REPRODUCTION
RESEARCH

Bovine piRNA-like RNAs are associated with both transposable elements and mRNAs

Stewart Russell[1], Mehool Patel[1], Graham Gilchrist[1], Leanne Stalker[1], Daniel Gillis[1], David Rosenkranz[2] and Jonathan LaMarre[1]

[1]*Biomedical Sciences, University of Guelph, Guelph, Ontario, Canada and* [2]*Anthropologie, Johannes Gutenberg-Universität Mainz, Mainz, Germany*

Correspondence should be addressed to J LaMarre; Email: jlamarre@uoguelph.ca

Abstract

PIWI proteins and their associated piRNAs have been the focus of intensive research in the past decade; therefore, their participation in the maintenance of genomic integrity during spermatogenesis has been well established. Recent studies have suggested important roles for the PIWI/piRNA system outside of gametogenesis, based on the presence of piRNAs and PIWI proteins in several somatic tissues, cancers, and the early embryo. Here, we investigated the small RNA complement present in bovine gonads, gametes, and embryos through next-generation sequencing. A distinct piRNA population was present in the testis as expected. However, we also found a large population of slightly shorter, 24–27 nt piRNA-like RNA (pilRNAs) in pools of oocytes and zygotes. These oocyte and embryo pilRNAs exhibited many of the canonical characteristics of piRNAs including a 1U bias, the presence of a 'ping-pong' signature, genomic clustering, and transposable element targeting. Some of the major transposons targeted by oocyte and zygote pilRNA were from the LINE RTE and ERV1 classes. We also identified pools of pilRNA potentially derived from, or targeted at, specific mRNA sequences. We compared the frequency of these gene-associated pilRNAs to the fold change in the expression of respective mRNAs from two previously reported transcriptome datasets. We observed significant negative correlations between the number of pilRNAs targeting mRNAs, and their fold change in expression between the 4–8 cell and 8–16 cell stages. Together, these results represent one of the first characterizations of the PIWI/piRNA pathway in the translational bovine model, and in the novel context of embryogenesis.

Reproduction (2017) **153** 305–318

Introduction

PIWI-interacting RNAs (piRNAs) comprise a class of 24–32 nucleotide (nt) small non-coding RNAs that have recently become an intense area of focus in small RNA (sRNA) biology. Originally described in the context of animal gametogenesis (Deng *et al.* 2002, Girard *et al.* 2006, Vagin *et al.* 2006), additional roles for piRNAs have been recognized in tissue regeneration (Reddien 2005, Rizzo *et al.* 2014), tumor biology (Cheng *et al.* 2011, Chen *et al.* 2013) and, more recently, embryogenesis (Rouget *et al.* 2010, Roovers *et al.* 2015, Russell *et al.* 2016). PiRNAs bind a subclass of Argonaute proteins called PIWI proteins, which have been demonstrated to be indispensable for spermatogenesis in both mouse and *Drosophila*, with oogenesis also disrupted in the latter (Deng *et al.* 2002, Kuramochi-Miyagawa *et al.* 2004, Carmell *et al.* 2007). The mature PIWI protein–piRNA complex is referred to as the piRNA-induced silencing complex (piRISC) (Siomi *et al.* 2011). The PIWI pathway was first identified as a mechanism to constrain endogenous transposable elements (TEs) during genomic remodeling periods (Vagin *et al.* 2006, Carmell *et al.* 2007); however, recent research has implicated PIWI proteins in the regulation of mRNA transcripts as well (Rouget *et al.* 2010, Gou *et al.* 2014, Gebert *et al.* 2015, Zhang *et al.* 2015). PiRISCs bind target transcripts through complementarity with the piRNA, and effect silencing by either slicing target RNA or recruiting repressive chromatin marks to the source loci (Brower-Toland *et al.* 2007, Gunawardane *et al.* 2007).

To understand the emerging roles of the PIWI pathway in early embryos, it is important to highlight the parallels between gamete and embryo development. Genomic reprogramming occurs during two different stages of the reproductive life cycle: following fertilization of the oocyte and during the establishment of primordial germ cells in the developing fetus (Morgan *et al.* 2005). These periods of reprogramming are defined by the erasure of epigenetic imprinting and the re-establishment of cellular potency. In both cases, reprogramming exposes the genetic material by removing protective epigenetic marks, making

ISSN 1470–1626 (paper) 1741–7899 (online)
DOI: 10.1530/REP-16-0620
Online version via www.reproduction-online.org

it susceptible to TE activation, which must be tightly regulated to maintain genomic integrity (Tanaka *et al.* 2012, Fadloun *et al.* 2013). Following reprogramming, epigenetic marks are restored at specific loci, a process that is partially directed by the PIWI pathway during spermatogenesis (Aravin & Bourc'his 2008, Kuramochi-Miyagawa *et al.* 2008). In embryos, DNA methylation is restored during the maternal-to-zygotic transition, between the 4- and 16-cell stage in both human and bovine development (Memili & First 2000). Despite some differences in reprogramming, both gametogenesis and embryogenesis require temporal regulation of TEs and mRNA expression.

Although dispensable for murine embryogenesis, multiple roles for the PIWI pathway during embryogenesis in several other model species have recently emerged. In *Drosophila* embryos, piRNAs complementary to the 3′ untranslated region (3′UTR) of the nanos transcript are required for deadenylation and translational repression, highlighting the importance of the PIWI pathway for proper anterior–posterior patterning (Rouget *et al.* 2010). Similarly, Zili (PIWIL2) in zebra fish antagonizes transforming growth factor (TGF)-β signaling and is required for dorsal–ventral patterning (Sun *et al.* 2010). More recently, a study by Roovers *et al.* identified the presence of a population of piRNA-like RNAs (pilRNAs) in the oocytes of several species, including the bovine, and demonstrated the presence of an uncharacterized PIWI protein (PIWIL3), which may be an important participant in mammalian embryogenesis (Roovers *et al.* 2015). Mouse models suggest that there is a sRNA population present in the oocyte and early embryo, primarily composed of endo-siRNAs and piRNAs, which are associated with repeat elements (García-López *et al.* 2014). However, their function is likely dependent on the mouse-specific dicer isoform, dicer^{O}, hinting at a different mechanism in other mammals (Flemr *et al.* 2013). Several studies have investigated the roles of sRNAs during mammalian oocyte and embryo development in non-rodent models (Yang *et al.* 2012, Abd El Naby *et al.* 2013, Gilchrist *et al.* 2016); however, there are many questions still unanswered in this critical context.

These studies led us to postulate that piRNAs are present in bovine oocytes and zygotes, and that they have the potential to regulate the expression of genes and TEs during the maternal-to-zygotic transition. In the present study, we identified and characterized the pilRNAs present in bovine gametes, gonads, and zygotes. We report a unique pool of smaller pilRNAs to be present in oocytes and zygotes compared to testis. Striking differences between pilRNAs present in male and female gametes are evident with respect to the originating loci in the genome and clusters expressed. Finally, we have demonstrated for the first time the potential for mammalian embryonic pilRNAs to regulate endogenous gene expression.

Materials and methods

Tissue collection

All experimental samples were obtained from licensed local abattoirs. Testes and ovaries from sexually mature bovids were collected for RNA extraction within 30 min of slaughter. Ovaries were transported in phosphate buffered saline (PBS) at 35°C and oocytes were collected within 2 h of removal from the animal, and subsequently used for *in vitro* production techniques as described previously (Russell *et al.* 2016). Testis tissue was harvested on site and snap-frozen for future RNA extraction. Semen was obtained from reproductively fit bulls and stored in liquid nitrogen in cryoprotectant until swim-up to select competent sperm. All tissues for RNA analysis were snap-frozen in liquid nitrogen and stored at −80°C until use.

In vitro *embryo production*

Oocyte collection and *in vitro* fertilization procedures were conducted as previously described (Russell *et al.* 2016). Cumulus–oocyte complexes (COCs) were aspirated from follicles between 2 and 8 mm in diameter into HEPES-buffered Nutrient Mixture F-10 Ham (Sigma-Aldrich) media supplemented with 0.5% PenStrep (Invitrogen), 2% steer serum, and 2 units/mL heparin sodium salt (Sandoz). COCs with homogenous cytoplasm and at least two layers of cumulus were selected under a dissecting microscope and washed twice in HEPES-buffered TCM-199 maturation medium (Caisson Labs, Smithfield, UT, USA) containing 11 µg/mL sodium pyruvate (Sigma-Aldrich), 0.2 M L-glutamine (Sigma-Aldrich) and 0.5% PenStrep (Invitrogen). Selected COCs were incubated in groups of 15 under mineral oil in 80 µL drops of maturation medium containing 0.5 µg/mL FSH, 1 µg/mL LH, and 1 µg/mL estradiol (NIH) for 22 h at 38.5°C in a humidified 5% CO_2 atmosphere.

Matured oocytes were washed in HEPES-TALP (Caisson Labs) and fertilized in groups of 15–20 in 80 µL drops of TL-Fert (Caisson Labs) fertilization medium, supplemented with 20 µg/mL heparin sodium salt (Sigma-Aldrich) and 0.96 µg/mL albumin from bovine serum albumin (BSA) (Sigma-Aldrich). Cryopreserved bull semen (EastGen, Guelph, ON, Canada) was washed in modified HEPES-TALP (Caisson Labs) and competent sperm selected by swim-up method. Sperm concentration was qualitatively assessed and approximately 1×10^5 sperm were added to each drop. The sperm/COC cultures were incubated for 18 h at 38.5°C in a humidified 5% CO_2 atmosphere.

After fertilization, remaining cumulus cells were removed from the presumptive zygotes by vortexing in modified HEPES-TALP. These 1-cell embryos are considered presumptive zygotes because they were not selected based on second polar body extrusion, but rather by homogenous and non-granular cytoplasm, however runs with lower than 80% cleavage rates were not included. Zygotes were then cultured in 30 µL of *in vitro* culture (IVC) medium, consisting of Synthetic Oviduct Fluid (Caisson Labs) supplemented with 0.96 µg/mL BSA (Sigma-Aldrich), 44.3 µg/mL sodium pyruvate (Sigma-Aldrich), 2% non-essential amino acids (Sigma-Aldrich), 1% essential amino acids (Sigma-Aldrich), 0.5% gentamicin (Sigma-Aldrich), and 2% serum under mineral oil. Groups of 20

were incubated at 38°C in a humidified 5% CO_2 and 5% O_2 atmosphere until collection.

