) were collected from adult wild-type or knock-out C57/BL6 mice. Upon collection, samples were immediately snap frozen in liquid nitrogen followed by storage at −80 °C until RNA isolation. Wild-type Sertoli cells were collected from postnatal day 6 mice as previously reported (16). Amh-cre control and cKO Sertoli cells from 2-day-old mice were isolated by FACS based on GFP expression from the ROSA26-mTmG reporter (20). ICCs were purified from adult mice using FACS as described (21).

qPCR Analyses of Small Noncoding RNAs Expression—Small noncoding RNAs (sncRNAs) were isolated from purified Sertoli cells, ICCs, and small intestine, and sncRNA cDNAs were prepared as previously described (22). qPCR analyses were performed using the 7900HT Fast Real-Time PCR System (Applied Biosystems). A TaqMan-based method was implemented using TaqMan probe (5′-FAM-CTCGAGTCCACTACTGTC-MGB-3′ (where FAM is 6-carboxyfluorescein and MGB is minor groove binder)) and U6 snRNA as an endogenous control (23). The levels in all somatic tissues were normalized to those in testis and knock-out samples to respective wild-type control samples. Primers for sncRNA qPCR are included in Table 1.

Small RNA Sequencing—Sequencing libraries prepared using sncRNAs from Sertoli cells, ICCs, and small intestine were sequenced on the 454 GS20 platform (454 Life Sciences, Roche, Branford, CT) as previously reported (16). Sequencing data has been deposited to the NCBI Gene Expression Omnibus database with an accession number of GSE48388 (24). 454 Sertoli pilRNA annotations were previously published and can be found with accession number GSE40692 (16).

Two biological replicates of FACS-purified Sertoli cells from control, Dicer cKO, and Drosha cKO testes (n = 2) were sequenced using the Ion Torrent Personal Genome Machine (PGM) (Invitrogen) as described (17, 18). Briefly, sncRNAs were isolated using the mirVana miRNA isolation kit (Invitrogen), and cDNA libraries were generated using the Ion Total RNA-Seq kit v2 (Invitrogen). Emulsion PCR was carried out on the Ion OneTouch (Invitrogen) using the Ion OneTouch 200 template kit v2 (Invitrogen) for cDNA amplification on ion sphere particles. Samples were enriched for positive spheres using the Ion OneTouch enrichment system (Invitrogen). After enrichment, sequencing was performed on the PGM using the Ion PGM 200 sequencing kit (Invitrogen) and Ion 316 chips (Invitrogen). Stomach samples (n = 2) were sequenced in a similar fashion with the exception of 318 chips being used in lieu of 316 chips. Stomach data were submitted to the NCBI Gene Expression Omnibus under accession number GSE53780 (24).

pilRNA Annotation—For 454 sequencing reads, annotation analyses were conducted as described (16). Briefly, alignment to known small RNA libraries (miRBase, the functional RNA database (2010/10/04 16:15:05 GMT +9), the noncoding RNA database (v2.0), the Genomic tRNA database (downloaded January 5, 2011), and Repbase (v14.10)) allowing for up to two mismatches was performed with Sequery v1.0 (16, 26–30). Remaining unmatched sequences were aligned to the mouse genome build mm9 with a perfect match requirement in Sequery v1.0 (31). Sequences that aligned to the genome were categorized as novel endo-siRNA, miRNA, snoRNA, and pilRNA based on length, nucleotide composition, and secondary structures using Sequery v1.0, MIREAP 0.2, and snoREPORT 1.0 (32, 33). Remaining sequences that aligned to the genome but were not categorized were kept as unidentifiable small RNAs.

Sequence reads from Ion Torrent sequencing of Sertoli cell sncRNA libraries were aligned to the mouse genome (build mm9) using the Ion Torrent Server (Suite 2.2) t-map aligner.
Somatic piRNA-like RNA Characterization

| Cluster | Primer | Primer sequence | Expected amplicon size |
|---------|--------|-----------------|-----------------------|
| 2-qE1-35981 | Primer 1, forward | CCTTTCTCTGTTTCGAT | 580 bp |
| 2-qE1-35981 | Primer 1, reverse | AGTGAGAAATTCTGTTCTG | |
| 2-qE1-35981 | Primer 2, forward | CCTCTTTGCGGCTGCAAGA | 628 bp |
| 2-qE1-35981 | Primer 2, reverse | AAAAGCGACTGTTGGAGCACTG | |
| 9-C-31469 | Primer 1, forward | CTTGCGAATCTAATGAAGA | 721 bp |
| 9-C-31469 | Primer 1, reverse | TCCACAGCCTATCAAGATCTT | |
| 9-C-31469 | Primer 2, forward | AGGTTGGGACTGTCGACC | 983 bp |
| 9-C-31469 | Primer 2, reverse | GCTTCTGCAGGTTT7AGATCA | |
| 17-qA3.3-27363 | Primer 1, forward | CACGTGTCAAGCGGAGACTAG | 793 bp |
| 17-qA3.3-27363 | Primer 1, reverse | CCAACATACGCATTCTAA | |
| 17-qA3.3-27363 | Primer 2, forward | CGAGGAGCAAGGACTT | 919 bp |
| 17-qA3.3-27363 | Primer 2, reverse | TGGTCTAAGACGAGGCTGGA | |
| 17-qA3.3-26735 | Primer 1, forward | ACCATCGCATTCTAAGA | 695 bp |
| 17-qA3.3-26735 | Primer 1, reverse | GCCAAGCTGAGACCTCA | |
| 17-qA3.3-26735 | Primer 2, forward | GGTCTCTTATACGATCCACT | |
| 17-qA3.3-26735 | Primer 2, reverse | 17-qA3.3-26735 Primer 2, reverse |

