

# Comparative profiling of small RNAs of pig seminal plasma and ejaculated and epididymal sperm

Cai Chen<sup>1,2</sup>, Han Wu<sup>2</sup>, Dan Shen<sup>2</sup>, Saisai Wang<sup>2</sup>, Li Zhang<sup>2</sup>, Xiaoyan Wang<sup>2</sup>, Bo Gao<sup>2</sup>, Tianwen Wu<sup>1</sup>, Bichun Li<sup>2</sup>, Kui Li<sup>1</sup> and Chengyi Song<sup>1,2</sup>

<sup>1</sup>The Key Laboratory for Domestic Animal Genetic Resources and Breeding of the Ministry of Agriculture of China, Institute of Animal Science, Chinese Academy of Agricultural Sciences, Beijing, China and <sup>2</sup>Joint International Research Laboratory of Agriculture and Agri-Product Safety, College of Animal Science & Technology, Yangzhou University, Yangzhou, Jiangsu, China

Correspondence should be addressed to C Song; Email: [cysong@yzu.edu.cn](mailto:cysong@yzu.edu.cn)

## Abstract

The similarities and differences of small RNAs in seminal plasma, epididymal sperm and ejaculated sperm remain largely undefined. We conducted a systematic comparative analysis of small RNA profiles in pig ejaculated sperm, epididymal sperm and seminal plasma and found that the diversity distribution of small RNA species was generally similar, whereas the abundance of small RNAs is dramatically different across the three libraries; miRNAs and small RNAs derived from rRNA, tRNA, small nuclear RNA, 7SK RNA, NRON RNA and cis-regulatory RNA were enriched in the three libraries, but piRNA was absent. A large population of small RNAs from ejaculated sperm are ejaculated sperm specific, and only 8–30% of small RNAs overlapped with those of epididymal sperm or seminal plasma and a small proportion (5–18%) of small RNAs were shared in the three libraries, suggesting that, in addition to the testes, sperm RNAs may also originate from seminal plasma, epididymis as well as other resources. Most miRNAs were co-distributed but differentially expressed across the three libraries, with epididymal sperm exhibiting the highest abundance, followed by ejaculated sperm and seminal plasma. The prediction of target genes of the top 10 highly expressed miRNAs across the three libraries revealed that these miRNAs may be involved in spermatogenesis, zygote development and the interaction between the environment and animals. Our study provides the first description of the similarities and differences of small RNA profiles in ejaculated sperm, epididymal sperm and seminal plasma and indicates that sperm RNA may have origins other than the testes.

*Reproduction* (2017) 153 785–796

## Introduction

For a long time, sperm RNA has been considered the residue of spermatogenesis and lacking in function because of its limited amount and special anatomical position. However, both coding and non-coding RNAs have been well characterized in sperm (Jodar *et al.* 2013), and it has been revealed that sperm can carry diverse RNA species (including mRNAs and small RNAs, some of which are sperm specific) to the fertilized zygote (Ostermeier *et al.* 2004, Stoeckius *et al.* 2014), indicating that these RNAs may play a unique paternal role in the development of zygotes.

The presence of coding RNAs in sperms of humans (Lambard *et al.* 2004), pigs (Yang *et al.* 2009), bovines (Gilbert *et al.* 2007, Lalancette *et al.* 2008), mice (Fang *et al.* 2014) and stallions (Das *et al.* 2013) is well established. Even though the functions of the majority of the sperm mRNAs remain enigmatic, it has been proposed that sperm mRNAs may play roles in the regulation of sperm morphology, motility and

capacitation and that sperm transcriptional profiles might provide clinical markers for male fertility (Lambard *et al.* 2004, Li *et al.* 2007, Bansal *et al.* 2015). Similarly, sperm transcriptome studies have been initiated in bulls and have revealed the differences between the mRNA profiles of high- and low-fertility males (Lalancette *et al.* 2008, Bissonnette *et al.* 2009). Recently, the sperm-borne mRNA Wnt4 was found to be delivered into the fertilized oocyte and translated, which represents evidence of a paternal effect on zygotic development (Fang *et al.* 2014); however, the functional significance of sperm mRNAs in the development of zygotes remains largely unknown.

The complexity of non-coding RNA species has also been well characterized in the sperm (Krawetz *et al.* 2011, Peng *et al.* 2012, Jodar *et al.* 2013, Sandler *et al.* 2013, García-lópez *et al.* 2015, Pantano *et al.* 2015, Schuster *et al.* 2016), and it is now commonly accepted that non-coding RNAs are critical for the process of embryogenesis (Liu *et al.* 2012a, Chen *et al.* 2016a,b, Yuan *et al.* 2016). The miRNA family is the best



characterized of the small RNA species in sperm. They were first identified in humans (Ostermeier *et al.* 2005), followed by confirmation of their presence in mice (Amanai *et al.* 2006) and have now been detected in various mammals, including porcine (Curry *et al.* 2011, Kasimanickam & Kastelic 2016), bulls (Stowe *et al.* 2014), stallions (Das *et al.* 2013), rats and rabbits (Schuster *et al.* 2016). It has been demonstrated that mature miRNAs that are carried into the zygote by sperm during the fertilization are relatively rare compared with oocyte miRNAs (Yang *et al.* 2016). Accumulating evidence shows that miRNAs are critical for fertilization and early embryonic development (Liu *et al.* 2012a,b, Cheong *et al.* 2014, Yuan *et al.* 2016). However, the role of some miRNAs in the development of early embryos remains controversial (Amanai *et al.* 2006, Yuan *et al.* 2015). piRNAs are the second best-known small RNA species in sperm (Krawetz *et al.* 2011, Pantano *et al.* 2015). It has been shown that the human sperm contains over tens of thousands of piRNAs with over 400 clusters and that piRNAs appear to be the most abundant class of small RNAs in mature human sperm and the most abundant sperm piRNAs targeting LINE1 transposons. In addition, pseudogene-derived piRNAs complementary to protein-coding genes, which may regulate the expression of their parent gene in the male germ line, have been identified in the human sperm (Krawetz *et al.* 2011, Pantano *et al.* 2015). Moreover, Stoeckius and coworkers found that approximately half of all assayed endogenous piRNAs and siRNAs of 1-cell embryos in *Caenorhabditis elegans* are of sperm origin (Stoeckius *et al.* 2014). The enrichment of tRNA-derived small RNAs (tsRNAs) in the sperm was first described in mice (Peng *et al.* 2012), and then extended to humans, rats and rabbits (Schuster *et al.* 2016). Currently, it is suggested as a paternal epigenetic factor that may mediate intergenerational inheritance (Chen *et al.* 2016a). The sperm-borne endo-siRNAs also appear to be important for the fertilization and the pre-implantation embryonic development (Yuan *et al.* 2016). However, the function of most sperm-borne small RNA populations in the development of zygotes is less clear.

Although extensive studies of sperm RNAs have been reported in mammal species, especially in humans and mice, and although they are all called sperm RNAs, these sperm species actually represent different statuses. All human sperm is purified from ejaculated sperm, whereas all mice sperm is collected from the epididymis, because it is almost technically impossible to collect human epididymal sperm and mouse ejaculated sperm samples, which renders them unavailable for RNA sequencing. Therefore, it is impossible to compare RNA profiles between ejaculated sperm and epididymal sperm in the same species (humans or mice). Seminal plasma is also relatively easy to collect in humans than that in mice, and several RNA species have been identified within seminal plasma (Wu *et al.* 2012, Xiong 2014, Hong *et al.* 2016). However, comparatively

little is known about the distribution of RNAs in epididymal sperm, seminal plasma and ejaculated sperm. Pigs, as an important animal model with physiological parameters and high genomic conservation similar to the humans, have unique reproductive features and produce a large volume of semen with a high amount of sperm in each ejaculation. Moreover, it is relatively easier to collect ejaculated sperm, seminal plasma and seminal plasma samples from pigs than those from other mammals. In the current study, we conducted a systematic comparative analysis of the distribution of small RNAs in fully matured sperm within the cauda epididymis, ejaculated sperm and seminal plasma in pigs. Integrating analyses of the small RNA content across these samples could reveal similarities and differences of small RNA profiles, which currently remain unknown, and may allow the origin of these small RNAs to be understood, in addition to their potential functions and significance. Our study provides the first description of the broad variety of small RNAs present in pig ejaculated sperm, epididymal sperm and seminal plasma and of the differences and similarities of small RNA profiles among these samples.