Oocyte and embryo collection

Immature oocytes were collected immediately after selection and mature oocytes were collected after 22 h of maturation. COCs were denuded by manual dissociation (pipetting), washed in PBS containing 0.01% poly vinyl alcohol, immediately flash frozen in liquid nitrogen, and stored at −80°C. Zygotes were washed to remove sperm and cumulus cells 18 h after insemination, followed by freezing and storage at −80°C. Only morphologically healthy oocytes/zygotes with a homogenous cytoplasm were collected. Groups of oocytes/zygotes were pooled to 600/700 per stage for deep sequencing, while groups of 40 were collected for qRT-PCR. For deep sequencing, oocytes/zygotes from 4 or more different days (runs) were pooled to increase starting amount for RNA extraction.

RNA extraction, cDNA synthesis, and quantitative PCR

Total RNA was extracted from tissues after homogenization using the miRNeasy Micro Kit (Qiagen) including on-column DNase digestion with the RNase-free DNase Set (Qiagen). *C. elegans* miR-39-3p was added to pools of oocytes and embryos used for quantitative PCR (qPCR) of piRNA targets as a spike-in control. Embryos were lysed by brief sonication and RNA was isolated with the miRNeasy Micro Kit (Qiagen). cDNA was generated with the miScript Plant RT Kit (Qiagen) for qPCR from total RNA from indicated oocyte/embryo pools. qPCR was conducted using miScript SYBR Green PCR Kit (Qiagen) for piRNA targets, the primers for which are listed in Table 1.

Small RNA deep sequencing and analysis of short reads

RNA from testis, ovary, and pools of 630–680 immature oocytes, mature oocytes, and zygotes was extracted as described above and submitted for sequencing at The Next-Generation Sequencing Facility in The Centre for Applied Genomics at the Hospital for Sick Children in Toronto. Generation and miRNA analysis of these datasets has been published previously (Gilchrist *et al.* 2016). SRNA populations were prepared for sequencing using a NEBNext Multiplex Small RNA Library Prep Set for Illumina (New England BioLabs, Whitby, ON, Canada) following the manufacturer's protocol. 3′ adaptor primers were ligated to input RNA for 1 h at 25°C, followed by 5′ adaptor denaturation and ligation to the RNA for 1 h at 25°C. RT reaction was performed for 1 h at 50°C and PCR amplification was performed at 94°C for 30 s followed by 12 cycles of 94°C for 15 s, 62°C for 30 s, and 70°C for 15 s, with a final extension 70°C for 5 min. Sequencing was performed on a Illumina HiSeq 2500 system with TrueSeq v3 chemistry (Illumina, San Diego, CA, USA) using the multiplex single read protocol (50 bases).

Table 1 piRNA primers used with miScript SYBR Green PCR Kit.

piR-1	TCTGACATCGGTTGCTCTGCTTC
piR-2	GTTCAGTGGTAGAATTCTCGCCT
piR-3	TGTAAGTTTCCGCGTTACTCTGTA
piR-4	TAGTCACACTGTGTCCTTTCAACT
piR-5	TCTCAGTCCTGAGAACTGTGTG
piR-6	TCCTTTAGGATAGACTGGTCTGATC

Preprocessing was performed in house using perl scripts available from http://www.smallrnagroup.uni-mainz.de/software.html (Jiang & Wong 2008, Rosenkranz & Zischler 2012, Roovers *et al.* 2015). First, 3′ adapters beginning with 'AGATCGGAAGAGC' were trimmed, followed by filtering out reads below 18 and above 35 nt in length. Reads with a greater than 10% chance of a miscalled base were removed, and the sequences were collapsed to unique reads. Non-coding RNAs were annotated using the Annotation Wizard (D Rosenkranz 2016, Bioinformatics, under review), with up-to-date sequences from the genomic tRNA (Chan & Lowe 2009), ensemble (Guttman *et al.* 2009), SILVA rRNA (Quast *et al.* 2013), and miRBase databases (Kozomara & Griffiths-Jones 2014), without mapping to known piRNA clusters. Unannotated sequences between 24 and 32 nt were considered pilRNAs and used for mapping with sRNAmapper to the unmasked bovine genome (bosTau7). During mapping, 0 mismatches were allowed in a seed substring of length 18, 2 mismatches in the 3 tailing bases and 1 between the seed and tail. Next, 5′–5′ overlaps were calculated to identify potential 'ping-pong' signatures from the map files. Sequence logos were made using Seq2Logo (Thomsen & Nielsen 2012) to observe positional nucleotide bias in the mapped pilRNA candidates. Mapping files were then used to reallocate pilRNAs based on unique genomic mapping and frequency, using options '5000 1000 b 0' (reallocate.pl) (Rosenkranz 2016). piRNA clusters were then predicted using proTRAC v2.1.2, and compared by identifying overlapping genomic intervals with the Galaxy 'Join' tool (usegalaxy.org). ProTRAC was run on default parameters with the exception of increasing the allowable pilRNA range to 32 nt (-pimax 32). MicroRNA (exclusively) expression analysis from oocytes at different maturation stages has been previously described elsewhere (Gilchrist *et al.* 2016). A flowchart outlining the steps for data processing is presented in Supplementary Fig. 1, see section on supplementary data given at the end of this article.

Genomic origin and targeting potential

Mapping of pilRNAs to TEs was performed using a custom perl script (RMvsMAP.pl) against RepeatMasker-annotated repeats for *Bos taurus* (bosTau7). Multiple mapping pilRNAs were divided evenly among mapping locations and therefore evenly among potentially targeted TEs. RepBase (www.girinst.org/repbase) was used to determine the TE class for each mapped TE, and the number of pilRNA hits per class was tallied. Genomic locations of mRNA were obtained from RefSeq (www.ncbi.nlm.nih.gov/refseq) and used to determine pilRNA overlap. When the mapping direction (e.g., +/-) for a given pilRNA–mRNA pair were opposite, the relationship was considered antisense, and when strands were the same, the pilRNA was considered sense to the putative mRNA target. Genomic origins of candidate piRNAs were determined by annotations from RefSeq.

Table 2 Summary of reads after each processing step.

	Immature oocytes	Mature oocytes	Zygotes	Testis	Ovary	Sperm
Raw reads	62,645,902	57,096,532	55,198,336	50,803,017	61,102,252	57,668,024
Step 1	42,082,708	37,763,231	31,935,130	46,200,935	49,272,673	27,606,789
	67.18%	66.14%	57.86%	90.94%	80.64%	47.87%
Step 2	37,976,078	34,272,449	28,830,099	36,314,705	6,957,506	18,986,269
	90.24%	90.76%	90.28%	78.60%	14.12%	68.77%
Step 3	26,478,497	23,618,949	18,911,966	31,771,460	2,555,336	5,900,637
	69.72%	68.92%	65.60%	87.49%	36.73%	31.08%
Step 4	23,366,727	20,716,168	16,257,836	28,326,504	1,311,347	3,059,703
	88.25%	87.71%	85.97%	89.16%	51.32%	51.85%
Step 5	19,813,284	17,402,708	13,403,715		1,239,959	1,115,080
	84.79%	84.01%	82.44%	0.00%	94.56%	36.44%
PilRNAs/step 1 (% pilRNAs of clean reads)	55.53%	54.86%	50.91%	61.31%	2.66%	11.08%

Percentages represent the fraction of reads remaining after each step. Raw read data has been reported previously (Gilchrist *et al.* 2016). Step 1, Adapter, 18–35 nt filter and quality processing; Step 2, Annotate and remove known small non-coding RNAs; Step 3, Length filter 24–32 nt; Step 4, Genome-mapped piRNA-like RNAs; Step 5, Clustered piRNAs (after weighting).

Lists of mapped piRNAs (alternatively mapped with Bowtie against bosTau7) that overlap with RefSeq genes, while considering piRNA and mRNA orientation, were generated as described above. Differentially regulated transcripts in bovine early embryogenesis identified in two recent next-generation sequencing (NGS) studies (Graf *et al.* 2014*a*, Jiang *et al.* 2014*a*) were used to determine the relationship between piRNA hits and mRNA expression. PiRNAs mapping to mRNA loci were tallied and compared to the lists of significantly differentially expressed genes for each stage comparison of embryogenesis.

Statistical analysis

All data were analyzed using the R statistical software package, version 3.12 ((R Development Core Team 2013) www.R-project.org). Specifically, the cor.test function was used to determine the Spearman's rank correlation coefficient and associated *P*-value for the entire dataset ($n = 11{,}949$) between the number of piRNA hits per mRNA, and the corresponding $\log_2$ fold change. Since the number of observations was large, the Spearman rank correlation coefficient was determined for a random sample (without replacement) of 500 records. A total of 10,000 such random samples were drawn from the full dataset, with the Spearman rank correlation coefficient and associated *P*-value recorded for each. The minimum, mean, and maximum observed correlations were identified, as well as the percentage of random samples that provided a significant (<0.05) *P*-value. This process was repeated using random samples based on the following variables: dataset of origin (i.e., Graf *et al.* 2014*a*, Jiang *et al.* 2014*a*), direction of mapping (sense or antisense), and stages being compared (e.g., 4-cell vs 8-cell). In the latter case, only stages with $n > 500$ were explored. Finally, a Poisson regression model was used to investigate the effect of $\log_2$ fold change and the other variables (i.e., dataset of origin, direction of mapping, and stages being compared) on the number of piRNA hits per mRNA.