(Invitrogen) (31). Known piRNA-Seq data analysis was performed using Partek Genom Suite (version 6.6; Partek, St. Louis, MO). Aligned reads were clustered into at least 15 nucleotide groups including overlapping reads. The resulting groups were aligned to mature piRNA sequences from the functional RNA database (2010/10/04 16:15:05 GMT +9), which contains a collection of small RNAs found associated with mouse PIWI proteins (30). Resulting alignments to known germ cell piRNAs were classified as pilRNAs. To account for sequencing depth variation, the abundance of each piRNA was normalized to the total aligned reads obtained from each sample, multiplied by variation, the abundance of each pilRNA was normalized to the total aligned reads obtained from each sample, multiplied by 10^6, and averaged between biological replicates. pilRNAs with ≤1 read in all six libraries were removed from analysis.

Resulting reads from PGM sequencing of wild-type stomach samples were mapped to mm9 using Ion Torrent Server Suite 3.2.1 t-map aligner (Invitrogen) (31). Recently redefined piRNA producing transcript genomic loci were identified by RNA-Seq on 1X DNase I digestion grade (Invitrogen) or RNase A (Sigma-Aldrich). DNase treatment was performed according to the manufacturer’s recommendations. RNase A treatment was performed at 37 °C for 60 min with 0.5 μg of RNase A in 1X DNase digestion buffer, followed by the same reaction conditions used for the DNase treatment (without the DNase). Using 1 μg of treated RNA, reverse transcription was performed using the SuperScript III first-strand synthesis system kit with random hexamers (Invitrogen). A reaction volume of 20 μl was used, consisting of 10 μl of 2X GoTaq (Promega), 10 pmol of forward and reverse primer, and 1 μl of CDNA. PCR conditions were as follows: 1 cycle at 95 °C/2min, 28 cycles of amplification (95 °C for 30 s, 58 °C for 30 s, and 72 °C for 45 s), followed by 1 cycle of 72 °C for 5 min. PCR product was run on a 2% agarose gel with the 100-bp DNA ladder H3 (Caisson Labs). Primers and anticipated amplicon sizes for the precursor transcript RT-PCR are included in Table 2.

PiRNA Post-Transcription Target Analysis—Annotated piRNA sequences were reverse complement aligned to all known murine 3’-UTRs, 5’-UTRs, and coding sequences (CDSs) using Sequery v1.0 (16). 3’-UTRs, 5’-UTRs, and CDSs were collected from ENSEMBL release 71 (35). Reverse complement analysis was first performed allowing up to two mismatches and then repeated allowing up to five mismatches with no gaps allowed. Gene features aligned to were considered potential post-transcriptional targets of pilRNAs.

Ping-Pong Analysis—Reads were subjected to a pairwise analysis of 5’ ends with a 10- or 11-nucleotide overlap using Sequery v1.0 (16). Reads that reverse complemented at their 5’ ends with a 10-nt overlap were considered ping-pong pilRNAs.

RESULTS

Identification of pilRNAs in Somatic Cells/Tissues—We initially searched for pilRNAs from sequence reads of three small noncoding RNA (sncRNA) libraries prepared using murine Sertoli cells purified from postnatal day 6 testes, ICCs of the murine gastrointestinal tract, and small intestine based on criteria described previously (16). Among 696 nonredundant pilRNAs identified, 117 were from the ICC sncRNA library, 263 were found in Sertoli cells, and 316 were found in the small intestine; ~40–60% of these pilRNAs have sequences identical to known germ cell piRNAs. The abundance of pilRNAs varied among the three somatic cell sncRNA libraries, with relatively more piRNAs identified in ICCs than in small intestine or Sertoli cells (Fig. 1A and supplemental Tables S1–S4). The size of germ cell piRNAs ranges from 24 to 32 nt with MILI-, MIWI2-, and MIWI-associated piRNAs being ~26, 28, and 30 nt long, respectively (6). The lengths of identified pilRNAs ranged from 20 to 32 nt with the majority at 30 nt (Fig. 1B), which is similar to the major size of pachytene/MIWI-associated piRNAs.