## Methods

### Sample preparation

The experimental procedure was approved and supervised by the Animal Care Commission of Yangzhou University, and all methods were performed in accordance with the relevant guidelines and regulations. Fresh ejaculated sperm was collected from three adult Meishan boars (aged over 2 years), filtered and then evaluated under a microscope. Only samples that exhibited more than 75% of the sperm with good motility and normal morphology were retained and diluted in the CXM medium (50 g/L, Guangzhou Bai Xu United Animal Husbandry Science and Technology Development Ltd. Co., Guangzhou, China) at the rate of 1:2 (v:v). The sperm was then purified using a two-step (swim-up and density-gradient centrifugation) protocol. In brief, the sperm swam up in a CO<sub>2</sub> incubator at 37°C for 60 min, which contained 5% CO<sub>2</sub> and which was placed at a 45° angle. Subsequently, three-quarters of the solution described previously was collected and centrifuged at 679 g for 7 min at 17°C to collect the precipitate. Finally, the sperm pellet was resuspended with 2 mL of CXM for the next step of purification by density-gradient centrifugation with PureSperm according to the manufacturer's instructions (Nidacon International AB, Gothenburg, Sweden). In brief, 80% and 40% PureSperm solutions were first prepared, and then, 2 mL of 80% PureSperm was added into a 15 mL conical tube, and 2 mL of 40% PureSperm was carefully added on top of the 80% layer. Finally, the resuspended sperm pellet was carefully added onto the surface of the gradient and centrifuged at 300 g for 20 min at 17°C. The surface layer was removed carefully to leave the sperm pellet within 4–6 mm of the 80% PureSperm layer untouched. Subsequently, the sperm pellet was transferred to a new tube and resuspended in 5 mL PureSperm Wash for a final centrifugation at 500 g for

10 min at 17°C. Finally, the PureSperm Wash supernatant was removed carefully and the pellet was collected, and a small amount of sperm precipitate was resuspended for detection by microscopy. The remaining sperm pellet was immediately frozen in liquid nitrogen and subjected to RNA extraction.

The seminal plasma samples were collected as described previously, with small modifications (Li *et al.* 2012). In brief, three samples of fresh semen were collected from the same three boars, filtered as described previously and diluted with an equal volume of CXM medium. The diluted semen was centrifuged at 1000g for 20 min at 4°C to remove the sperm and other cells. The supernatant was transferred to new 2 mL tubes and centrifuged twice at 10,000g for 10 min at 4°C to remove the remaining cells and cell debris. The supernatant (seminal plasma) was then carefully collected and immediately frozen in liquid nitrogen.

The cauda epididymis was dissected from the same three boars. Sperm was squeezed out and flushed from the cauda epididymis and then filtered and diluted according to the protocol used for the ejaculated sperm. The same two-step purification protocol described previously was used to purify the epididymal sperm. The final clean epididymal sperm pellet was frozen in liquid nitrogen.

### Small RNA library construction and sequencing

Total RNA was extracted using the TRIzol (Invitrogen, Thermo Fisher Scientific) according to the manufacturer's protocol, and each sample was treated with RNase-free DNase I and assessed for the absence of genomic DNA contamination by RT-PCR. Then, the RNA samples were subjected to purity qualification by RT-PCR amplification of *CD4*, *E-cadherin* and *c-kit*, which are markers of leukocytes, epithelial cells and testicular germ cells respectively, as described previously (Song *et al.* 2011). *GAPDH* and *AQN-3* were used as positive controls for sperm RNA and testis cDNA. The GenBank accession number and primer sequences for the *CD4*, *c-kit*, *E-cadherin*, *GAPDH* and *AQN-3* genes are listed in [Supplementary Table 1](#) (see section on [supplementary data](#) given at the end of this article). Only RNA samples without leukocyte or epithelial and germ cell contamination were used for further analysis. RNA integrity was confirmed using a 2100 Bioanalyzer (Agilent Technologies). The validated RNA samples from three boars (1000 ng RNA from each) were equally pooled, and then subjected to 12% (w/v) denaturing polyacrylamide gel electrophoresis (PAGE), and the small RNA fragments 18–40 nucleotides were isolated. Then, 100 ng of each small RNA sample was used to prepare for the small RNA library construction with the small RNA Sample Prep Kit (Illumina) according to the manufacturer's recommendations. In brief, after ligation to the 5' and 3' oligonucleotide RNA adaptors using T4 RNA ligase, total RNA was subjected to RT-PCR with 15 cycles of amplification. The adaptor-ligated small RNA (140–160 nucleotides) was purified and used directly for the cluster generation and sequencing analysis using an Illumina/Solexa sequencer (Shanghai Oebiotech Co. Ltd, Shanghai, China). Finally, the samples were analyzed again after small RNA library construction to confirm the size of the products. Only samples that passed each quality control step were

processed for the next-generation sequencing. Sequencing was processed on an Illumina HiSeq 2500 apparatus in single-ended mode 1 × 36 bp.

### Data analysis

Small RNA reads were generated by Illumina HiSeq analysis. All sequencing data underwent stringent quality control measures prior to the analysis. In brief, low-quality reads, reads shorter than 18 nucleotides or longer than 40 nucleotides and reads with 5' primer contaminants and ambiguous bases were removed from the raw data, and the clean reads were obtained and mapped against the pig genomes. These clean reads were then annotated with the following databases and programs: (1) miRBase (version 21.0: <http://www.mirbase.org/index.shtml>) to identify known miRNAs using the Bowtie software; (2) Rfam (<http://rfam.xfam.org/>), to identify other known small RNA clusters such as ribosomal RNA (rRNA), transfer RNA (tRNA), small nuclear RNA (snRNA), small nucleolar RNAs (snoRNAs) and so forth by the BLASTN program; (3) pig mRNA and Unigene databases to identify the degradation fragments of mRNA using the Bowtie software and (4) Rebase to identify repeat sequences using the RepeatMasker program. Subsequently, the remaining unannotated clean reads that could be aligned to the genome were analyzed by miRDeep2 to predict novel miRNAs. Based on the hairpin structure of the pre-miRNA and the miRBase databases, the corresponding miRNA start sequence was also identified. The final remaining unannotated clean reads were aligned to the known piRNAs in the pig gonads (Kowalczykiewicz *et al.* 2014) to identify the putative piRNA species using the Bowtie software.

The abundances of miRNAs were calculated using transcript per million reads (TPM).  $TPM = (\text{the number of reads that can be matched to each miRNA}) / (\text{the number of total miRNA reads}) \times 10^6$ . *P* values were determined using the Audic–Claverie formula, and miRNAs with  $P < 0.05$  and TPM fold change  $> 2$  between two libraries were defined as differentially expressed miRNAs.

The miRanda software was used to predict target genes of miRNAs. Functional annotation of the predicted target genes of miRNAs was performed using the Gene Ontology (GO, <http://amigo.geneontology.org/amigo>) and the Kyoto Encyclopaedia of Genes and Genomes (KEGG, <http://www.genome.jp/kegg/>) programs. Acquired genes were subjected to GO and KEGG enrichment analyses to determine the pathways that were significantly represented by the miRNA group. GO and KEGG enrichments were assessed using the Database for Annotation, Visualization and Integrated Discovery (DAVID, v6.7, <http://david.abcc.ncifcrf.gov/home.jsp>). Only gene categories with  $P < 0.05$  were considered significantly enriched.

## Results

### Evaluation of the purity of sperm and RNA samples

After the purification of sperm using the swim-up and density-gradient centrifugation, the purity of ejaculated and epididymal sperm samples was evaluated under a microscope, and all sperms had normal morphology

**Table 1** Overview of sequencing data.

Sample	Raw reads	Clean reads (% of raw reads)	Number of unique sequences	Annotated reads (% of clean reads)
Epididymal sperm	6,417,531	5,554,557 (86.55%)	300,987	3,478,686 (62.63%)
Ejaculated sperm	6,741,724	5,080,463 (75.36%)	227,288	1,336,554 (26.31%)
Seminal plasma	6,734,025	4,866,597 (72.27%)	64,419	115,939 (2.38%)

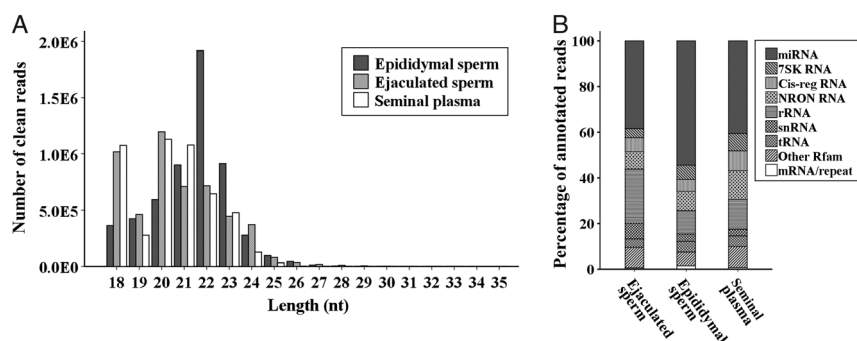
profiles after dilution. Cells other than sperm cells were not observed in the sample as showed in the representative figure of the ejaculated sperm sample (Supplementary Fig. 1A), and the contamination from leukocytes, epithelial cells or testicular germ cells were excluded by RT-PCR against the specific markers *CD4*, *E-cadherin* or *c-kit* respectively, whereas primers against the *GAPDH* mature transcript amplified a fragment of the correct length (Supplementary Fig. 1B). The overall profile of the extracted RNAs based on analysis using Agilent 2100 confirmed the absence of intact ribosomal RNA and the predominance of short RNA molecules, which are well-known characteristics of the sperm RNA and also precludes RNA contamination from somatic cells (Supplementary Fig. 1C, D and E).