Results

Deep sequencing of sRNA from bovine reproductive tissues

SRNA from bovine testis, ovary, sperm, immature oocytes, mature oocytes, and zygotes was sequenced using Illumina HiSeq 2500. Pools of 600–700 oocytes/embryos were used to generate a sRNA pool sufficient for deep sequencing. Notably, similar sized pools of embryos from the 2-, 4-, and 8-cell stages failed to produce libraries due to low input RNA. Each of our libraries contained ~50–60 million raw reads, however only 57–67% of raw reads remained in the oocyte/embryo pools after adapter and quality processing in contrast to 91% in the testis (Table 2). After trimming, quality processing, and length filtering, we determined the read length distribution for each of the datasets and observed a bimodal peak in the testis at 22- and 30-nucleotides (nt), a large peak in the ovary at 22 nt, and decreased number of reads with increased read length in the sperm with no defined single peak. The distributions between the oocyte and zygote datasets are similar, each displaying a prominent peak at 25–26 nt (Fig. 1). Next, we annotated the sRNAs present within our datasets using sequences from available databases (Fig. 2). The ovary dataset contained the largest proportion of known sRNA (85.9%) of which the majority were miRNA (88.6%), corresponding to the 22 nt peak (Fig. 1). The majority of reads from the other 5 datasets were unannotated and most of the annotated reads mapped to miRNAs and rRNAs. Of the unannotated reads between 24 and 32 nt, >86% were successfully mapped to the bovine genome from the testis, oocytes, and zygotes, and ~51% in the

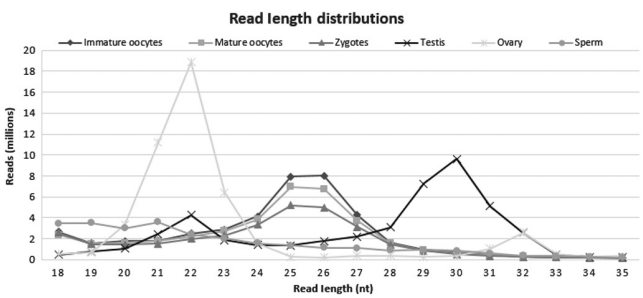

Figure 1 Read length distributions of six deep-sequenced samples. Graph shows the distribution of reads for each given read length after trimming, 18–35 nt length, and quality filtering steps.

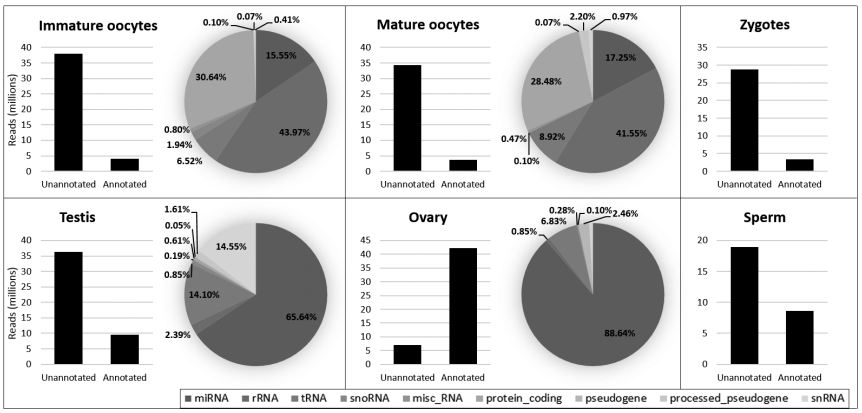

Figure 2 Small RNA annotation. Graphs on the left show the number of annotated and unannotated reads, and the charts on the right show the distribution of annotated small RNAs between classes. Unannotated reads contain the putative piRNAs.

ovary and sperm. At this point, genome-mapped sRNAs between 24 and 32 nt were considered pilRNAs and were assessed for features of canonical piRNAs.

Characterization pilRNAs

Of the total clean reads, 50.9–55.5% were classified as pilRNAs in the oocyte and zygote pools while 61.3% were classified as such in the testis. A common, but not ubiquitous, feature of piRNAs is the presence of a 'ping-pong signature' in the sequence reads. This is a 5′ to 5′ 10 nt overlap that results from secondary piRNA biogenesis via the 'ping-pong' amplification cycle (Supplementary Fig. 2). Identification of this signature within our pool suggests that a fraction of sRNA present was generated through this cycle. Four of our six filtered pilRNA pools demonstrate strong ping-pong signatures, with the ovary and sperm being the exceptions (Fig. 3). Another conserved feature of piRNAs across species is a uracil bias in the 1′ position (1U bias). We observe an 72–89% 1U bias in our pilRNA pools from oocytes and embryos, which is similar to the testis sample at 83.5%. Both the sperm and ovary pilRNAs show weak or absent 1U bias, with 30.6% and 16.1%, respectively. These biases are graphically represented as sequence logos in Figure 4. Unannotated ovary reads predominantly consisted of a single sequence without full-length hits to any known coding or non-coding RNAs, comprising approximately 50% of ovary pilRNAs. We also observed a slight but consistent adenosine bias in the 10th position in all the oocyte/zygote samples (31–36%), a feature of secondary piRNAs produced by the ping-pong cycle.

Cluster identification

Primary piRNAs have been shown to originate from specific clusters in the genome; they are transcribed into long precursors from piRNA dense loci within these clusters. We identified a total of 403 different putative piRNA clusters within the six datasets using a weighted proTRAC cluster analysis (Rosenkranz & Zischler 2012, Rosenkranz *et al.* 2015). In total, we detected 1104 different clusters from the 6 datasets. We observed 184 common clusters which accounted for >98% of the clustered pilRNAs in the immature oocytes, mature oocytes, and zygotes. Next, we analyzed the expression of pilRNAs from clusters across chromosomes (Fig. 4). In oocytes and zygotes there are only minor differences in expression, and the vast majority of clustered pilRNAs arise from chromosomes 3, 6, 13, 14, 17, and 23. This is in contrast to the testis in which cluster expression was more evenly distributed across chromosomes (Fig. 4). The majority of these clusters were unistranded; however, we also identified between 12 and 18% bidirectional clusters in the oocytes and zygotes. An example of each is shown in Supplementary Fig. 2. Interestingly, only 10% of the testis clusters were bidirectional. Within these clusters, the 1U nt bias was

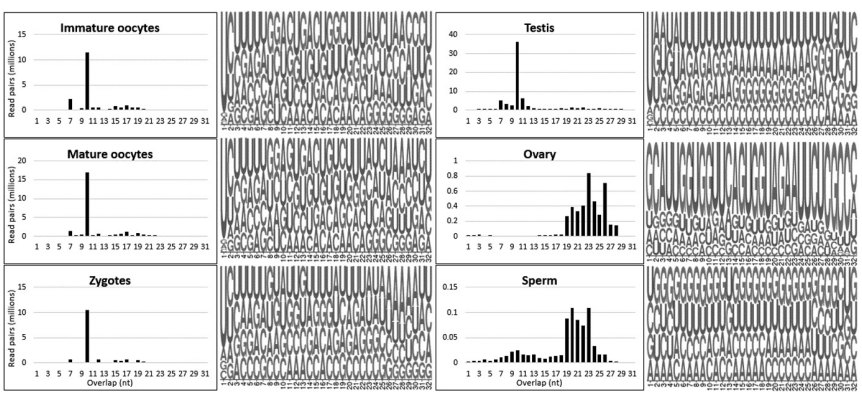

Figure 3 Identification of piRNA-like RNAs (pilRNAs) by ping-pong signature and 1U bias. Graphs on the left represent the number of pairs of reads with a 5′-to-5′ overlap in the datasets. A peak at 10 nucleotides is characteristic of piRNAs produced by the ping-pong cycle (see Supplementary Fig. 2). The sequence logos on the right represent nucleotide biases at each position of the pilRNAs, with a 1′ uracil bias characteristic of piRNAs bound to a PIWI protein.

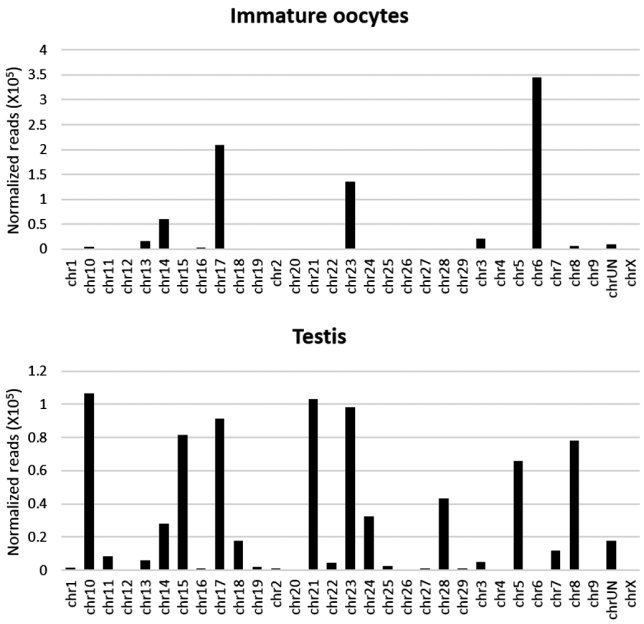

Figure 4 PiIRNA cluster expression across chromosomes. Values represent the number of normalized, clustered piIRNAs mapping to each chromosome in immature oocytes (top) and testis (bottom) datasets. Immature oocyte cluster distribution is representative of the mature oocyte and zygote distributions.

above 75%, or the 10A nt bias was above 50%. In order to compare the cluster expression between samples, the clusters were joined on their genomic intervals. As mentioned, the oocytes and zygotes expressed 184 common clusters between the three samples, with the adjacent developmental stages having the most clusters in common (Fig. 5). Interestingly, the zygotes only shares 10 clusters with the sperm, none of which are unique from those shared between the mature oocytes and sperm. Finally, we observed more clusters in common between the immature oocytes and testis than with its parent tissue, the ovary.