Recent data redefined murine genomic loci of piRNA-producing transcripts using a series of sequencing techniques including small RNA-Seq, sequencing of samples enriched for pilRNAs based on 2’-O-methyl-modified 3’ termini, cap analysis of gene expression sequencing, and polyadenylation site
These techniques clarified the termini of transcripts that produced piRNAs and were combined with temporal analyses allowing for the classification of piRNA clusters as pre-pachytene, hybrid, or pachytene piRNAs (34). Using these redefined clusters and our recent stomach small RNA deep sequencing data, we were able to map nearly 1.6 million stomach sncRNA reads to these germ cell piRNA clusters (Fig. 1C and supplemental Table S5). This piRNA population was the largest classifiable population, taking up 18% of the stomach small RNA sequencing reads. Coinciding with our initial piRNA results, stomach piRNAs aligned to the redefined germ cell piRNA clusters were predominantly 30 nt in length, similar to the pachytene/MIWI-associated piRNAs (Fig. 1D). Stomach replicates had a $R^2$ of 0.98, showing good correlation and corroboration of the stomach piRNA results (Fig. 1E).

**Validation of piRNA Expression in Somatic Tissues**—We verified the expression of three novel piRNAs identified from murine small intestine (piR-in3, piR-in18, and piR-in87) in 10 different tissues using qPCR (Fig. 2). As expected, the two germ cell-expressed piRNAs (piR-118029 and piR-126541) were predominantly detected in the testis. Interestingly, all three small intestine piRNAs also showed higher expression in the testis than in most of the somatic tissues analyzed. The highest levels of piRNA-in3 were detected in brain, whereas piRNA-in18...
was most abundant in heart. The top three organs expressing pilRNA-in87 included testis, brain, and heart (Fig. 2).

**pilRNAs Resemble Pachytene piRNAs**—Repeat-associated pilRNAs are produced through the ping-pong mechanism involving MILI and MIWI2 in mice (6, 9, 11). These pilRNAs include both primary (contain 5' uracil) and secondary (contain 10th position adenine) piRNA transcripts, and there is a consistent 10-nt complementary overlap between primary and secondary transcripts (11). piRNAs produced through the ping-pong mechanism tend to display an accumulation of secondary piRNAs, which contain an adenine at the 10th nt position (6).

We analyzed the pilRNA sequences and found that the 10th nt preference for adenine was absent. In addition, a strong preference for the signature 5' uracil of primary piRNA transcripts was detected in over 50% of pilRNAs (Fig. 3A). Stomach pilRNAs aligned to redefined germ cell piRNA loci confirmed this data with most reads preferentially containing a 5' uracil opposed to the 10th nt adenine (Fig. 3B). The nucleotide preferences of novel pilRNAs and pilRNAs that align to known piRNA sequences were investigated separately in ICCs, Sertoli cells, and small intestine (Fig. 4, A–C, respectively). A bias toward a 5' uracil and a lack of a 10th nt adenine preference was observed in both novel pilRNAs and pilRNAs that align to known piRNA sequences. Together, these data suggest that...
piRNAs resemble pachytene piRNAs and are predominantly primary piRNAs.

**piRNAs Tend to Cluster with Pachytene piRNAs**—Clustering is a key characteristic of piRNAs. However, lone piRNAs derived from unique, unclustered genomic locations have been found in large quantities as well (12, 14). To properly classify the piRNAs identified, we collected genomic coordinates containing murine pachytene and pachytene piRNA clusters, and piRNAs were aligned and categorized within these clusters (9, 12). 60 of 117 ICC piRNAs (~51%) were found to cluster with pachytene piRNA clusters, whereas none belonged to pre-pachytene piRNA clusters (Fig. 3C). Although 3 of 263 Sertoli cell piRNAs were found to belong to pachytene piRNA clusters, these 3, together with 31 additional Sertoli pilRNAs, also fell into pachytene piRNA clusters (Fig. 3C). Similarly, 117 of 316 small intestine pilRNAs (~37%) belonged to pachytene piRNA clusters, with only 2 also found in pre-pachytene piRNA clusters (Fig. 3C). The preferential clustering with pachytene piRNAs again supports the notion that pilRNAs are more like pachytene piRNAs. We confirmed that these results with stomach reads aligned to the redefined germ cell piRNA clusters. Of ~1.6 million, reads only 7,850 aligned to pachytene clusters and 27,945 to hybrid clusters, whereas 98% (1.54 million) of the stomach piRNAs aligned to pachytene clusters (Fig. 3D). Recent data have demonstrated that pachytene piRNA clusters are predominantly intergenic (34). This was found to be true for stomach piRNAs as well, with an average piRNA density of 15,800 piRNA per intergenic cluster in contrast to 422 per nonintergenic cluster (supplemental Table S5).