#### Diverse small RNA classes were enriched in ejaculated sperm, seminal plasma and epididymal sperm

The total small RNA reads and genome-mapping statistical data (Table 1) showed that, although the total clean reads were similar across the three libraries, the abundance of small RNAs was dramatically different, with the highest annotated reads in the library of epididymal sperm (3,478,686), which accounted for 62.63% of the total number of clean reads and was in contrast with the libraries of ejaculated sperm and seminal plasma, in which annotated reads only made up 26.31% (1,336,554) and 2.38% (115,939) of the total number of clean reads respectively. The abundance of so many unannotated small RNAs in seminal plasma indicated that there may be many novel small RNA species that remain undiscovered. The overall length distribution of small RNAs revealed that the dominant reads from pig seminal plasma, epididymal sperm and ejaculated sperm are 18–26 nucleotides (Fig. 1A),

which is distinct from that observed in adult testes (Kowalczykiewicz *et al.* 2014, Gebert *et al.* 2015), where two peaks (18–25 nucleotides and 26–32 nucleotides) are observed, and in human sperm (24–31 nucleotides) (Pantano *et al.* 2015). The majority of the 26–32 nucleotide small RNAs in the pig testes and the human sperm are piRNAs (Kowalczykiewicz *et al.* 2014, Gebert *et al.* 2015, Pantano *et al.* 2015), which are actively involved in the retrotransposon silencing that protects the integrity of the genome. In addition, piRNAs can be observed in human seminal plasma (Hong *et al.* 2016). To exclude the existence of piRNAs in these samples, the remaining unannotated reads of them were also aligned to the known piRNAs in the pig gonads (Kowalczykiewicz *et al.* 2014), and we found that only 0.03% of the unannotated reads of epididymal sperm were identified as piRNAs, whereas piRNA was not detectable in seminal plasma and ejaculated sperm samples. The absence of piRNAs in seminal plasma, ejaculated and epididymal sperm was somewhat surprising and suggests that the distribution of sperm piRNAs varies from species to species. In addition, the peak length value of seminal plasma and ejaculated sperm small RNAs was 20 nucleotides, which is slightly different from that observed for epididymal sperm, in which the overwhelming majority of the small RNAs are concentrated at 22 nucleotides, followed by 21 and 23 nucleotides (Fig. 1A).

Further annotation of these small RNAs revealed that, although the abundance of the annotated small RNAs was significantly different across the three libraries, the diversity of the small RNAs within the three libraries was similar. Several small RNA classes (Fig. 1B and Supplementary Table 2), including miRNA and small RNAs derived from rRNA, tRNA, small nuclear RNA (snRNA), 7SK RNA, NRON RNA (known as a non-

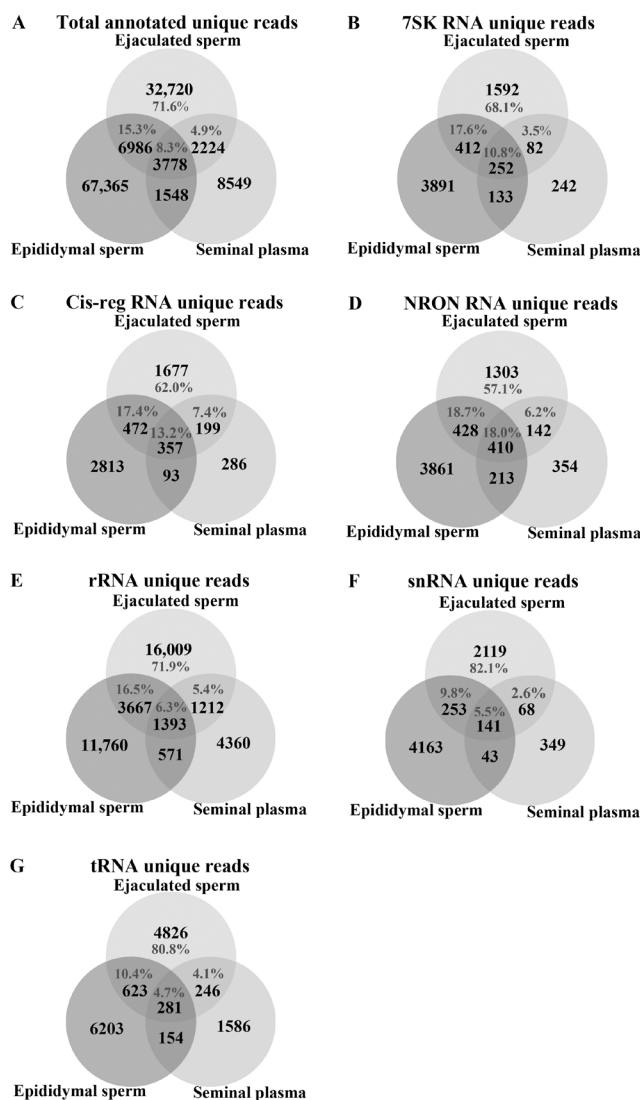


**Figure 1** Identification and characterization of small RNAs in pig seminal plasma, epididymal sperm and ejaculated sperm. (A) Length distribution of small RNAs in seminal plasma, epididymal sperm and ejaculated sperm. (B) Sources of small RNAs in each library.

coding RNA repressor of the nuclear factor of activated T cells) and cis-regulatory RNA (cis-reg RNA), were enriched in boar epididymal sperm, ejaculated sperm and seminal plasma. Most of these small RNAs were clustered in the miRNA group, which represented 54.44, 38.40 and 40.54% of the total reads annotated in boar epididymal sperm, ejaculated sperm and seminal plasma respectively. The second most abundant group was that of rRNA-derived small RNAs, representing 10.14, 23.59 and 13.15% of the total reads annotated in boar epididymal sperm, ejaculated sperm and seminal plasma respectively. The other small RNAs derived from NRON RNA, cis-reg RNA, tRNA, snRNA and 7SK RNA also represented a substantial portion (2.83–12.52%) of the reads annotated across the three libraries. The NRON RNA is a non-coding RNA involved in repressing the nuclear factor of activated T cells (NFAT) (Willingham *et al.* 2005), which is a general name given to a family of transcription factors that are important in the immune response (Macian 2005). The 7SK RNA is an abundant small nuclear RNA found in the metazoans that play a role in the regulation of transcription by controlling the positive transcription elongation factor P-TEFb within a small nuclear ribonucleoprotein complex (snRNP) (Wassarman & Steitz 1991, Diribarne & Bensaude 2009, Peterlin *et al.* 2012). In contrast, cis-reg RNAs are RNA motifs that regulate nucleic acid regions on the same molecule and are believed to be involved in the ribosome-mediated control of gene expression (Xue & Barna 2015). The enrichment of small RNA species derived from NRON RNA, cis-reg RNA and 7SK RNA in a mammal sperm was reported for the first time in this study; however, the functional significance of these small RNAs in sperm remains unknown.

### Most small RNAs of ejaculated sperm are ejaculated sperm specific

A comparative analysis of the distribution of small RNAs between seminal plasma, epididymal sperm and ejaculated sperm revealed the following: over 70% of the total unique reads (71.6%) of small RNAs in ejaculated sperm library were ejaculated sperm specific, and only 8.3% of the unique reads (71.6%) of small RNAs in ejaculated sperm library were shared among the three libraries, approximately 25% of the unique reads of ejaculated sperm overlapped with those of epididymal sperm, and approximately 13% of the unique reads of ejaculated sperm overlapped with those of seminal plasma (Fig. 2A). An in-depth comparison of the characterized classes of small RNAs showed that most unique reads of small RNAs derived from 7SK RNA (68.1%), cis-reg RNA (62.0%), NRON RNA (57.1%), rRNA (71.9%), snRNA (82.1%) and tRNA (80.8%) were ejaculated sperm specific; only 5–18% of the unique reads of these small RNA classes were co-distributed among the three libraries; and 15–30% of the unique



**Figure 2** Pairwise comparisons of the small RNA reads in the three libraries. (A) Pairwise comparisons of the total annotated unique reads of small RNAs in the three libraries. Pairwise comparisons of the unique reads of small RNAs derived from 7SK RNA (B), cis-reg RNA (C), NRON RNA (D), rRNA (E), snRNA (F) and tRNA (G).

reads of these small RNA classes were co-distributed only between ejaculated sperm and epididymal sperm, 8–20% of the unique reads of these small RNA classes were co-distributed only between ejaculated sperm and seminal plasma (Fig. 2B, C, D, E, F and G). These data revealed that only a small population of small RNAs of ejaculated sperm is derived from epididymis and seminal plasma, and a large population of small RNAs of ejaculated sperm may have other origins.