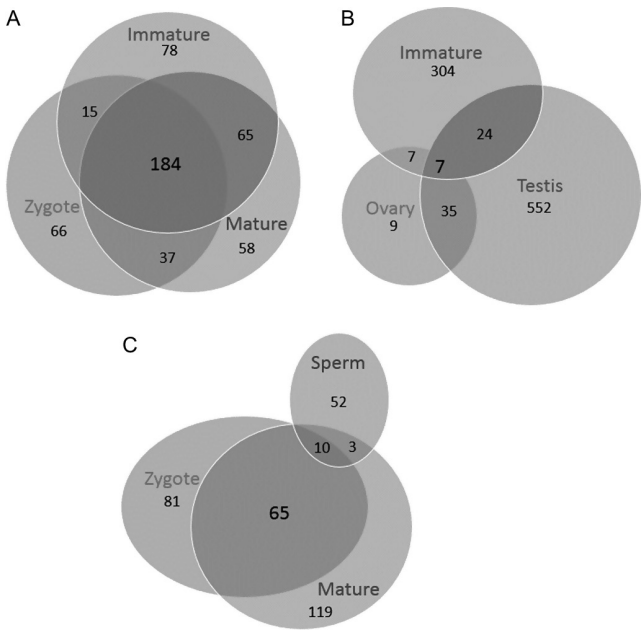

Figure 5 Cluster overlap between samples. Clusters were compared on genomic intervals, comparing the expressed clusters between (A) immature and mature oocytes, and zygotes; (B) immature oocytes, ovary, and testis; (C) mature oocytes, zygotes, and sperm.

Validation of piIRNA Expression by qPCR

To confirm the expression of piIRNAs identified by deep sequencing in the samples and validate the sequencing data, we conducted RNA adapter-based qRT-PCR on sRNA from independent pools of 40 oocytes and embryos at the equivalent stages of maturation to those used for sequencing. Six piRNA targets and one spike-in, *C. elegans* miR-39-3p added during RNA extraction (Fig. 6) were evaluated. In order to determine suitable reference targets, we employed the geneNorm algorithm (Vandesompele *et al.* 2002; https://genorm.cmgg.be) on all seven targets and determined piR-4 and piR-18 to be the most stable between stages. The expression of each piIRNA target was reproducibly detectable in the samples. Surprisingly, changes in the level of expression of only 2 of 6 targets (bta-piR-4 and bta-piR-9) correlated with the changes predicted by our deep sequencing data. Potential explanations for this observation are explored in the discussion.

Potential transposon and mRNA targeting by piIRNAs

Candidate piIRNAs were mapped to bovine TEs annotated by RepeatMasker to determine their potential

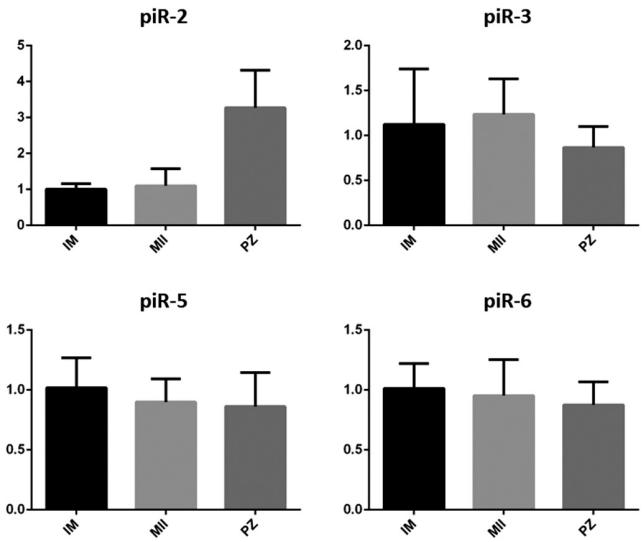

Figure 6 Adapter-based qRT-PCR of four piRNA targets in zygotes, immature, and mature oocytes. Fold change values determined by normalization to the geometric norm of piR-4, piR-18, and *C. elegans* miR-39-3p, represented by mean and standard error ($n = 3$).

target transcripts. The long interspersed nuclear element (LINE) RTE and SINE2/tRNA-derived repeats are the most abundantly represented TE classes in the bovine testis, oocyte, and zygote pilRNAs, representing over 50% of all TE mapped reads (Fig. 7). Within the LINE RTE class, pilRNAs map predominantly to BovB repeats in the antisense direction (Fig. 7). Within the SINE2/tRNA class, the most highly represented TEs are the Core-RTE and Bov-tA elements, both of which are short, non-autonomous repeats. Interestingly, the DNA TE Tigger was substantially represented, with >99% of the pilRNAs associated with the antisense strand. These reads were aligned to BovB to observe the putative targeting profile (Supplementary Fig. 4). The profiles of mapping location and frequency were very similar across development from the immature oocyte to zygote. Overall, pilRNAs consistently mapped to substantially different loci when compared to those from testis.

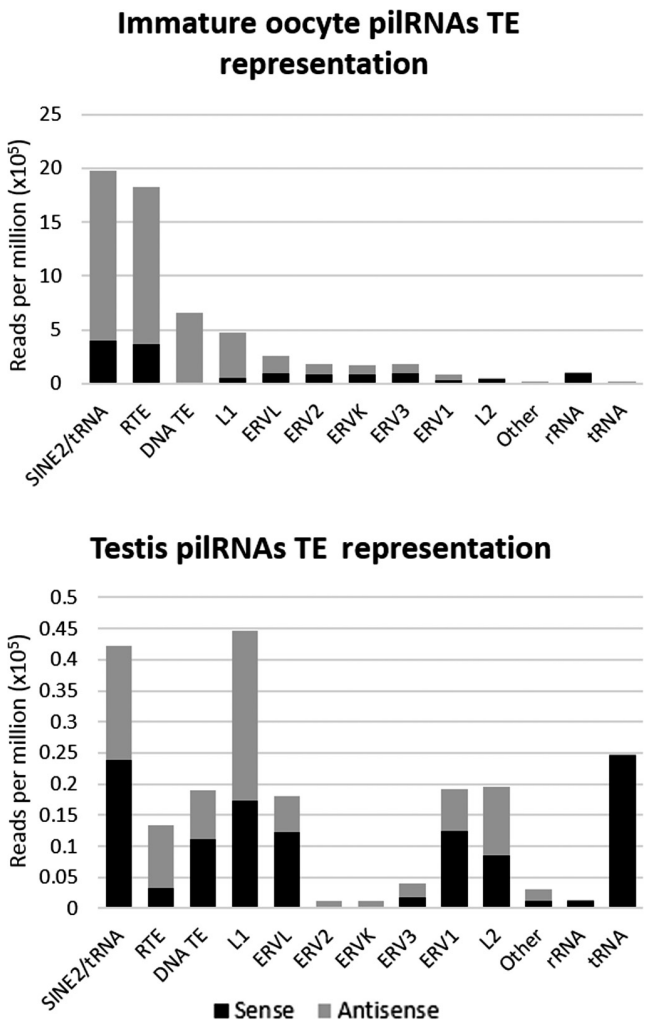

Figure 7 TE representation within immature oocyte and testis pilRNA datasets. PilRNAs were mapped against genomic transposable elements (TEs). Results from the testis and immature oocytes (representative of mature oocytes and zygotes) show the classes of TEs represented in the pilRNA populations in both the sense (black) and antisense (gray) orientations.

Recent studies have suggested that piRNAs have the potential to target mRNAs in addition to their traditional RT targets (Rouget *et al.* 2010, Gou *et al.* 2014, Zhang *et al.* 2015). We therefore investigated the relationship between pilRNAs mapped to the genome and mRNAs expressed in these tissue contexts, taking into account both the sense and antisense orientations of the piRNAs. Those pilRNAs mapping in the sense direction were considered as potentially originating from mRNAs. We also compared the number of pilRNA overlapping with mRNAs derived from the genome, to the gene expression data presented in two previous studies (Graf *et al.* 2014*b*, Jiang *et al.* 2014*a*). We observed a statistically significant negative correlation (rho = −0.1783165, P-value < 2.2×10^{-16}) between the number of piRNA hits for any specific mRNA target and the $\log_2$ fold change of the differentially expressed genes (Fig. 8). This test was also conducted within the variables of mapping direction, origin of transcriptome data, and stage comparison; the statistics of which are presented in Table 3. As a control, the fold change values were randomized per gene while maintaining the pilRNA hit values, which resulted in none of the above correlations, suggesting that the observed relationship between numbers of piRNAs mapped to a gene and its $\log_2$ fold change was not a result of the datasets having a net negative $\log_2$ fold change. Finally, pilRNA candidates that map to mRNAs were analyzed for a ping-pong signature, and show a significant (Z-score = 23.82) 10 nt overlap suggestive of participation in the ping-pong cycle (Fig. 9).

Discussion

PIWI proteins and piRNAs represent key players in a complex regulatory system that has only recently been recognized to play important roles outside of gametogenesis (Rouget *et al.* 2010, Jiang *et al.* 2014*b*, Roovers *et al.* 2015). This study represents one of the first characterizations of piRNA biology in bovine gonads, gametes, and zygotes. The results presented here support potential roles for piRNAs in the classical control of transposon expression during spermatogenesis in the bovine and expand the scope of recent studies in murine and *Drosophila* models that implicate piRNAs in mRNA turnover during gametogenesis (Rouget *et al.* 2010, Gou *et al.* 2014). In addition, the pattern of piRNA expression in bovine oocytes and zygotes is highly consistent with very recent studies by others (Roovers *et al.* 2015) and supports the likelihood that this regulatory pathway also participates in transposon control in oocytes and early embryos. Most importantly, we have identified piRNAs in oocytes that appear associated with target genes that are subsequently destined for turnover in the early bovine embryo.