**piRNA Clusters Are Similar to Pachytene piRNA Clusters**—Like pachytene piRNAs, somatic piRNAs predominantly originate from intergenic genomic loci. Fig. 5A shows an example of a piRNA cluster from the stomach and testis, supporting the notion that pilRNAs are more like pachytene piRNAs. Unlike pachytene piRNAs, pilRNAs that originate from these clusters are largely 30 nt in length and have a preference for 5’ uracil similar to pachytene piRNAs (Fig. 5, B and C). RT-PCR detected the presence of several precursor transcripts in both the stomach and testis, suggesting that these clusters are also transcribed in somatic cells (Fig. 5D).

Although piRNAs Are Similar to Pachytene piRNAs, *piRNAs Exhibit Ping-Pong Activity*—To evaluate whether piRNAs are in fact products of the ping-pong mechanism, we performed pairwise analyses of the 5’ end of piRNA pairs. In the small intestine snRNA library, one single ping-pong match was identified between piR-141849 and piRNA-in87, which, in addition to having a 10-nt overlap, also had the signature 5’ uracil and 10th nt adenine. The ICC and Sertoli cells did not have any ping-pong signature detected. In contrast, the stomach had a large fraction (28%) of piRNAs that had the canonical 10-nt overlap at their 5’ ends (Fig. 6A). A similar analysis looking for 11-nt overlaps had markedly reduced alignments (7%), indicating that a 10-nt overlap of 5’ ends is indeed prominent in the stomach piRNAs. The stomach piRNAs that generate pairs with a 10-nt overlap also have a preference for a 5’ uracil, as well as the 10th nt adenine indicative of ping-pong activity (Fig. 6B). These ping-pong piRNAs predominantly align to pachytene clusters (Fig. 6C). In addition, the presence of secondary transcripts with 10th nt adenines do not seem to be dependent on the type of cluster the piRNA is derived from (Fig. 6D). In fact all ping-pong piRNAs, regardless of the cluster they originate from, have similar nucleotide profiles.

**piRNA Biogenesis Is Independent of Known miRNA or piRNA Biogenetic Pathways**—DICER and DROSHA are two RNase III enzymes essential for the biogenesis of the best-studied small RNAs, miRNAs (36, 37). In addition, DICER is also required for the production of endo-siRNAs (37). Because of their necessity in miRNA and/or endo-siRNA production, we examined the effects of DICER and DROSHA depletion on piRNA expression in murine Sertoli cells. Using Sertoli cells purified from Sertoli cell-specific Dicer (Amh-Cre; Dicer<sup>lox/lox</sup>; Rosa26-mTmG<sup>14/18</sup>) and Drosha (Amh-Cre; Drosha<sup>lox/lox</sup>; Rosa26-mTmG<sup>14/18</sup>) conditional knock-out testes, we performed

---

**FIGURE 4. Nucleotide preference of piRNAs that align to known piRNA and novel piRNAs.** A, ICC piRNAs show a 5’ uracil preference and no 10<sup>th</sup> nt preference in both groups. B, Sertoli cell piRNAs that align known piRNAs show a 5’ uracil preference for both groups but less so in novel piRNAs. Both groups show no 10<sup>th</sup> nt preference. C, small intestine pilRNAs show a 5’ uracil preference and no 10<sup>th</sup> nt preference in both groups. Nt, nucleotide; Sm Int, small intestine.

---

**Somatic piRNA-like RNA Characterization**

---

---

---
sncRNA deep sequencing (sncRNA-Seq) (17, 38). Sequencing reads aligned to known piRNAs were used for differential piRNA expression analyses. Although levels of Sertoli cell piRNAs were slightly altered, almost all piRNAs were still present at similar levels to control, as shown in the heat map representing the piRNA transcriptomes of control, Dicer- and Drosha-null Sertoli cells (Fig. 7, A and B, and supplemental Table S6). The persistent expression of piRNAs in the absence of either DICER or DROSHA suggests that piRNA production is independent of the two proteins. To see whether piRNA production is dependent on piRNA biogenetic pathways, we examined the expression of two piRNAs similar to known piRNAs in 10 organs of MIWI, MIWI2, and MOV10L1 universal KO mice (Fig. 7 C and supplemental Table S7). As expected, the two known piRNAs showed minimal or no expression in the testes of the three KO lines. The three novel piRNAs were also detected in testes but consistently decreased in the knock-outs. In comparison with the somatic tissues, testis samples had the largest decrease in expression between KO and controls. This is possibly due to the decrease in germ cell content in KO testis, a phenomenon observed previously in MIWI2 KO lines (19). Despite the fact that there was some differential piRNA expression between control and KO lines in all somatic tissues, piRNAs were still detected at variable levels, suggesting that their biogenesis in somatic cells does not rely on the known piRNA biogenetic pathways.