### Most annotated miRNA species of ejaculated sperm overlapped with epididymal sperm and seminal plasma

The miRNAs from 47,075, 1,913,178 and 516,589 aligned reads within seminal plasma, epididymal sperm

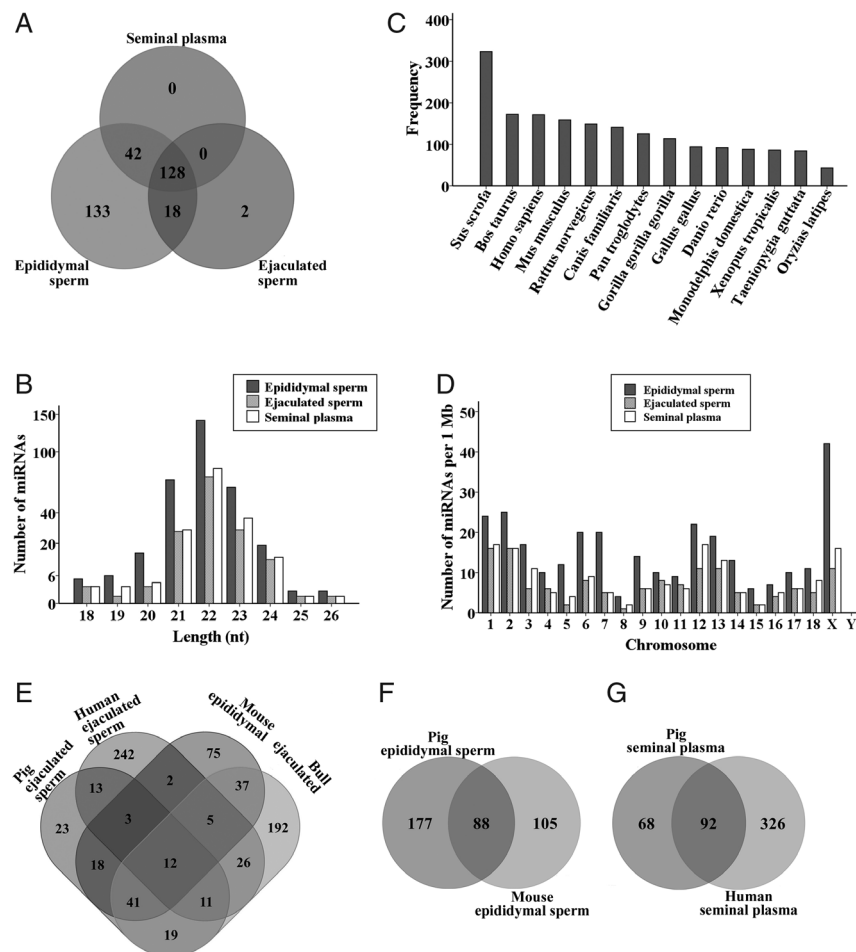
**Table 2** Total number of known and novel miRNAs from the three libraries.

Sample name	Epididymal sperm	Ejaculated sperm	Seminal plasma	Total
Aligned reads (total)	1,913,178	516,589	47,075	2,476,842
Number of known miRNAs	265	140	160	266
Number of novel miRNAs	56	8	10	57
Total miRNA	321	148	170	323

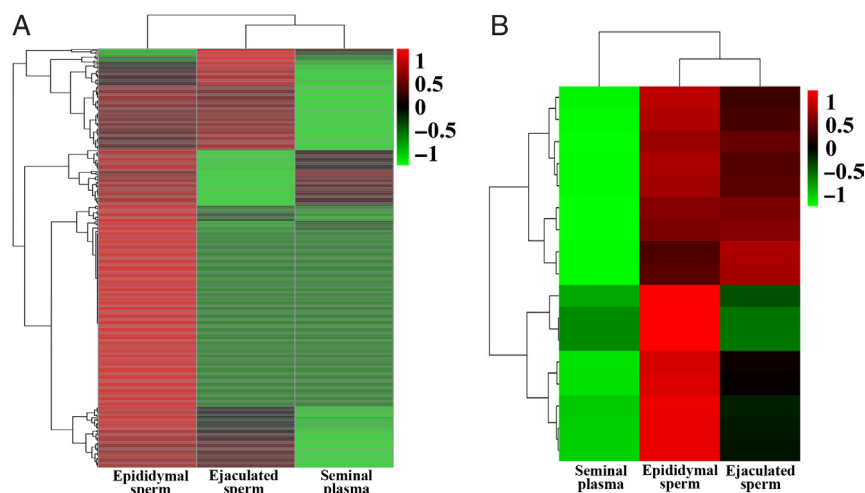
and ejaculated sperm respectively were characterized using the miRBase 21.0 database and the miRDeep2 software. In total, 323 miRNAs, including 266 known and 57 novel miRNAs in pig, are annotated (Table 2 and Supplementary Table 3). The greatest number of miRNA species was detected within epididymal sperm followed by seminal plasma and ejaculated sperm. A Venn diagram (Fig. 3A) revealed that approximately 40% of the miRNA species were co-expressed in the three libraries; 98.65% and 86.47% of miRNA species of ejaculated sperm overlapped with those of epididymal sperm and seminal plasma respectively; 100.00% and 75.29% of seminal plasma miRNA species overlapped with those of epididymal sperm and ejaculated sperm respectively and 133 and 2 miRNAs were epididymal sperm and ejaculated sperm specific respectively (no specific miRNA was identified in seminal plasma). These data suggest that

most miRNA species of ejaculated sperm overlap with those of epididymal sperm and seminal plasma.

The length distribution of these miRNAs ranged from 18 to 26 nucleotides (Fig. 3B), with the majority of sequences having between 21 and 24 nucleotides and with a large representation at 22 nucleotides, which accounted for 43.92, 50.71 and 47.64% of the total miRNA reads in epididymal sperm, ejaculated sperm and seminal plasma respectively. This length is highly consistent with the common size of miRNAs from Dicer digestion products. The sequence conservation of the miRNAs identified revealed that over one-third were conserved across most mammals investigated (Fig. 3C). The density analysis of the miRNAs across the porcine chromosomes revealed that the miRNAs in seminal plasma, epididymal sperm and ejaculated sperm exhibited a similar chromosome distribution pattern



**Figure 3** Identification and characterization of miRNAs in seminal plasma, epididymal sperm and ejaculated sperm of pigs. (A) Pairwise comparisons of the miRNA species in three libraries. (B) Length distributions of miRNAs in seminal plasma, epididymal sperm and ejaculated sperm. (C) The sequence conservation of the identified miRNAs in vertebrates. (D) Chromosome distribution of all identified miRNAs in seminal plasma, epididymal sperm, and ejaculated sperm. (E) Pairwise comparisons of the miRNA species in the ejaculated sperm of human, pig and bull and the epididymal sperm of mice. (F) Pairwise comparisons of the miRNA species in the pig and mouse epididymal sperm. (G) Pairwise comparisons of the miRNA species in the seminal plasma of pigs and humans.



**Figure 4** Expression of all identified miRNAs (A) and the top 10 most highly expressed miRNAs (B) in seminal plasma, epididymal sperm and ejaculated sperm.

(Fig. 3D) and are mapped to almost all chromosomes, with the exception of the Y chromosome, but showed a high preference for chromosome 12 and, to a lesser extent, for chromosomes 2, 10, 11, 17, 18 and X. A significant fraction of miRNAs that originated from epididymal sperm were enriched on the X chromosome compared with those from epididymal sperm and seminal plasma. The enrichment of miRNAs on the X chromosome has been noted previously (Guo *et al.* 2009), and these miRNAs may play roles in spermatogenesis. Moreover, a large proportion of the miRNAs identified in pig seminal plasma, epididymal sperm and ejaculated sperm were also present in human seminal plasma (Wang *et al.* 2011) and ejaculated sperm (Krawetz *et al.* 2011), mouse epididymal sperm (Amanai *et al.* 2006) and bull ejaculated sperm (Stowe *et al.* 2014) (Fig. 3E, F, G, and Supplementary Table 3). These results suggest that miRNAs in sperm serve as an ancient paternal element with evolutionarily conserved functions.