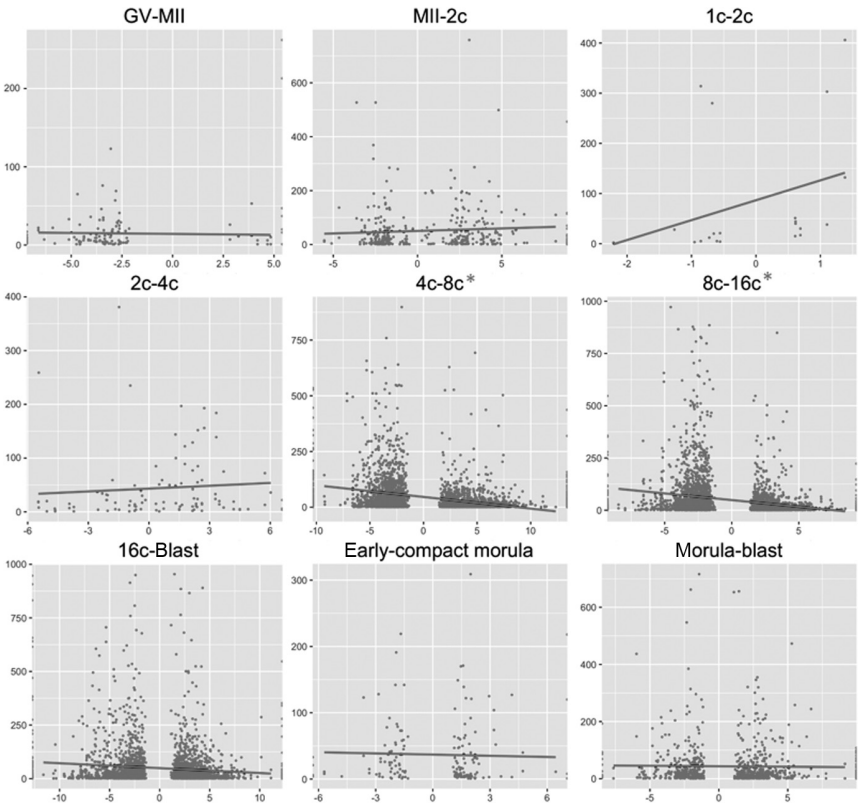

Figure 8 Scatter plots depicting piRNA hits vs $\log_2$ fold changes in the expression of the corresponding mRNA. Data from two studies of mRNA profiling during embryogenesis were used in a comparison of number of piRNA hits (*y*-axis) to mRNA $\log_2$ fold change (*x*-axis) between the stages indicated, allowing visualization of potential correlations. Red bars represent linear regressions and red stars indicate stages with 100% significant tests for correlation. The absence of data points between ~1 and −1 are the result of comparing pilRNAs to genes with a significantly different expression between stages (GV, germinal vesicle oocyte; MII, metaphase II oocyte; 2c, 2-cell embryo; 4c, 4-cell embryo; 8c, 8-cell embryo; mor, morula; blast, blastocyst).

One of the primary challenges in this study was isolating sufficient sRNA from oocytes and zygotes to successfully prepare a library for NGS. To this end 600–700 oocytes or zygotes for each of the stages, representing a minimum of 4 separate runs of *in vitro* production within each stage, were collected and pooled. Although this resulted in a single sequencing replicate per stage, the depth of our sequencing and the large number of individuals within each group increases our confidence that the counted reads accurately reflect the population of sRNAs from each of the stages. This is further supported by the high similarity in mapping patterns between the pilRNAs in our data and the recent independent work of others (Roovers *et al.* 2015).

Difficulties in producing high-quality libraries from the limited RNA present in the gamete samples resulted in decreased yields after processing compared to the tissue samples (testis, ovary). The read length distributions show clear peaks corresponding to miRNAs in the testis and ovary, and the testis shows an additional prominent peak at the expected piRNA size of 28–32 nt (Brennecke *et al.* 2007, Russell *et al.* 2016). Interestingly, the peak of ~24–27 nt present in the immature and mature oocytes and zygote reads that appears to represent a population of shorter piRNAs. After adapter trimming and quality filtering, known sRNAs were annotated using several databases. The remaining unannotated reads were considered pilRNA if they mapped to the genome and had a length between 24 and 32 nt. Unsurprisingly, this accounted for 61% of testis reads, a tissue known to express PIWI pathway components (Roovers *et al.* 2015, Russell *et al.* 2016). However, large pilRNA populations were also identified in the oocytes/zygotes (>50%), which led us to investigate their properties in that context.

To characterize these pilRNA, we began by investigating common piRNA features, one of which is the presence of a 'ping-pong' signature (Brennecke *et al.* 2007, Gunawardane *et al.* 2007). This signature arises from the slicer activity of the PIWI–piRNA complex directed at the 10th nucleotide of its target transcript, producing secondary piRNAs with

Table 3 Results of repeated measures of piRNA hit vs mRNA $\log_2$ fold change for each variable.

Variable	**Total** (all variables)	**Graf *et al.* 2014**	**Jiang *et al.* 2014**	**Antisense mapping**	**Sense mapping**	**4-cell vs 8-cell**	**8-cell vs 16-cell**	**16-cell vs blastocyst**	**Morula vs blastocyst**
Number of observations	10,446	6935	3511	4224	6222	3091	3117	2734	788
Minimum correlation	−0.3256	−0.3301	−0.3564	−0.2929	−0.3708	−0.4187	−0.3754	−0.2655	−0.1113
Mean correlation	−0.1796	−0.1682	−0.1980	−0.1366	−0.2266	−0.2741	−0.2036	−0.1290	−0.0902
Maximum correlation	−0.0878	−0.0878	−0.0878	−0.0877	−0.0927	−0.1222	−0.0884	−0.0877	0.0977
% significant of 10,000 tests	97.88	96.09	99.59	78.34	99.94	100.00	99.78	68.36	0.48

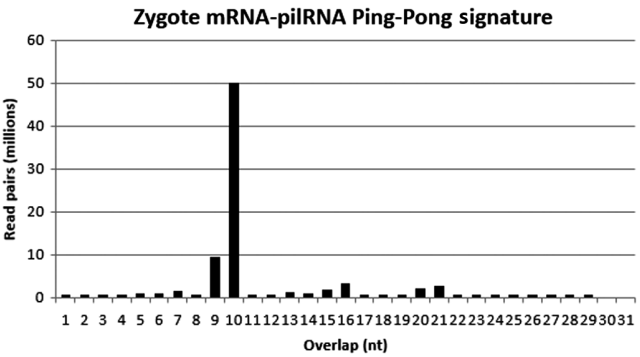

Figure 9 Ping-pong signature of mRNA mapped pilRNAs. Candidate pilRNAs from the zygote were mapped to bovine mRNAs, and the resulting mapped sequences were checked for a ping-pong signature. A significant peak at 10 nucleotides suggests that mRNA mapped pilRNAs from the zygote may participate in the slicer-mediated amplification loop (*Z*-score = 23.82).

a 5′–5′ overlap of 10 nt. Through this amplification loop, the piRNA pathway generates new PIWI–piRNA complexes from target sequences (Supplementary Fig. 2). We detected a ping-pong signature in our testis pilRNAs, similar to that seen in mature testis tissue from mice (Brennecke *et al.* 2007). Ovarian and sperm-born pilRNAs did not display a ping-pong signature in our samples, consistent with a recent study suggesting that piRNA enrichment by periodate elimination may be necessary to detect their presence (and the accompanying signature) in tissues such as the ovary where the overall abundance is low (Roovers *et al.* 2015). Intriguingly, we did observe a very strong ping-pong signature in our oocyte/zygote samples. This finding suggests that there is likely to be an uncharacterized PIWI–piRNA pathway functioning in the early embryo context.

A second canonical feature of piRNAs is the presence of a uridine bias in the first position of the 5′ end, known as the 1U bias, generated as a result of preferential binding of piRNA intermediates displaying a uridine at the 5′ terminus (Kawaoka *et al.* 2011, Cora *et al.* 2014). This bias was present in all samples showing a ping-pong signature. Experiments investigating the source of the 1U bias suggest that intrinsic properties of the PIWI protein partner are responsible for binding and stabilization of 1U containing piRNAs (Cora *et al.* 2014). Additionally, HSP90 has been shown to confer 1U bias to piRNAs loaded onto PIWI proteins *in vitro* (Izumi *et al.* 2013). This suggests that oocyte and zygote pilRNAs may also be loaded by the chaperone HSP90, similar to the canonical biogenesis in the testis. However, this does not seem to be the case for sperm and ovary pilRNA. This is one of the few studies to investigate oocyte and sperm-borne pilRNAs (García-López *et al.* 2014, Pantano *et al.* 2015), and we suggest that the pilRNA content is undetectable among the noise of other 24–32 nt reads.

Mapping of piRNA to the genome allows the identification of piRNA-rich loci called piRNA clusters (Aravin *et al.* 2006, Girard *et al.* 2006, Brennecke *et al.* 2007). These clusters produce single-stranded precursors that contain multiple piRNAs, which are exported from the nucleus, processed into intermediate fragments, loaded onto PIWI proteins, and trimmed into mature piRNAs. Several accepted methods of cluster identification from piRNA containing datasets have been developed, and we utilized an established cluster detection software called proTRAC (Rosenkranz & Zischler 2012). This approach resulted in the identification of 1,104 clusters in the bovine genome, of which the oocytes and zygotes express about 30%, testis 56%, and ovary and sperm 6%. The expression of pilRNA from these clusters is similar between immature oocyte, mature oocyte, and zygote samples in expressed clusters, and these common clusters produce the majority of the piRNAs (~98%). The relative homogeneity in cluster expression between the oocytes and zygotes suggests that pilRNA precursor regulation may be significantly restricted during oocyte maturation, and that the oocyte pilRNA repertoire is generated during oogenesis prior to the period when germinal vesicle oocytes are normally collected. In contrast, the cluster profile of pilRNAs from the testis is distinct from that of the oocyte/zygote, with only 31 common clusters, and a more evenly distributed pilRNA frequency between clusters.