**piRNAs Can Target 3’-UTRs by Partial Annealing**—Post-transcriptional regulation is one of the most important functions of small RNAs, e.g. miRNAs and endo-siRNAs (39–41). miRNAs recognize their targets by partial complementary annealing to their target sequences mostly located in 3’-UTRs of mRNAs, whereas endo-siRNAs tend to bind to 3’-UTRs in a perfect complementary manner (40–42). If piRNAs also function as a post-transcriptional regulator in a manner similar to that of miRNAs and endo-siRNAs, we would expect to see that...
these piRNAs could bind mRNAs, especially 3’-UTRs. We therefore performed bioinformatic analyses by aligning ICC, Sertoli, and small intestine piRNAs in a reverse complementary manner to all known mouse 3’-UTRs, 5’-UTRs, and CDSs (35). We first allowed a 1–2-nt mismatch and found a small number of piRNAs (24 of 263 Sertoli cells and 5 of 316 small intestine piRNAs) could complementarily anneal to murine mRNAs with the majority (88% Sertoli cells and 80% small intestine) matched to 3’-UTRs (Fig. 8A and supplemental Tables S8 and S9). We then increased the mismatch allowance to 1–5 nt. Interestingly, a drastic increase in alignments was observed in piRNAs of all three tissue-origins, with 12 ICC, 96 Sertoli cells, and 66 small intestine piRNAs annealed to mRNA sequences (Fig. 8A and supplemental Tables S10–S12). The targeting sites were mainly located in 3’-UTRs, which displayed the highest percentage and were potentially targeted by 67% ICC, 79% Sertoli cells, and 62% small intestine piRNAs when one to five mismatches were allowed (Fig. 8A). An increase in CDS-targeting piRNAs was also noted when the mismatch allowance was increased to five, with many of the same piRNAs annealing to both the CDSs and 3’-UTRs of mRNAs. We examined the number of mRNAs that could potentially be targeted by piRNAs in a manner of partial complementary annealing, and a preference of potential piRNA binding sites in 3’-UTRs was observed (Fig. 8B). This was further exemplified by the fact that at least 2.4 times more mRNAs displayed potential 3’-UTR piRNA binding sites in comparison with those in 5’-UTRs and CDSs combined (Fig. 8B).

**DISCUSSION**

Given that three PIWI proteins are all expressed exclusively in germ cells in gonads, piRNAs have long been regarded as germ cell-specific (1, 3, 5–8). Several reports have documented the detection of piRNA-like RNAs in somatic cell types (13–16), but systematic characterization of piRNA structural features, genomic origins, biogenesis, and functions have not been reported. Approximately 6 years ago, we performed sncRNA deep sequencing on two somatic cell types (Sertoli cells purified from postnatal day 6 testes and interstitial cells of Cajal purified from gastrointestinal tract of adult mice) and somatic murine small intestine tissue using the 454 sequencing platform. We identified a total of 696 piRNAs from sncRNA libraries of Sertoli cells, ICCs, and small intestine. Our recent Ion Torrent PGM-based sncRNA-Seq analyses on the murine...
stomach resulted in 1.6 million reads aligned to germ cell piRNA clusters. By analyzing sequence reads from the two types of sncRNA-Seq analyses, we defined the pilRNA transcriptomes in two murine somatic cell types (i.e., Sertoli cell and ICC) and two murine organs (i.e., stomach and small intestine). Interestingly, pilRNAs are very similar to non-repeat-associated intergenic piRNAs expressed abundantly and predominantly in pachytene spermatocytes and round spermatids in adult testes. This notion is supported by the following findings: 1) the major size of pilRNAs is 30 nt, which is closer to that of pachytene piRNAs; 2) pilRNAs largely fall within known pachytene piRNA instead of prepachytene piRNA clusters; and 3) similar to pachytene piRNAs, pilRNAs are mainly derived from intergenic regions (6). In contrast to pachytene piRNAs, which only have ~2% ping-pong activity, pilRNAs have a definite ping-pong signature with ~30% of stomach pilRNAs having a 10-nt overlap at their 5’ ends (12). Therefore, we propose that pilRNAs are similar to germ cell pachytene piRNAs but are different in the fact that they have ping-pong activity like prepachytene piRNAs. It remains unknown how pachytene piRNAs are produced and what roles pachytene piRNAs play in spermatogenesis. Elucidation of piRNA biogenesis and functions would help us understand those of pilRNAs, and vice versa.

Although the function of pilRNAs remains elusive, their dynamic, tissue-specific expression patterns strongly suggest that pilRNAs are under tight regulation and thus may play an important role in cellular functions. Given that significantly more hits were observed when one to five mismatches were allowed, pilRNAs may be able to target mRNAs by partial annealing to their binding sites located in 3’-UTRs of mRNAs. Upon increasing the mismatch allowance from up to two to up to five, an increase in CDS-targeting pilRNAs was also observed.