#### Abundance analyses of miRNAs in epididymal sperm, seminal plasma and ejaculated sperm

The abundance of miRNAs in epididymal sperm, seminal plasma and ejaculated sperm was evaluated based on transcript per million reads (TPM) and is summarized in Fig. 4 and Table 3. A heat map revealed that most miRNAs were highly expressed in epididymal sperm, followed by ejaculated sperm and seminal plasma (Fig. 4A). Among the 323 miRNAs identified in the three libraries, 221, 259 and 136 miRNAs were expressed differentially in the ejaculated sperm vs epididymal sperm, seminal plasma

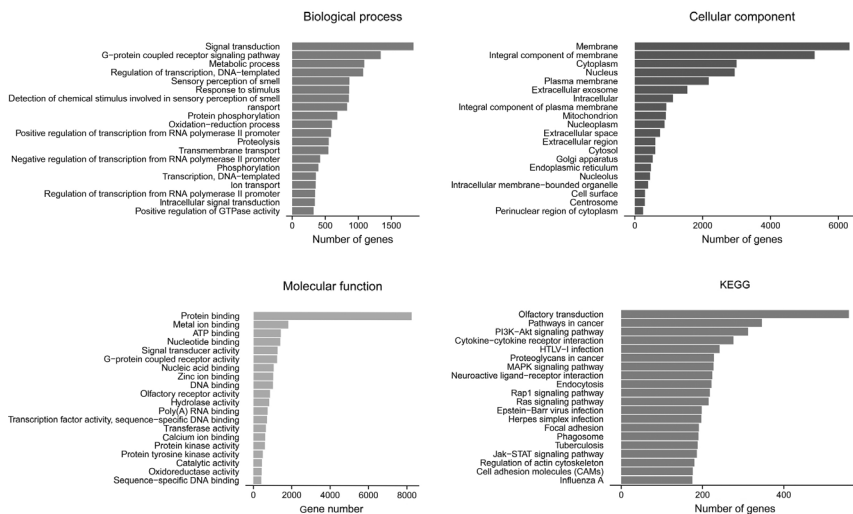
vs epididymal sperm and seminal plasma vs ejaculated sperm libraries respectively. Moreover, compared with epididymal sperm and ejaculated sperm, 254 and 195 miRNAs were decreased and only 5 and 19 miRNAs were increased in seminal plasma. Compared with epididymal sperm, 117 miRNAs were downregulated, 26 were upregulated and 102 were unchanged in ejaculated sperm (Table 3). The top 10 most expressed miRNAs in seminal plasma, epididymal sperm and ejaculated sperm are summarized in the Supplementary Table 4 and accounted for 64.11, 67.56 and 63.78% of the total miRNA reads of seminal plasma, epididymal sperm and ejaculated sperm respectively, resulting in a total of 17 highly expressed miRNAs across the three libraries. These highly expressed miRNAs also displayed a different abundance distribution across the three libraries. These were most abundant in epididymal sperm, followed by ejaculated sperm and seminal plasma (Fig. 4B). Three miRNAs (ssc-let-7a, ssc-miR-26a and ssc-miR-10b) were shared in the top 10 most abundant miRNAs in the three libraries. These 17 highly expressed miRNAs were conserved across mammals: 13 and 5 miRNAs were present in human seminal plasma and ejaculated sperm respectively and 12 and 10 miRNAs were present in bull ejaculated sperm and mouse epididymal sperm respectively (Supplementary Table 4), which suggests a conservation of function.

#### Predicted target analysis of miRNAs

The identification of putative miRNA targets and target enrichment in biological processes and regulatory

**Table 3** Summary of differences in expressed miRNAs between each of the two libraries.

Control	Case	Number of miRNAs			
		Increased	Decreased	No change	Total
Epididymal sperm	Seminal plasma	5	254	62	259
Ejaculated sperm	Seminal plasma	19	117	54	136
Epididymal sperm	Ejaculated sperm	26	195	102	221



**Figure 5** Enrichment analysis of genes targeted by the top 10 most highly expressed miRNAs in the three libraries using gene ontology (GO) and the Kyoto Encyclopaedia of Genes and Genomes (KEGG).

networks are two critical steps that can help elucidate the biological function of newly identified miRNAs. Here, the miRanda software was used to predict the target genes of the top 10 highly expressed miRNAs in the three libraries. The acquired genes were subjected to a GO and KEGG enrichment analysis to determine the functions and pathways that were significantly represented by this group. The top 20 gene clusters involving biological processes, cellular components, molecular functions and pathway analyses predicted by GO and KEGG are summarized in Fig. 5. The most significant biological processes involving the targets of these highly abundant miRNAs were signal transduction and transport pathways, metabolic processes, regulation of transcription, sensory perception of smell and response to stimuli. The most significant gene clusters involving cellular components included the membrane, nucleus, integral component of membranes, extracellular exosome, cytoplasm and nucleoplasm. The most significant molecular function involving the targets of these highly abundant miRNAs were protein binding, followed by metal ion binding, ATP binding, nucleotide binding and signal transducer activity. A pathway analysis performed using KEGG showed that the olfactory transduction, PI3K–Akt, MAPK, Jak–STAT, Rap1 and Ras signaling pathways were most significant. Most of these gene clusters involving biological processes and pathways appeared to be relevant for the sperm metabolism and activity and may be necessary for the spermatogenesis processes, significant enrichment of genes involving olfactory transduction and sensory perception of smell and response to stimulus, which suggests that these miRNAs are more likely to be relevant for the interaction between the environment and animals. In contrast, some of these categories are more likely to be pertinent to the zygote development rather than to the sperm metabolism, such as the gene cluster pathways that were enriched in the regulation of transcription.

## Discussion

The current study provides examples of both similarities and differences in small RNA profiles across the fully matured sperm in the cauda epididymis, ejaculated sperm and seminal plasma of the pig. We compared the abundance and diversity of the small RNAs identified in ejaculated sperm with those found in epididymal sperm and seminal plasma to improve the understanding of their origins and regulatory roles. Regardless of their detection in all libraries studied, globally, the abundance of small RNAs varied significantly across the three libraries, with epididymal sperm being the most abundant, followed by ejaculated sperm and seminal plasma. Most miRNA species were highly expressed in epididymal sperm compared with ejaculated sperm and seminal plasma; and seminal plasma exhibited the lowest abundance. Conversely, the current study revealed that the diversity of small RNAs across the three libraries is generally similar.

In addition to mRNAs, the diversity of the population of the sperm RNAs has now been extended to several non-coding RNAs, including miRNAs, piRNAs (Amanai *et al.* 2006, Krawetz *et al.* 2011, Pantano *et al.* 2015), endo-siRNAs (Yuan *et al.* 2016), small RNAs derived from tRNA, rRNA, snRNA and snoRNA (Peng *et al.* 2012, Schuster *et al.* 2016) and lncRNA (Sendler *et al.* 2013). The sperm miRNA has been reported most extensively and identified in all mammal species examined, including humans, mice, rats, rabbits, bulls, pigs and stallions (Amanai *et al.* 2006, Cornwall 2009, Krawetz *et al.* 2011, Das *et al.* 2013, Stowe *et al.* 2014, Kasimanickam & Kastelic 2016, Schuster *et al.* 2016). Recently, breed-specific differential expression of miRNAs in the pig sperm was detected using the real-time PCR (Kasimanickam & Kastelic 2016). We identified 148 and 321 miRNAs in pig ejaculated sperm and epididymal sperm respectively. We also extended the complexity of small RNA species in sperm



and discovered several other classes of small RNAs, including those derived from rRNA, NRON RNA, cis-reg RNA, tRNA, snRNA and 7SK RNA, which were enriched in pig ejaculated sperm, epididymal sperm and seminal plasma. Small RNAs derived from NRON RNA, cis-reg RNA and 7SK RNA are novel classes of small RNAs that were identified in seminal plasma, ejaculated and epididymal sperm and were reported for the first time. However, we did not detect any piRNA species, which have been extensively reported in human ejaculated sperm and seminal plasma (Krawetz *et al.* 2011, Pantano *et al.* 2015, Hong *et al.* 2016) and in pig testis libraries (Kowalczykiewicz *et al.* 2014, Gebert *et al.* 2015). In addition, the most abundant class of small RNAs in human sperm is piRNAs (Pantano *et al.* 2015), whereas here we found that miRNAs represent the most abundant small RNA species in pig sperm. These data indicate that the small RNA profiles of sperm are significantly different between different mammal species.