Previous studies suggest that piRNAs present in the sperm may be delivered to the oocyte upon fertilization, albeit in small quantities (García-López *et al.* 2014, Pantano *et al.* 2015). Comparing cluster expression between the sperm and mature oocytes shows that only 13 pilRNA clusters are expressed in both tissues. One might expect this proportion to increase in the zygote if sperm-'delivered' pilRNAs form uniquely expressed clusters upon fertilization. However, the sperm–zygote overlap remains restricted to 10 clusters upon fertilization, actually less than the sperm–mature oocyte overlap, suggesting minimal contribution in terms of cluster diversity from sperm-borne pilRNAs. We cannot, however, rule out the possibility that the 'dose' of individual or multiple pilRNAs from clusters might be augmented through sperm-mediated delivery. Alternatively, it is possible that few sperm-born pilRNAs 'kick-start' cluster expression after the zygote stage. Most notably, the greatest similarity in cluster expression is found between the immature oocyte, mature oocyte, and fertilized zygote, where pilRNAs from 184 clusters are common across stages. These common clusters account for ~98% of the normalized, clustered pilRNAs in each of the stages. This suggests that pilRNA expression is relatively static during oocyte maturation and immediately after fertilization. We suggest that this relatively stable pilRNA population likely participates in RNA turnover and transposon control during subsequent early stages of bovine embryogenesis.

In an attempt to validate the piRNA expression pattern observed in the NGS results, we adapted a plant miRNA Sybr-based qPCR assay for detection of piRNA from bovine oocytes and embryos. Plant miRNAs carry a 2′-*O*-methylation at their 3′ ends, equivalent to that seen in mammalian piRNAs. Of 6 piRNA targets, only two followed the trend predicted by the sequencing data when measured by qPCR. Despite identical conditions between *in vitro* embryo production runs, modest variability in oocyte and embryo competence has been observed, which could contribute to a lack of trending between NGS and qPCR. This variability can be attributed to the quality of the ovaries obtained from the abattoir, which depends on the health of the animals on any given day. An additional explanation for this lack of trending could be variations in the 3′ ends of piRNA, which makes primer design difficult and decreases the reliability of amplification of specific piRNA targets. Finally, piRNAs canonically possess a 2′-*O*-methyl (2′-*O*-Me) group on their 3′ end, protecting them from degradation (Kirino & Mourelatos 2007, Saito *et al.* 2007). However, a recent study suggested that piRNA from oocytes and embryos may be un-methylated, due to depletion of oocyte piRNAs after periodate oxidation treatment (Roovers *et al.* 2015). To remain robust in our detection of piRNAs, our technique involves RNA adaptor ligation, designed to function in the presence of 2′-*O*-Me. If these piRNAs are indeed un-methylated, the process of adaptor ligation may have introduced a bias into downstream qPCR measurements of target piRNAs.

PIWI proteins and piRNAs are potent regulators of both transcriptional and posttranscriptional transposon and mRNA expression (Brennecke *et al.* 2007, Gunawardane *et al.* 2007, Shoji *et al.* 2009, Rouget *et al.* 2010, Gou *et al.* 2014). Previous studies have demonstrated that an increase in TE expression is a defining, and potentially essential, event during the early stages of embryogenesis (Peaston *et al.* 2004, Beraldi *et al.* 2006, Bui *et al.* 2009). However, the deleterious effects of widespread TE expression are well-documented, suggesting that expression during the reprogramming phase of embryogenesis must be tightly controlled. By mapping piRNAs to known bovine and ancestral TEs, we identified the various classes and their representation in each of our datasets. Unsurprisingly, the most commonly targeted TE in our piRNA-rich datasets is the RTE LINE element BovB. The abundance of piRNAs with the potential to target the LINE1 RTE BovB is likely due to the presence of 376 thousand copies comprising a total of 10.7% of the bovine genome (Adelson *et al.* 2009), second only to LINE1 repeats with 11.3% coverage. Previous research by Bui *et al.* suggests the primary retrotransposon class expressed in pre-implantation bovine embryos is ERV1 (Bui *et al.* 2009). However, ERV1s account for only 1% of our mapped piRNAs compared to ERV2s, which account for 4%. BTLTR1 (an ERV2 repeat) has previously been shown to have the fewest expressed sequence tags present in morula embryos (Bui *et al.* 2009), but is potentially targeted by as much as 2% of our mapped piRNAs. Taken together, these data suggest an inverse relationship between retrotransposon expression and the piRNAs that map to them in the pre-implantation embryo. It will be important to ultimately compare a more comprehensive characterization of embryonic TE expression to the pattern of expressed piRNA that target them in order to understand the roles of this relationship on embryonic development.

Some striking differences in TE representation were evident when considering TE class representation comparisons between samples. The coverage and frequency of TE mapping of piRNAs between oocyte and zygote samples are very similar, but very different from the testis. It is well-documented that the oocyte is transcriptionally inactive shortly after reaching the antral follicle stage, and few changes in gene expression occur before the 4–8 cell stage (Sirard & Blondin 1996, Hyttel *et al.* 1997, Graf *et al.* 2014*a*). The majority of piRNAs from oocyte/zygote piRNAs align to the LINE RTEs, SINE2/tRNAs, and ERVK/L repeat classes, whereas the testis shows a greater mapping to the LINE family (L1, L2) and endogenous retroviruses (ERV1, ERVL). Additionally, there is a strong bias for antisense mapping of oocyte and zygote piRNAs (>80%) compared to the testis (~50%). The lack of TE sense-derived piRNAs in the oocyte/zygote context could represent a population that is prepared for TE defense rather than participating in active TE destruction. These data suggest that the oocytes and zygotes contain a relatively static TE targeting piRNA repertoire that is distinct from that found in the testis. This phenomenon is likely the result of the diverse pattern of TE repression required across different tissues due to the differing reprogramming and pattern of TE expression.

Using the location and frequency of piRNAs mapping along the length of the primary LINE RTE target, BovB, a pattern we have termed a ‘targeting profile’ emerges. Between the germinal vesicle oocyte, mature oocyte, and zygote, both the position and magnitude of piRNA mapping to BovB remains highly consistent, a trend seen across multiple targeted TEs. However, comparing the profile between these relatively static piRNA populations and those of the testis, ovary, and sperm, we observe large variations in both the position and frequency of potential piRNA targets. Additionally, several sites exhibiting a ping-pong signature of sense and antisense reads are present, suggesting that BovB is a target in the piRNA amplification loop. The differences observed in targeting profiles of TEs such as BovB between different tissue samples likely reflects a unique biology responsible for the balance between TE expression and silencing in different cellular contexts.

In addition to their roles in TE silencing, piRNAs have very recently been shown to be essential for the

widespread, targeted elimination of mRNA transcripts during pachytene spermatogenesis, a process highly similar to the maternal-to-embryonic transition (MET) in the mammalian embryo (Gou *et al.* 2014). Based on the results of the present study, we postulate that pilRNAs present in testis, oocytes, and early embryos have the potential to target both TE and mRNA transcripts. Since the discovery of *nanos* transcript de-adenylation directed by the PIWI pathway in *Drosophila* (Rouget *et al.* 2010), and later the illuminating work done by Rajasethupathy *et al.* showing transcriptional regulation of CREB2 by piRNAs in *Aplysia* neurons (Rajasethupathy *et al.* 2012), evidence has accumulated concerning the potential of PIWI–piRNA complexes to regulate expression of mRNAs in addition to TEs. Most notably, pull-down experiments in *Drosophila* and mouse led to the identification of a host of mRNA genes regulated by the PIWI pathway (Gou *et al.* 2014, Barckmann *et al.* 2015). We therefore hypothesized that pilRNAs in the zygote may participate in the elimination of maternal transcripts during embryogenesis resulting in a negative correlation between pilRNAs and target gene expression. To this end, differentially regulated transcripts present during bovine embryogenesis were obtained from two published datasets (Graf *et al.* 2014*a*, Jiang *et al.* 2014*a*), and we compared these to the list of pilRNA mapping to coding genes in the present study. Using the compiled pilRNA hits per gene and comparing it to the $\log_2$ fold change for the differentially expressed genes, we observed a statistically significant −0.18 overall correlation between the numbers of pilRNAs mapped to any given gene and the relative fold change of that transcript. The strongest relationships with the highest confidence intervals were present in the transcript changes between the 4-cell vs 8-cell and 8-cell vs 16-cell stages, with statistically significant mean correlations of −0.27 and −0.20, respectively. These correlations were internally validated through repeated random sampling and significance testing of the pilRNA hits vs gene expression changes (10,000 tests of 500 randomly selected genes). These repeated measures were only conducted for the variables with a sufficient number of observations. Strikingly, almost 100% of the tests from both the 4-cell vs 8-cell and 8-cell vs 16-cell stages were significant, in contrast to the later stage comparison of 16-cell vs blastocyst, of which 68.4% of the tests were significant with a weaker mean correlation of −0.13. Most importantly, for the changes in gene expression between the morula and blastocyst stages, there were very few significant tests and almost no correlation observed. Gene expression changes were also scrambled across the dataset to confirm that the negative correlation was not a product of many data points and/or an overall net negative gene expression change. The scrambled control abolished the correlations both globally and within the random repeated sampling.

Given the large number of observations, the variability in transcripts, and the confidence of the repeated sampling, the observed correlation between increased numbers of targeting piRNAs and decreasing gene expression between the 4-cell and 16-cell stages is unlikely to represent a product of random piRNA overlap with coding genes. We therefore suggest that the pilRNA repertoire present in zygotes contain the sequences that are likely to participate in mRNA elimination or transcriptional gene silencing during the MET. To further strengthen our hypothesis, we looked specifically at pilRNAs mapping to mRNAs and checked whether they display a ping-pong signature. Indeed, a significant peak at 10 nts suggests that the mRNA-associated pilRNAs from the zygote also participated in the ping-pong cycle, and therefore have the capacity to silence mRNAs through a slicer-mediated mechanism. Posttranscriptional regulation of mRNAs by miRNAs has recently emerged as essential for embryogenesis in *Drosophila*, mouse, zebra fish, and cows (Bernstein *et al.* 2003, Giraldez *et al.* 2006, Tang *et al.* 2007, Bushati *et al.* 2008, Tripurani *et al.* 2011); however, a similar role for piRNA-guided mRNA repression has not been investigated in mammalian embryos. We attribute the relatively low correlation between the magnitude of change and the number of targeting pilRNAs to the overall complexity of mRNA turnover in the early embryo, where multiple pathways (miRNAs, endo-siRNAs, RNA-binding proteins) control mRNA decay (Mondou *et al.* 2012). An additional explanation for our significant, but relatively low correlation values may be that our mapping criteria are too stringent to capture the breadth of potential piRNA–mRNA interactions, and therefore underestimate the number or specificity of target transcripts. Indeed, mRNA elimination during late stage spermatogenesis in the mouse is mediated by much lower stringency of piRNA–target complementarity than previously shown to be required for TE repression (Reuter *et al.* 2011, Gou *et al.* 2014, Zhang *et al.* 2015).