FIGURE 7. pilRNA production is independent of the known miRNA/endo-siRNA or piRNA biogenetic pathways. A, heat map representing the pilRNA transcriptome in wild-type, Dicer- or Drosha-null murine Sertoli cells based on small RNA sequencing. Expression is based on a Log10 scale. B, scatter plots depicting log2(normalized value) expression between replicates from FACS-purified Sertoli cells of Amh-cre control, Dicer conditional knock-out, and Drosha conditional knock-out mice. R² values were at least 0.82, indicating relatively good correlation between replicates. Samples had an average depth of ~450,000 aligned reads. C, heat maps representing qPCR analyses of expression levels of three intestinal pilRNAs (pilRNA-in3, pilRNA-in18, and pilRNA-in87) and two known germ cell piRNAs (piR-118029 and piR-126541) in 10 organs of MIWI, MIWI2, and MOV10L1 global KO mice. U6 snRNA was used as an endogenous control, and values in KO samples were relative to respective values of wild-type samples.
observation may be the result of CDSs having longer sequences than UTRs, thereby increasing the random chance of a complementary match. Further bioinformatic analyses may reveal the rules that piRNAs follow to recognize their mRNA targets, and their effects on mRNA fate need to be confirmed experimentally in the future.

Although both DICER and DROSHA are ubiquitously expressed, DROSHA substrates must contain either the stem-loop structure or double-stranded RNA. Because neither piRNAs nor pachytene piRNAs contain such a structure, it is not surprising to see that inactivation of either Dicer or Drosha, despite changes in piRNA levels probably caused by secondary effects, does not block piRNA biogenesis. We know pachytene piRNAs require MIWI and several other proteins including MOV10L1, MITOPLD, and GASZ, which are exclusively expressed in male germ cells (25, 43–45). In addition, our lab has confirmed the lack of MIWI, MIW12, and MILI expression in several somatic tissues including murine stomach and confirmed the absence of these proteins by Western blot (data not shown). Therefore, it is expected that inactivation of MIWI/MIW12 or MOV10L1 would not abolish piRNA production. This was also anticipated, because although piRNAs are similar to pachytene piRNAs, they do have a key difference in that they have ping-pong mechanism activity. However, given the high degree of similarities between piRNAs and pachytene piRNAs, similar machinery is likely operating in somatic cells for piRNA production. Thus, it is critical to identify the proteins responsible for somatic piRNA production. The elucidation of piRNA biogenesis will further help define the physiological role of piRNAs in somatic cell lineages.

In summary, our data confirm that somatic cells indeed express abundant piRNAs, and these piRNAs are similar to pachytene piRNAs with the exception of ping-pong activity. piRNAs are synthesized using a yet to be defined pathway, and piRNAs may function in a manner similar to miRNAs as post-transcriptional regulators. With more understanding of piRNAs, the biogenesis and function of piRNAs, the somatic counterparts of germ cell piRNAs, will be elucidated in the near future.

FIGURE 8. Partial complementary alignment of 696 piRNAs identified in 454 sequencing to all known mRNAs. A, number of piRNAs potentially targeting 3′-UTRs, 5′-UTRs, and CDSs of all known murine mRNAs. Shown are the results of reverse complementary match of piRNAs to mRNAs when one to two or one to five mismatches were allowed. B, number of mRNA transcripts potentially targeted within the 3′-UTRs, 5′-UTRs, and CDSs by piRNAs. Sm Int, small intestine.

REFERENCES

1. Aravin, A., Gaidatzis, D., Pfeffer, S., Lagos-Quintana, M., Landgraf, P., Iovino, N., Morris, P., Brownstein, M. J., Kuramochi-Miyagawa, S., Nakano, T., Chien, M., Russo, J. J., Ju, J., Sheridan, R., Sander, C., Zavolan, M., and Tuschi, T. (2006) A novel class of small RNAs bind to MILI protein in mouse testes. Nature 442, 203–207

2. Girard, A., Sachidanandam, R., Hannon, G. J., and Carmell, M. A. (2006) A germ-line-specific class of small RNAs binds mammalian Piwi proteins. Nature 442, 199–202

3. Grivina, S. T., Beyret, E., Wang, Z., and Lin, H. (2006) Identification and characterization of two small RNA-generating loci in mouse embryonic stem cells. Genes Dev. 20, 1709–1714

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

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

6. Aravin, A. A., Sachidanandam, R., Bourchis, D., Schaefer, C., Pezic, D., Toth, K. F., Bestor, T., and Hannon, G. J. (2008) A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. Mol. Cell 31, 785–799

7. Deng, W., and Lin, H. (2002) miwi, a murine homolog of piwi, encodes a cytoplasmic protein essential for spermatogenesis. Dev. Cell 2, 819–830