It is believed that the sperm RNA is residual to spermatogenesis, and, according to this hypothesis, sperm RNA is a 'by-product' of the gradual shutdown of the nuclear transcriptional apparatus, and sperm does not support any nuclear gene transcription. However, Pittoggi and coworkers found that the transcription and splicing activities are occasionally initiated in mature sperm (Pittoggi *et al.* 2006). Moreover, a very high abundance of sperm adhesins and most CRISP and Fn-type II mRNAs were found in ejaculated sperm, most of which exhibited weak or no expression in the testes (Song *et al.* 2010, 2011). A comparative analysis of the transcripts between bovine spermatids and mature sperm also revealed that although 36.81% were common in both cell types, only 4.5% were specific to mature sperm (Gilbert *et al.* 2007). These data question the hypothesis that the RNA in sperm is leftover from spermiogenesis and suggest that sperm RNAs have other origins. In the current study, we found that a large population of small RNAs, including the enriched small RNA groups, from ejaculated sperm was ejaculated sperm specific, and a small population of small RNAs from ejaculated sperm was shared with those of seminal plasma or epididymal sperm. These data indicate that some sperm-borne small RNAs may originate from seminal plasma and from structures other than the testes.

It has been demonstrated that RNAs are also present in most (if not all) biological fluids, such as plasma, serum, saliva, urine and breast milk (Weber *et al.* 2010). The seminal plasma is a mixture of fluids secreted by the sexual glands (seminal vesicles, the prostate and the bulbourethral glands) and other male sexual organs (testis and epididymis) and provides a nutritive and protective medium for sperm after ejaculation (Juyena & Stelletta 2012). The presence of RNA molecules in human seminal plasma was first identified by Huang and coworkers (Huang *et al.* 2009), and then confirmed

by several studies (Huang *et al.* 2009, Weber *et al.* 2010, Wang *et al.* 2011, Wu *et al.* 2012, Xiong 2014). In seminal plasma, these RNA molecules display remarkable stability and resistance to degradation, possibly because of their association with protein complexes or because they are enclosed within membrane-bound nanovesicles referred to as 'exosomes' (Li *et al.* 2012) that are produced by the diverse sexual glands and organs (Sullivan *et al.* 2005, Piehl *et al.* 2013). Furthermore, exosomes can promote intercellular communication by shuttling proteins and RNAs among cells (Ludwig & Giebel 2012). Both mRNAs and small RNAs in seminal plasma, either existing in the form of exosomes or complexes bound to proteins, can be transferred to the sperm, indicating that, in addition to the testes, seminal plasma may be an alternative origin of sperm RNAs.

Moreover, the epididymis is a curved structure located at the back of the testicles in which sperm matures and is stored. Sperm is formed in the testes and then undergoes maturation processes, including a series of morphological, biochemical and physiological changes, to acquire the function of progressive motility and the ability to fertilize an ovum (Cornwall 2009). During their transit and storage in the epididymis, many epididymal-secreted proteins are transferred to the sperm plasma membrane mediated by exosomes, which are also called epididymosomes and exist within the epididymal lumen secreted by the epididymis (Sullivan *et al.* 2005). Furthermore, epididymosomes can deliver RNAs to immature sperm (Sharma *et al.* 2016). It also has been confirmed that the sperm RNA profile is altered during the epididymal maturation in mice (Nixon *et al.* 2015) and that epididymosomes can directly facilitate such changes (Reilly *et al.* 2016). Therefore, in addition to the testes, the epididymis may be an important origin of the sperm RNA.

However, considering a large population of small RNAs (over 70% of the unique reads of the total annotated small RNAs and 57.1–82.1% of the unique reads in the enriched small RNA classes) from the ejaculated sperm is ejaculated sperm specific indicating that sperm-borne small RNAs may still have other origins in addition to testis, epididymis and seminal plasma. These origins may include the sexual glands (seminal vesicles, the prostate and the bulbourethral glands). These RNAs could be delivered to the sperm mediated with exosomes secreted from these tissues during the ejaculation process, which also deserves further validation. In addition, the diverse origins of sperm RNAs mediated with exosomes may also be one of the explanation of why some small RNA species (such as small RNAs derived from rRNA) in ejaculated sperm represent even higher proportion than that in the epididymal sperm. Two functional roles of sperm RNAs have been suggested. One possibility is that they are required by the sperm itself; another is that the sperm RNA may affect the offspring as a paternal factor by involving in the regulation of early development and

epigenetic inheritance. The functional significance of some miRNAs still seems to be compelling, although it is now commonly accepted that small RNAs play a pivotal role in fertilization and embryonic development (Liu *et al.* 2012a, Chen *et al.* 2016a,b, Yuan *et al.* 2016). An early study suggested that sperm-borne miRNAs play a restricted role in mammalian fertilization (Amanai *et al.* 2006); however, a subsequent study suggested that miRNAs play an important role in fertilization and pre-implantation embryonic development (Yuan *et al.* 2016). In addition, based on the attenuation of the first cleavage division *in vitro* via the injection of a 'miR-34c inhibitor' into zygotes, Liu and coworkers suggested that miR-34c plays an essential role in the first cleavage division in mice (Liu *et al.* 2012b). However, the validity of the function of miR-34c *in vivo*, as assessed using miR-34c-knockout mice, revealed that miR-34c is essential for normal spermatogenesis and male fertility, but its presence in sperm is dispensable for fertilization and pre-implantation development (Yuan *et al.* 2015). In the current study, we identified more than 300 miRNA species, over one-third of which were highly conserved across mammal species. miR-34c, which is also present in humans, mice and bull sperm (Amanai *et al.* 2006, Krawetz *et al.* 2011, Stowe *et al.* 2014), was detected in all the three libraries from pig ejaculated sperm, epididymal sperm and seminal plasma, indicating that this miRNA species is highly conserved across mammals and confirming that it plays a significant biological role in spermatogenesis or in early embryonic development. We also found that three miRNAs (ssc-let-7a, ssc-miR-26a and ssc-miR-10b) were shared in the top 10 most abundant miRNAs in the three libraries, and one of them (miR-let-7a) was detected in mouse epididymal sperm and human seminal plasma (Liu *et al.* 2012a, Wu *et al.* 2012, Cheong *et al.* 2014) and shown to be involved in the regulation of embryo implantation in mice (Liu *et al.* 2012a, Cheong *et al.* 2014). Furthermore, the expression level of let-7a in human seminal plasma was significantly associated with male infertility (Wu *et al.* 2012). Studies have shown that both miR-26a and miR-10b, which are also present in mouse and bull sperms (Amanai *et al.* 2006, Bissonnette *et al.* 2009), appear to be associated with carcinogenesis (Ma *et al.* 2007, Kato *et al.* 2015), but their role in spermatogenesis and in early embryonic development remains unknown. The functional significance of most other miRNAs remains largely unknown. However, based on the prediction of target genes of the top 10 highly expressed miRNAs across the three libraries using the GO and KEGG programs, most of the target gene clusters were enriched in biological processes and pathways that appear to be relevant to the sperm metabolism and activity and the regulation of gene transcription, suggesting that these miRNAs are necessary for spermatogenesis processes and zygote development. Moreover, a significant enrichment of genes involved in the olfactory

transduction, sensory perception of smell and response to stimuli was observed, suggesting that these miRNAs appear to be more relevant in the interaction between the environment and animals.

The functional significance of sperm RNAs has now been extended to the position that some non-coding RNAs act as epigenetic modifiers, perhaps playing a role in transgenerational epigenetic inheritance. Alterations in the expression profiles of sperm tsRNAs, which were found to be significantly enriched in human, mouse, rat and rabbit sperm (Schuster *et al.* 2016), have been observed after paternal exposure to dietary changes, and the modifications of tsRNAs resulted in metabolic disorders in the F1 offspring and altered gene expression of metabolic pathways in early embryos and islets of F1 offspring (Chen *et al.* 2016a), which supports the notion that modifications in the sperm RNA can influence offspring phenotypes via epigenetic inheritance. In the current study, we found that tsRNAs were significantly enriched in pig ejaculated sperm, epididymal sperm and seminal plasma, indicating that the function of sperm tsRNAs as an epigenetic factor that regulates early development and influences offspring phenotypes may be conserved across most mammals.

A cluster of piRNAs was detected in both human ejaculated sperm and seminal plasma, and one study has demonstrated that the expression levels of five piRNAs in human seminal plasma were significantly associated with male infertility (Hong *et al.* 2016), which indicates that sperm-borne piRNAs may also play important and direct roles in spermatogenesis and male infertility. However, the biological significance of sperm-borne piRNAs in early embryonic development remains unknown. Here, as discussed previously, we found that piRNAs are totally absent in pig epididymal sperm, ejaculated sperm and seminal plasma, which indicates that the presence of piRNAs in sperm is not conserved across mammals and that sperm-borne piRNAs play a limited role, if any, in mammalian fertilization or early embryonic development. Although the function of most other classes of small RNAs (NRO RNA, cis-reg RNA and 7SK RNA) enriched in pig sperm remains unclear at the moment, both cis-reg RNA (Xue & Barna 2015) and 7SK RNA (Wassarman & Steitz 1991, Diribarne & Bensaude 2009, Peterlin *et al.* 2012) have been suggested to play roles in the regulation of the transcription and translation of genes. These small RNAs may also play roles in spermatogenesis and in early embryonic development, which deserves further attention.