The stability of this population in read length distribution, transposon targeting, cluster expression, and quantity as determined by qPCR leads us to propose that these pilRNA are likely deposited during oogenesis and remain in the zygote as a genetic defense mechanism with the potential to regulate TEs that arise during embryonic reprogramming. Moreover, significant negative correlations with putative gene targets suggest that these pilRNAs may also participate in gene regulation during the MET between the 4-cell and 16-cell stages, similar to the involvement of the PIWI–piRNA pathway in mRNA elimination during late spermatogenesis (Gou *et al.* 2014).

Supplementary data

This is linked to the online version of the paper at http://dx.doi.org/10.1530/REP-16-0620.

Declaration of interest

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Funding

This work was supported by grants from the Natural Sciences and Engineering Council of Canada (to J L) and the Ontario Ministry of Agriculture, Food and Rural Affairs (to J L).

Acknowledgements

We thank Liz St John for excellent technical assistance; Monica Antenos, Allison Tscherner, and Sayed Ahmed for helpful discussions; and the Canadian Centre for Computational Genomics (c3g) for their help and advice in the early phase of small RNA analysis.

References

Abd El Naby WS, Hagos TH, Hossain MM, Salilew-Wondim D, Gad AY, Rings F, Cinar MU, Tholen E, Looft C, Schellander K *et al*. 2013 Expression analysis of regulatory microRNAs in bovine cumulus oocyte complex and preimplantation embryos. *Zygote* **21** 31–51. (doi:10.1017/S0967199411000566)

Adelson DL, Raison JM & Edgar RC 2009 Characterization and distribution of retrotransposons and simple sequence repeats in the bovine genome. *PNASs* **106** 12855–12860. (doi:10.1073/pnas.0901282106)

Aravin AA & Bourc'his D 2008 Small RNA guides for de novo DNA methylation in mammalian germ cells. *Genes & Development* **22** 970–975. (doi:10.1101/gad.1669408)

Aravin A, Gaidatzis D, Pfeffer S, Lagos-Quintana M, Landgraf P, Iovino N, Morris P, Brownstein MJ, Kuramochi-Miyagawa S, Nakano T *et al*. 2006 A novel class of small RNAs bind to MILI protein in mouse testes. *Nature* **442** 203–207. (doi:10.1038/nature04916)

Barckmann B, Pierson S, Dufourt J, Papin C, Armenise C, Port F, Grentzinger T, Chambeyron S, Baronian G, Desvignes JP *et al*. 2015 Aubergine iCLIP reveals piRNA-dependent decay of mRNAs involved in germ cell development in the early embryo. *Cell Reports* **12** 1205–1216. (doi:10.1016/j.celrep.2015.07.030)

Beraldi R, Pittoggi C, Sciamanna I, Mattei E & Spadafora C 2006 Expression of LINE-1 retroposons is essential for murine preimplantation development. *Molecular Reproduction and Development* **73** 279–287. (doi:10.1002/mrd.20423)

Bernstein E, Kim SY, Carmell MA, Murchison EP, Alcorn H, Li MZ, Mills AA, Elledge SJ, Anderson KsV & Hannon GJ 2003 Dicer is essential for mouse development. *Nature Genetics* **35** 215–217. (doi:10.1038/ng1253)

Brennecke J, Aravin AA, Stark A, Dus M, Kellis M, Sachidanandam R & Hannon GJ 2007 Discrete small RNA-generating loci as master regulators of transposon activity in drosophila. *Cell* **128** 1089–1103. (doi:10.1016/j.cell.2007.01.043)

Brower-Toland B, Findley SD, Jiang L, Liu L, Yin H, Dus M, Zhou P, Elgin SCR & Lin H 2007 Drosophila PIWI associates with chromatin and interacts directly with HP1a. *Genes and Development* **21** 2300–2311. (doi:10.1101/gad.1564307)

Bui LC, Evsikov A V, Khan DR, Archilla C, Peynot N, Hénaut A, Le Bourhis D, Vignon X, Renard JP & Duranthon V 2009 Retrotransposon expression as a defining event of genome reprograming in fertilized and cloned bovine embryos. *Reproduction* **138** 289–299. (doi:10.1530/REP-09-0042)

Bushati N, Stark A, Brennecke J & Cohen SM 2008 Temporal reciprocity of miRNAs and their targets during the maternal-to-zygotic transition in drosophila. *Current Biology* **18** 501–506. (doi:10.1016/j.cub.2008.02.081)

Carmell MA, Girard A, van de Kant HJG, Bourc'his D, Bestor TH, de Rooij DG, Hannon GJ, Kant H & Rooij D 2007 MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. *Developmental Cell* **12** 503–514. (doi:10.1016/j.devcel.2007.03.001)

Chan PP & Lowe TM 2009 GtRNAdb: a database of transfer RNA genes detected in genomic sequence. *Nucleic Acids Research* **37** 93–97. (doi:10.1093/nar/gkn787)

Chen C, Liu J & Xu G 2013 Overexpression of PIWI proteins in human stage III epithelial ovarian cancer with lymph node metastasis. *Cancer Biomarkers : Section A of Disease Markers* **13** 315–321.

Cheng J, Guo J-M, Xiao B-X, Miao Y, Jiang Z, Zhou H & Li Q-N 2011 piRNA, the new non-coding RNA, is aberrantly expressed in human cancer cells. *Clinica Chimica Acta* **412** 1621–1625. (doi:10.1016/j.cca.2011.05.015)

Cora E, Pandey RR, Xiol J, Taylor J, Sachidanandam R, McCarthy AA & Pillai RS 2014 The MID-PIWI module of Piwi proteins specifies nucleotide- and strand-biases of piRNAs. *RNA* **20** 773–781. (doi:10.1261/rna.044701.114)

Deng W, Lin H & Carolina N 2002 miwi, a murine homolog of piwi, encodes a cytoplasmic protein essential for spermatogenesis. *Developmental Cell* **2** 819–830. (doi:10.1016/S1534-5807(02)00165-X)

Fadloun A, Le Gras S, Jost B, Ziegler-Birling C, Takahashi H, Gorab E, Carninci P & Torres-Padilla M-E 2013 Chromatin signatures and retrotransposon profiling in mouse embryos reveal regulation of LINE-1 by RNA. *Nature Structural & Molecular Biology* **20** 332–338. (doi:10.1038/nsmb.2495)

Flemr M, Malik R, Franke V, Nejepinska J, Sedlacek R, Vlahovicek K & Svoboda P 2013 A retrotransposon-driven dicer isoform directs endogenous small interfering RNA production in mouse oocytes. *Cell* **155** 807–816. (doi:10.1016/j.cell.2013.10.001)

García-López J, Hourcade J de D, Alonso L, Cárdenas DB & Del Mazo J 2014 Global characterization and target identification of piRNAs and endo-siRNAs in mouse gametes and zygotes. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms* **1839** 463–475. (doi:10.1016/j.bbagrm.2014.04.006)

Gebert D, Ketting RF, Zischler H & Rosenkranz D 2015 piRNAs from pig testis provide evidence for a conserved role of the Piwi pathway in post-transcriptional gene regulation in mammals. *PLoS ONE* **10** e0124860. (doi:10.1371/journal.pone.0124860)

Gilchrist GC, Tscherner A, Nalpathamkalam T, Merico D & Lamarre J 2016 MicroRNA expression during bovine oocyte maturation and fertilization. *International Journal of Molecular Sciences* **17** 396. (doi:10.3390/ijms17030396)

Giraldez AJ, Mishima Y, Rihel J, Grocock RJ, Van Dongen S, Inoue K, Enright AJ & Schier AF 2006 Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. *Science* **312** 75–79. (doi:10.1126/science.1122689)

Girard A, Sachidanandam R, Hannon GJ & Carmell MA 2006 A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature* **442** 199–202. (doi:10.1038/nature04917)

Gou L-T, Dai P, Yang J-H, Xue Y, Hu Y-P, Zhou Y, Kang J-Y, Wang X, Li H, Hua M-M *et al*. 2014 Pachytene piRNAs instruct massive mRNA elimination during late spermiogenesis. *Cell Research* **24** 680–700. (doi:10.1038/cr.2014.41)

Graf A, Krebs S, Zakhartchenko V, Schwalb B, Blum H & Wolf E 2014*a* Fine mapping of genome activation in bovine embryos by RNA sequencing. *PNAS* **111** 4139–4144. (doi:10.1073/pnas.1321569111)

Graf A, Krebs S, Heininen-Brown M, Zakhartchenko V, Blum H & Wolf E 2014*b* Genome activation in bovine embryos: review of the literature and new insights from RNA sequencing experiments. *Animal Reproduction Science* **149** 46–58. (doi:10.1016/j.anireprosci.2014.05.016)

Gunawardane LS, Saito K, Nishida KM, Miyoshi K, Kawamura Y, Nagami T, Siomi H & Siomi MC 2007 A slicer-mediated mechanism for repeat-associated siRNA 5′ end formation in Drosophila. *Science* **315** 1587–1590. (doi:10.1126/science.1140494)

Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, Huarte M, Zuk O, Carey BW, Cassady JP *et al*. 2009 Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **458** 223–227. (doi:10.1038/nature07672)

Hyttel P, Fair T, Callesen H & Greve T 1997 Oocyte growth, capacitation and final maturation in cattle. *Theriogenology* **47** 23–32. (doi:10.1016/S0093-691X(96)00336-6)

Izumi N, Kawaoka S, Yasuhara S, Suzuki Y, Sugano S, Katsuma S & Tomari Y 2013 Hsp90 facilitates accurate loading of precursor piRNAs into PIWI proteins. *RNA* **19** 896–901. (doi:10.1261/rna.037200.112)