8. Unhavaithaya, Y., Hao, Y., Beyret, E., Yin, H., Kuramochi-Miyagawa, S., Nakano, T., and Lin, H. (2009) 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. 284, 6507–6519

9. Aravin, A. A., Sachidanandam, R., Girard, A., Fejes-Toth, K., and Hannon, G. J. (2007) Developmentally regulated piRNA clusters implicate MILI in transposon control. Science 316, 744–747

10. Carmell, M. A., Girard, A., van de Kant, H. J., Bourchis, D., Bestor, T. H., de Rooij, D. G., and Hannon, G. J. (2007) MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germ line. Dev. Cell 12, 503–514

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

12. Beyret, E., Liu, N., and Lin, H. (2012) piRNA biogenesis during adult spermatogenesis in mice is independent of the ping-pong mechanism. Cell Res. 22, 1429–1439

13. Wei, Z., Liu, X., and Zhang, H. (2012) Identification and characterization of piRNA-like small RNAs in the gonad of sea urchin (Strongylocentrotus nudus). Mar. Biotechnol. (NY) 14, 459–467

VOLUME 289 • NUMBER 47
JOURNAL OF BIOLOGICAL CHEMISTRY 32833

ASBMB

NOVEMBER 21, 2014 • VOLUME 289 • NUMBER 47
Somatic piRNA-like RNA Characterization

14. Yan, Z., Hu, H. Y., Jang, X., Maiherofer, V., Neb, E., He, L., Hu, Y., Hu, H., Li, N., Chen, W., and Khaitovich, P. (2011) Widespread expression of piRNA-like molecules in somatic tissues. *Nucleic Acids Res.* 39, 6596–6607

15. Ro, S., Park, C., Song, R., Nguyen, D., Jin, J., Sanders, K. M., McCarron, J. R., and Yan, W. (2007) Cloning and expression profiling of testis-expressed piRNA-like RNAs. *RNA* 13, 1693–1702

16. Ortogero, N., Hennig, G. W., Langille, C., Ro, S., McCarrey, J. R., and Yan, W. (2013) Computer-assisted annotation of murine Sertoli cell small RNA transcriptome. *Birol Reprod.* 88, 3

17. Wu, Q., Song, R., Ortogero, N., Zheng, H., Evanoff, R., Small, C. L., Griswold, M. D., Namekawa, S. H., Royo, H., Turner, J. M., and Yan, W. (2012) The RNAse III enzyme DROSHA is essential for microRNA production and spermatogenesis. *J. Biol. Chem.* 287, 25173–25190

18. Ro, S., Ma, H. Y., Park, C., Ortogero, N., Song, R., Hennig, G. W., Zheng, H., Lin, Y. M., Moro, L., Hsieh, J. T., and Yan, W. (2013) The mitochondrial genome encodes abundant small noncoding RNAs. *Cell Res.* 23, 759–774

19. Bao, J., Zhang, Y., Schuster, A. S., Ortogero, N., Nilsson, E. E., Skinner, M. K., and Yan, W. (2014) Conditional inactivation of Miwi2 reveals that Miwi2 is only essential for prospermatogonial development in mice. *Cell Deathiffer.* 21, 783–796

20. Muzumdar, M. D., Tasic, B., Miyamichi, K., Li, L., and Luo, L. (2007) A global double-fluorescent Cre reporter mouse. *Genesis* 45, 593–605

21. Park, C., Yan, W., Ward, S. M., Hwang, S. J., Wu, Q., Hatton, W. J., Park, J. K., Sanders, K. M., and Ro, S. (2011) MicroRNAs dynamically remodel gastrointestinal smooth muscle cells. *PLoS One* 6, e18628

22. Ro, S., and Yan, W. (2010) Small RNA cloning. In RNA Therapeutics: *Function, Design, and Delivery* (Sioud, M., ed) pp. 271–283, Humana Press Inc., Totowa, NJ

23. Song, R., Hennig, G. W., Wu, Q., Jose, C., Zheng, H., and Yan, W. (2011) Male germ cells express abundant endogenous siRNAs. *Proc. Natl. Acad. Sci. U.S.A.* 108, 13159–13164

24. Edgar, R., Domrachev, M., and Lash, A. E. (2002) Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* 30, 207–210

25. Watanabe, T., Chuma, S., Yamamoto, Y., Kuramochi-Miyagawa, S., Toki, y., Toyoda, A., Hoki, Y., Fujiyama, A., Shibata, T., Sado, T., Noce, T., Nakano, T., Nakatsuji, N., Lin, H., and Sasaki, H. (2011) MITOPLD is a mitochondrial protein essential for nuage formation and piRNA biogenesis. *Nucleic Acids Res.* 40, D84–D90