Overall, the present study revealed the content and profiles of small RNAs in pig seminal plasma, epididymal sperm and ejaculated sperm and highlighted the differences in the abundance and diversity of these small RNAs across the three libraries. This is the first study to confirm the diverse origins of sperm RNAs and report the enrichment of diverse non-miRNA small RNA

classes in pig sperm and seminal plasma. The biogenesis and function of these small RNAs are an interesting area that warrants future investigation.

## Supplementary data

This is linked to the online version of the paper at <http://dx.doi.org/10.1530/REP-17-0014>.

## 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 the Natural Science Foundation of China (31272406 and 31572364), Natural Science Foundation for youths of Jiangsu Province (BK20150440), the Priority Academic Program Development of Jiangsu Higher Education Institutions and the National Development and Reform Commission Special Breeding Projects.

## References

- Amanai M, Brahmajosyula M & Perry AC 2006 A restricted role for sperm-borne microRNAs in mammalian fertilization. *Biology of Reproduction* **75** 877–884. (doi:10.1095/biolreprod.106.056499)
- Bansal SK, Gupta N, Sankhwar SN & Rajender S 2015 Differential genes expression between fertile and infertile spermatozoa revealed by transcriptome analysis. *PLoS ONE* **10** e0127007. (doi:10.1371/journal.pone.0127007)
- Bissonnette N, Lévesque-Sergerie JP, Thibault C & Boissonneault G 2009 Spermatozoal transcriptome profiling for bull sperm motility: a potential tool to evaluate semen quality. *Reproduction* **138** 65–80. (doi:10.1530/REP-08-0503)
- Chen Q, Yan M, Cao Z, Li X, Zhang Y, Shi J, Feng G, Peng H, Zhang X, Zhang Y *et al.* 2016a Sperm tsRNAs contribute to intergenerational inheritance of an acquired metabolic disorder. *Science* **351** 397–400. (doi:10.1126/science.aad7977)
- Chen Q, Yan W & Duan E 2016b Epigenetic inheritance of acquired traits through sperm RNAs and sperm RNA modifications. *Nature Reviews Genetics* **17** 733–743. (doi:10.1038/nrg.2016.106)
- Cheong AWY, Pang RTK, Liu WM, Kottawatta KSA, Lee KF & Yeung WSB 2014 MicroRNA Let-7a and dicer are important in the activation and implantation of delayed implanting mouse embryos. *Human Reproduction* **29** 750–762. (doi:10.1093/humrep/det462)
- Cornwall GA 2009 New insights into epididymal biology and function. *Human Reproduction Update* **15** 213–227. (doi:10.1093/humupd/dmn055)
- Curry E, Safranski TJ & Pratt SL 2011 Differential expression of porcine sperm microRNAs and their association with sperm morphology and motility. *Theriogenology* **76** 1532–1539. (doi:10.1016/j.theriogenology.2011.06.025)
- Das PJ, McCarthy F, Vishnoi M, Paria N, Gresham C, Li G, Kachroo P, Sudderth AK, Teague S, Love CC *et al.* 2013 Stallion sperm transcriptome comprises functionally coherent coding and regulatory RNAs as revealed by microarray analysis and RNA-seq. *PLoS ONE* **8** e56535. (doi:10.1371/journal.pone.0056535)
- Diribarne G & Bensaude O 2009 7SK RNA, a non-coding RNA regulating P-TEFb, a general transcription factor. *RNA Biology* **6** 122–128. (doi:10.4161/rna.6.2.8115)
- Fang P, Zeng P, Wang Z, Liu M, Xu W, Dai J, Zhao X, Zhang D, Liang D, Chen X *et al.* 2014 Estimated diversity of messenger RNAs in each murine spermatozoa and their potential function during early zygotic development. *Biology of Reproduction* **90** 94.
- García-lópez J, Alonso L, Cárdenas DB, Artaza-alvarez H, De J, Hourcade D, Martínez S, Briño-enríquez MA & Mazo JDEL 2015 Diversity and functional convergence of small noncoding RNAs in male germ cell differentiation and fertilization. *RNA* **21** 946–962. (doi:10.1261/rna.048215.114)
- 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)
- Gilbert I, Bissonnette N, Boissonneault G, Vallée M & Robert C 2007 A molecular analysis of the population of mRNA in bovine spermatozoa. *Reproduction* **133** 1073–1086. (doi:10.1530/REP-06-0292)
- Guo X, Su B, Zhou Z & Sha J 2009 Rapid evolution of mammalian X-linked testis microRNAs. *BMC Genomics* **10** 97. (doi:10.1186/1471-2164-10-97)
- Hong Y, Wang C, Fu Z, Liang H, Zhang S, Lu M, Sun W, Ye C, Zhang CY, Zen K *et al.* 2016 Systematic characterization of seminal plasma piRNAs as molecular biomarkers for male infertility. *Scientific Reports* **6** 24229. (doi:10.1038/srep24229)
- Huang S, Li H, Ding X & Xiong C 2009 Presence and characterization of cell-free seminal RNA in healthy individuals: implications for noninvasive disease diagnosis and gene expression studies of the male reproductive system. *Clinical Chemistry* **55** 1967–1976. (doi:10.1373/clinchem.2009.131128)
- Jodar M, Selvaraju S, Sendler E, Diamond MP & Krawetz SA 2013 The presence, role and clinical use of spermatozoal RNAs. *Human Reproduction Update* **19** 604–624. (doi:10.1093/humupd/dmt031)
- Juyena NS & Stelletta C 2012 Seminal plasma: an essential attribute to spermatozoa. *Journal of Andrology* **33** 536–551. (doi:10.2164/jandrol.110.012583)
- Kasimanickam V & Kastelic J 2016 MicroRNA in sperm from Duroc, Landrace and Yorkshire boars. *Scientific Reports* **6** 32954.
- Kato M, Goto Y, Matsushita R, Kurozumi A, Fukumoto I, Nishikawa R, Sakamoto S, Enokida H, Nakagawa M, Ichikawa T *et al.* 2015 MicroRNA-26a/b directly regulate La-related protein 1 and inhibit cancer cell invasion in prostate cancer. *International Journal of Oncology* **47** 710–718.
- Kowalczykiewicz D, Świercz A, Handschuh L, Leśniak K, Figlerowicz M & Wrzesinski J 2014 Characterization of *Sus scrofa* small non-coding RNAs present in both female and male gonads. *PLoS ONE* **9** e113249. (doi:10.1371/journal.pone.0113249)
- Krawetz SA, Kruger A, Lalancette C, Tagett R, Anton E, Draghici S & Diamond MP 2011 A survey of small RNAs in human sperm. *Human Reproduction* **26** 3401–3412. (doi:10.1093/humrep/der329)
- Lalancette C, Thibault C, Bachand I, Caron N & Bissonnette N 2008 Transcriptome analysis of bull semen with extreme nonreturn rate: use of suppression-subtractive hybridization to identify functional markers for fertility. *Biology of Reproduction* **78** 618–635. (doi:10.1095/biolreprod.106.059030)
- Lambard S, Galeraud-Denis I, Martin G, Levy R, Chocat A & Carreau S 2004 Analysis and significance of mRNA in human ejaculated sperm from normozoospermic donors: relationship to sperm motility and capacitation. *Molecular Human Reproduction* **10** 535–541. (doi:10.1093/molehr/gah064)
- Li HC, Ding XF, Liao AH, Kong XB & Xiong CL 2007 Expression of CatSper family transcripts in the mouse testis during post-natal development and human ejaculated spermatozoa: relationship to sperm motility. *Molecular Human Reproduction* **13** 299–306. (doi:10.1093/molehr/gam009)
- Li H, Huang S, Guo C, Guan H & Xiong C 2012 Cell-free seminal mRNA and microRNA exist in different forms. *PLoS ONE* **7** e34566. (doi:10.1371/journal.pone.0034566)
- Liu W-M, Pang RTK, Cheong AWY, Ng EHY, Lao K, Lee K-F & Yeung WSB 2012a Involvement of microRNA lethal-7a in the regulation of embryo implantation in mice. *PLoS ONE* **7** e37039. (doi:10.1371/journal.pone.0037039)
- Liu W-M, Pang RTK, Chiu PCN, Wong BPC, Lao K, Lee K-F & Yeung WSB 2012b Sperm-borne microRNA-34c is required for the first cleavage division in mouse. *PNAS* **109** 490–494. (doi:10.1073/pnas.1110368109)
- Ludwig AK & Giebel B 2012 Exosomes: small vesicles participating in intercellular communication. *International Journal of Biochemistry and Cell Biology* **44** 11–15. (doi:10.1016/j.biocel.2011.10.005)