Jiang H & Wong WH 2008 SeqMap: mapping massive amount of oligonucleotides to the genome. *Bioinformatics* **24** 2395–2396. (doi:10.1093/bioinformatics/btn429)

Jiang Z, Sun J, Dong H, Luo O, Zheng X, Obergfell C, Tang Y, Bi J, O'Neill R, Ruan Y *et al*. 2014*a* Transcriptional profiles of bovine in vivo pre-implantation development. *BMC Genomics* **15** 756. (doi:10.1186/1471-2164-15-756)

Jiang S, Zhao L, Lu Y, Wang M, Chen Y, Tao D, Liu Y, Sun H, Zhang S & Ma Y 2014*b* Piwil2 inhibits keratin 8 degradation through promoting p38-induced phosphorylation to resist Fas-mediated apoptosis. *Molecular and Cellular Biology* **34** 3928–3938. (doi:10.1128/MCB.00745-14)

Kawaoka S, Izumi N, Katsuma S & Tomari Y 2011 3′ end formation of PIWI-interacting RNAs in vitro. *Molecular Cell* **43** 1015–1022. (doi:10.1016/j.molcel.2011.07.029)

Kirino Y & Mourelatos Z 2007 2′-O-methyl modification in mouse piRNAs and its methylase. *Nucleic Acids Symposium Series* **51** 417–418. (doi:10.1093/nass/nrm209)

Kozomara A & Griffiths-Jones S 2014 miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Research* **42** D68–D73. (doi:10.1093/nar/gkt1181)

Kuramochi-Miyagawa S, Kimura T, Ijiri TW, Isobe T, Asada N, Fujita Y, Ikawa M, Iwai N, Okabe M, Deng W *et al*. 2004 Mili, a mammalian member of piwi family gene, is essential for spermatogenesis. *Development* **131** 839–849. (doi:10.1242/dev.00973)

Kuramochi-Miyagawa S, Watanabe T, Gotoh K, Totoki Y, Toyoda A, Ikawa M, Asada N, Kojima K, Yamaguchi Y, Ijiri TW *et al*. 2008 DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. *Genes & Development* **22** 908–917. (doi:10.1101/gad.1640708)

Memili E & First NL 2000 Zygotic and embryonic gene expression in cow: a review of timing and mechanisms of early gene expression as compared with other species. *Zygote* **8** 87–96. (doi:10.1017/S0967199400000861)

Mondou E, Dufort I, Gohin M, Fournier E & Sirard M 2012 Analysis of microRNAs and their precursors in bovine early embryonic development. *Molecular Human Reproduction* **18** 425–434. (doi:10.1093/molehr/gas015)

Morgan HD, Santos F, Green K, Dean W & Reik W 2005 Epigenetic reprogramming in mammals. *Human Molecular Genetics* **14** R47–R58. (doi:10.1093/hmg/ddi114)

Pantano L, Jodar M, Bak M, Ballescà JL, Tommerup N, Oliva R & Vavouri T 2015 The small RNA content of human sperm reveals pseudogene-derived piRNAs complementary to protein-coding genes. *RNA* **21** 1085–1095. (doi:10.1261/rna.046482.114)

Peaston AE, Evsikov AV, Graber JH, de Vries WN, Holbrook AE, Solter D & Knowles BB 2004 Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. *Developmental Cell* **7** 597–606. (doi:10.1016/j.devcel.2004.09.004)

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J & Glöckner FO 2013 The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research* **41** D590–D596. (doi:10.1093/nar/gks1219)

R Development Core Team 2013 R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/. s*R Foundation for Statistical Computing*, *Vienna*, Austria.

Rajasethupathy P, Antonov I, Sheridan R, Frey S, Sander C, Tuschl T & Kandel ER 2012 A role for neuronal piRNAs in the epigenetic control of memory-related synaptic plasticity. *Cell* **149** 693–707. (doi:10.1016/j.cell.2012.02.057)

Reddien PW 2005 SMEDWI-2 is a PIWI-like protein that regulates planarian stem cells. *Science* **310** 1327–1330. (doi:10.1126/science.1116110)

Reuter M, Berninger P, Chuma S, Shah H, Hosokawa M, Funaya C, Antony C, Sachi danandam R & Pillai RS 2011 Miwi catalysis is required for piRNA amplification-independent LINE1 transposon silencing. *Nature* **480** 264–267. (doi:10.1038/nature10672)

Rizzo F, Hashim A, Marchese G, Ravo M, Tarallo R, Nassa G, Giurato G, Rinaldi A, Cordella A, Persico M *et al*. 2014 Timed regulation of P-element-induced wimpy testis-interacting RNA expression during rat liver regeneration. *Hepatology* **60** 798–806. (doi:10.1002/hep.27267)

Roovers EF, Rosenkranz D, Mahdipour M, Han CT, He N, Chuva de Sousa Lopes SM, van der Westerlaken LAJ, Zischler H, Butter F, Roelen BAJ *et al*. 2015 Piwi proteins and piRNAs in mammalian oocytes and early embryos. *Cell Reports* **10** 2069–2082. (doi:10.1016/j.celrep.2015.02.062)

Rosenkranz D 2016 piRNA cluster database: a web resource for piRNA producing loci. *Nucleic Acids Research* **44** D223–D230. (doi:10.1093/nar/gkv1265)

Rosenkranz D & Zischler H 2012 proTRAC - a software for probabilistic piRNA cluster detection, visualization and analysis. *BMC Bioinformatics* **13** 5. (doi:10.1186/1471-2105-13-5)

Rosenkranz D, Rudloff S, Bastuck K, Ketting RF & Zischler H 2015 Tupaia small RNAs provide insights into function and evolution of RNAi-based transposon defense in mammals. *RNA* **21** 911–922. (doi:10.1261/rna.048603.114)

Rouget C, Papin C, Boureux A, Meunier A-CA-C, Franco BB, Robine N, Lai EC, Pelisson A & Simonelig M 2010 Maternal mRNA deadenylation and decay by the piRNA pathway in the early Drosophila embryo. *Nature* **467** 1128–1132. (doi:10.1038/nature09465)

Russell SJ, Stalker L, Gilchrist G, Backx A, Molledo G, Foster RA & LaMarre J 2016 Identification of PIWIL1 isoforms and their expression in bovine testes, oocytes, and early embryos. *Biology of Reproduction* **94** 1–11. (doi:10.1095/biolreprod.115.131854)

Saito K, Sakaguchi Y, Suzuki T, Suzuki T, Siomi H & Siomi MC 2007 Pimet, the Drosophila homolog of HEN1, mediates 2′-O-methylation of PIWI-interacting RNAs at their 3′ ends. *Genes & Development* **21** 1603–1608. (doi:10.1101/gad.1563607)

Shoji M, Tanaka T, Hosokawa M, Reuter M, Stark A, Kato Y, Kondoh G, Okawa K, Chujo T, Suzuki T *et al*. 2009 The TDRD9-MIWI2 complex is essential for piRNA-mediated retrotransposon silencing in the mouse male germline. *Developmental Cell* **17** 775–787. (doi:10.1016/j.devcel.2009.10.012)

Siomi MC, Sato K, Pezic D & Aravin AA 2011 PIWI-interacting small RNAs: the vanguard of genome defence. *Nature Reviews. Molecular Cell Biology* **12** 246–258. (doi:10.1038/nrm3089)

Sirard MA & Blondin P 1996 Oocyte maturation and IVF in cattle. *Animal Reproduction Science* **42** 417–426. (doi:10.1016/0378-4320(96)01518-7)

Sun H, Li D, Chen S, Liu Y, Liao X, Deng W, Li N, Zeng M, Tao D & Ma Y 2010 Zili inhibits transforming growth factor-beta signaling by interacting with Smad4. *Journal of Biological Chemistry* **285** 4243–4250. (doi:10.1074/jbc.M109.079533)

Tanaka Y, Chung L & Park IH 2012 Impact of retrotransposons in pluripotent stem cells. *Molecules and Cells* **34** 509–516.

Tang F, Kaneda M, O'Carroll D, Hajkova P, Barton SC, Sun YA, Lee C, Tarakhovsky A, Lao K & Surani MA 2007 Maternal microRNAs are essential for mouse zygotic development. *Genes and Development* **21** 644–648. (doi:10.1101/gad.418707)

Thomsen MCF & Nielsen M 2012 Seq2Logo: a method for construction and visualization of amino acid binding motifs and sequence profiles including sequence weighting, pseudo counts and two-sided representation of amino acid enrichment and depletion. *Nucleic Acids Research* **40** 281–287. (doi:10.1093/nar/gks469)

Tripurani SK, Lee K-B, Wee G, Smith GW & Yao J 2011 MicroRNA-196a regulates bovine newborn ovary homeobox gene (NOBOX) expression during early embryogenesis. *BMC Developmental Biology* **11** 25. (doi:10.1186/1471-213X-11-25)

Vagin VV, Sigova A, Li C, Seitz H, Gvozdev V & Zamore PD 2006 A distinct small RNA pathway silences selfish genetic elements in the germline. *Science* **313** 320–324. (doi:10.1126/science.1129333)

Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A & Speleman F 2002 Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology* **3** RESEARCH0034. (doi:10.1186/gb-2002-3-7-research0034)

Yang C-X, Wright EC, Rothschild MF, Prather RS, Ross JW, Du Z-Q, Wright EC, Rothschild MF, Prather RS & Ross JW 2012 Small RNA profile of the cumulus-oocyte complex and early embryos in the pig. *Biology of Reproduction* **87** 117. (doi:10.1095/biolreprod.111.096669)

Zhang P, Kang J-Y, Gou L-T, Wang J, Xue Y, Skogerboe G, Dai P, Huang D-W, Chen R, Fu X-D *et al*. 2015 MIWI and piRNA-mediated cleavage of messenger RNAs in mouse testes. *Cell Research* **25** 193–207. (doi:10.1038/cr.2015.4)

Received 29 August 2016
First decision 30 September 2016
Revised manuscript received 20 November 2016
Accepted 13 December 2016