26. Chan, P. P., and Lowe, T. M. (2009) GtRNAdb: a database of transfer RNA genes. *Bioinformatics* 25, D112–D115

27. Jurka, J., Kapitonov, V. V., Pavlicek, A., Klonowski, P., Kohany, O., and Jurka, J. (2002) Identification and characterization of novel ambisome microRNAs by Solexa sequencing. *Genome Biol.* 10, R78

28. Kozomara, A., and Griffiths-Jones, S. (2011) miRBase: integrating miRNA data. *Nucleic Acids Res.* 39, D152–D157

29. Liu, C., Bai, B., Skogerbo, G., Cai, L., Deng, W., Zhang, Y., Bu, D., Zhao, Y., and Chen, R. (2005) NONCODE: an integrated knowledge database of non-coding RNAs. *Nucleic Acids Res.* 33, D112–D115

30. Mituyama, T., Yamada, K., Hattori, E., Okida, H., Ono, Y., Terai, G., Yoshizawa, A., Komori, T., and Asai, K. (2009) The Functional RNA Database 3.0: databases to support mining and annotation of functional RNAs. *Nucleic Acids Res.* 37, D89–D92

31. Waterston, R. H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J. F., Agarwal, P., Agarwala, R., Ainscough, R., Alexandersson, M., An, P., Antonarakis, S. E., Attwood, J., Baertsch, R., Bailey, J., Barlow, K., Beck, S., Berry, E., Birren, B., and Bloom, T. (2002) Initial sequencing and comparative analysis of the mouse genome. *Nature* 420, 520–562

32. Chen, X., Li, Q., Wang, J., Guo, X., Jiang, X., Ren, Z., Weng, C., Sun, G., Wang, X., Liu, Y., Ma, L., Chen, J. Y., Wang, J., Zen, K., Zhang, J., and Zhang, C. Y. (2009) Identification and characterization of novel ambisome microRNAs by Solexa sequencing. *Genome Biol.* 10, R78

33. Li, X. Z., Roy, C. K., Dong, X., Bolcun-Filas, E., Wang, J., Han, B. W., Xu, J., Moore, M. J., Schimenti, J. C., Weng, Z., and Zamore, P. D. (2013) An ancient transcription factor initiates the burst of piRNA production during early meiosis in mouse testes. *Mol. Cell* 50, 67–81

34. Frost, R. J., Hamra, F. K., Richardson, J. A., Qi, X., Bassel-Duby, R., and Gruenbaum, R. (2009) GASZ is essential for male meiosis and suppression of retrotransposon expression in the male germline. *Nature* 453, 798–802

35. Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Probst, P., Rådmark, O., Kim, S., and Kim, V. N. (2003) The nuclear RNAIII Drosophila initiates microRNA processing. *Nature* 425, 415–419

36. Bartel, D. P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116, 281–297

37. Papaioanou, M. D., Pitielli, I. L., Ro, S., Park, C., Aubry, F., Schaad, O., Veijnar, C. E., Kühne, F., Descombes, P., Zdobnov, E. M., McManus, M. T., Guilhou, F., Harfe, B. D., Yan, W., Jégou, B., and Nef, S. (2009) Sertoli cell Dicer is essential for spermatogenesis in mice. *Dev. Biol.* 326, 250–259

38. Lee, R. C., Feinbaum, R. L., and Ambros, V. (1993) The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* 75, 843–854

39. Watanabe, T., Totoki, Y., Toyoda, A., Kaneda, M., Kuramochi-Miyagawa, S., Ohata, Y., Chiba, H., Kohara, Y., Kono, T., Nakano, T., Surani, M. A., Sakaki, Y., and Sakaki, H. (2008) Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature* 453, 539–543

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

41. Lewis, B. P., Burge, C. B., and Bartel, D. P. (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* 120, 15–20

42. Ma, L., Buchold, G. M., Greenbaum, M. P., Roy, A., Burns, K. H., Zhu, H., Han, D. Y., Harris, R. A., Coarfa, C., Gunaratne, P. H. Y., Yan, W., and Matz, M. M. (2009) GASZ is essential for male meiosis and suppression of retrotransposon expression in the male germline. *PLoS Genet.* 5, e1000635

43. Frost, R. J., Hamra, F. K., Richardson, J. A., Qi, X., Bassel-Duby, R., and Olson, E. N. (2010) MOV10L1 is necessary for protection of spermatoocytes against retrotransposons by Piwi-interacting RNAs. *Proc. Natl. Acad. Sci. U.S.A.* 107, 11847–11852

44. Zheng, K., Xiol, J., Reuter, M., Eckardt, S., Lee, N. A., McLoughlin, K. J., Stark, A., Sachidanandam, R., Pillai, R. S., and Wang, P. J. (2010) Mouse MOV10L1 associates with Piwi proteins and is an essential component of the Piwi-interacting RNA (piRNA) pathway. *Proc. Natl. Acad. Sci. U.S.A.* 107, 11841–11846