- Ma L, Teruya-Feldstein J & Weinberg RA 2007 Tumour invasion and metastasis initiated by microRNA-10b in breast cancer. *Nature* **449** 682–688. (doi:10.1038/nature06174)
- Macian F 2005 NFAT proteins: key regulators of T-cell development and function. *Nature Reviews Immunology* **5** 472–484. (doi:10.1038/nri1632)
- Nixon B, Stanger SJ, Mihalas BP, Reilly JN, Anderson AL, Tyagi S, Holt JE & McLaughlin EA 2015 The microRNA signature of mouse spermatozoa is substantially modified during epididymal maturation. *Biology of Reproduction* **93** 91. (doi:10.1095/biolreprod.115.132209)
- Ostermeier GC, Miller D, Huntriss JD, Diamond MP & Krawetz SA 2004 Reproductive biology: delivering spermatozoan RNA to the oocyte. *Nature* **429** 154. (doi:10.1038/429154a)
- Ostermeier GC, Goodrich RJ, Moldenhauer JS, Diamond MP & Krawetz SA 2005 A suite of novel human spermatozoal RNAs. *Journal of Andrology* **26** 70–74.
- Pantano L, Jodar M, Bak M, Ballesca 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)
- Peng H, Shi J, Zhang Y, Zhang H, Liao S, Li W, Lei L, Han C, Ning L, Cao Y *et al.* 2012 A novel class of tRNA-derived small RNAs extremely enriched in mature mouse sperm. *Cell Research* **22** 1609–1612. (doi:10.1038/cr.2012.141)
- Peterlin BM, Brogie JE & Price DH 2012 7SK snRNA: a noncoding RNA that plays a major role in regulating eukaryotic transcription. *Wiley Interdisciplinary Reviews: RNA* **3** 92–103. (doi:10.1002/wrna.106)
- Piehl LL, Fischman ML, Hellman U, Cisale H & Miranda PV 2013 Boar seminal plasma exosomes: effect on sperm function and protein identification by sequencing. *Theriogenology* **79** 1071–1082. (doi:10.1016/j.theriogenology.2013.01.028)
- Pittoggi C, Beraldi R, Sciamanna I, Barberi L, Giordano R, Magnano AR, Torosantucci L, Pescarmona E & Spadafora C 2006 Generation of biologically active retro-genes upon interaction of mouse spermatozoa with exogenous DNA. *Molecular Reproduction and Development* **73** 1239–1246. (doi:10.1002/mrd.20550)
- Reilly J, McLaughlin E, Stanger S, Anderson A, Hutcheon K, Church K, Mihalas B, Tyagi S, Holt J, Eamens A *et al.* 2016 Characterisation of mouse epididymosomes reveals a complex profile of microRNAs and a potential mechanism for modification of the sperm epigenome. *Scientific Reports* **6** 31794. (doi:10.1038/srep31794)
- Schuster A, Tang C, Xie Y, Ortogero N, Yuan S & Yan W 2016 SpermBase: a database for sperm-borne RNA contents. *Biology of Reproduction* **95** 99. (doi:10.1095/biolreprod.116.142190)
- Sandler E, Johnson GD, Mao S, Goodrich RJ, Diamond MP, Hauser R & Krawetz SA 2013 Stability, delivery and functions of human sperm RNAs at fertilization. *Nucleic Acids Research* **41** 4104–4117. (doi:10.1093/nar/gkt132)
- Sharma U, Conine CC, Shea JM, Boskovic A, Derr AG, Bing XY, Belleannee C, Kucukural A, Serra RW, Sun F *et al.* 2016 Biogenesis and function of tRNA fragments during sperm maturation and fertilization in mammals. *Science* **351** 391–396. (doi:10.1126/science.aad6780)
- Song CY, Gao B, Wu H, Wang XY, Chen GH & Mao J 2010 Spatial and temporal expression of spermadhesin genes in reproductive tracts of male and female pigs and ejaculated sperm. *Theriogenology* **73** 551–559. (doi:10.1016/j.theriogenology.2009.09.030)
- Song CY, Gao B, Wu H, Wang XY, Zhou HY, Wang SZ, Li BC, Chen GH & Mao JD 2011 Spatial and temporal gene expression of Fn-type II and cysteine-rich secretory proteins in the reproductive tracts and ejaculated sperm of Chinese Meishan pigs. *Reproduction in Domestic Animals* **46** 848–853. (doi:10.1111/j.1439-0531.2011.01753.x)
- Stoekius M, Grün D & Rajewsky N 2014 Paternal RNA contributions in the *Caenorhabditis elegans* zygote. *EMBO Journal* **33** 1740–1750. (doi:10.15252/embj.201488117)
- Stowe HM, Calcaterra SM, Dimmick MA, Andrae JG, Duckett SK & Pratt SL 2014 The bull sperm microRNAome and the effect of fescue toxicosis on sperm microRNA expression. *PLoS ONE* **9** e113163. (doi:10.1371/journal.pone.0113163)
- Sullivan R, Saez F, Girouard J & Frenette G 2005 Role of exosomes in sperm maturation during the transit along the male reproductive tract. *Blood Cells, Molecules, and Diseases* **35** 1–10. (doi:10.1016/j.bcmd.2005.03.005)
- Wang C, Yang C, Chen X, Yao B, Yang C, Zhu C, Li M, Wang J, Li X, Shao Y *et al.* 2011 Altered profile of seminal plasma microRNAs in the molecular diagnosis of male infertility. *Clinical Chemistry* **57** 1722–1731. (doi:10.1373/clinchem.2011.169714)
- Wassarman DA & Steitz JA 1991 Structural analyses of the 7SK ribonucleoprotein (RNP), the most abundant human small RNP of unknown function. *Molecular and Cellular Biology* **11** 3432–3445. (doi:10.1128/MCB.11.7.3432)
- Weber JA, Baxter DH, Zhang S, Huang DY, Huang KH, Lee MJ, Galas DJ & Wang K 2010 The microRNA spectrum in 12 body fluids. *Clinical Chemistry* **56** 1733–1741. (doi:10.1373/clinchem.2010.147405)
- Willingham AT, Orth AP, Batalov S, Peters EC, Wen BG, Aza-Blanc P, Hogenesch JB & Schultz PG 2005 A strategy for probing the function of noncoding RNAs finds a repressor of NFAT. *Science* **309** 1570–1573. (doi:10.1126/science.1115901)
- Wu W, Hu Z, Qin Y, Dong J, Dai J, Lu C, Zhang W, Shen H, Xia Y & Wang X 2012 Seminal plasma microRNAs: potential biomarkers for spermatogenesis status. *Molecular Human Reproduction* **18** 489–497. (doi:10.1093/molehr/gas022)
- Xiong C 2014 Identification of microRNAs predominately derived from testis and epididymis in human seminal plasma. *Clinical Biochemistry* **47** 967–972. (doi:10.1016/j.clinbiochem.2013.11.009)
- Xue S & Barna M 2015 Cis-regulatory RNA elements that regulate specialized ribosome activity. *RNA Biology* **12** 1083–1087. (doi:10.1080/15476286.2015.1085149)
- Yang CC, Lin YS, Hsu CC, Wu SC, Lin EC & Cheng WTK 2009 Identification and sequencing of remnant messenger RNAs found in domestic swine (*Sus scrofa*) fresh ejaculated spermatozoa. *Animal Reproduction Science* **113** 143–155. (doi:10.1016/j.anireprosci.2008.08.012)
- Yang Q, Lin J, Liu M, Li R, Tian B, Zhang X, Xu B, Liu M, Zhang X, Li Y *et al.* 2016 Highly sensitive sequencing reveals dynamic modifications and activities of small RNAs in mouse oocytes and early embryos. *Science Advances* **2** e1501482. (doi:10.1126/sciadv.1501482)
- Yuan S, Tang C, Zhang Y, Wu J, Bao J, Zheng H, Xu C & Yan W 2015 Mir-34b/c and Mir-449a/b/c are required for spermatogenesis, but not for the first cleavage division in mice. *Biology Open* **4** 212–223. (doi:10.1242/bio.201410959)
- Yuan S, Schuster A, Tang C, Yu T, Ortogero N, Bao J, Zheng H & Yan W 2016 Sperm-borne miRNAs and endo-siRNAs are important for fertilization and preimplantation embryonic development. *Development* **143** 635–647. (doi:10.1242/dev.131755)

---

Received 7 January 2017

First decision 30 January 2017

Revised manuscript received 16 March 2017

Accepted 17 March 2017