Title

A novel class of tRNA-derived small RNAs extremely enriched in mature mouse sperm

Journal

Cell Research, 22(11)

ISSN

1001-0602

Authors

Peng, Hongying
Shi, Junchao
Zhang, Ying
et al.

Publication Date

2012-11-01

DOI

10.1038/cr.2012.141

Peer reviewed

Open
LETTER TO THE EDITOR

Cell Research (2012) :1-4.

www.nature.com/cr

A novel class of tRNA-derived small RNAs extremely enriched in mature mouse sperm

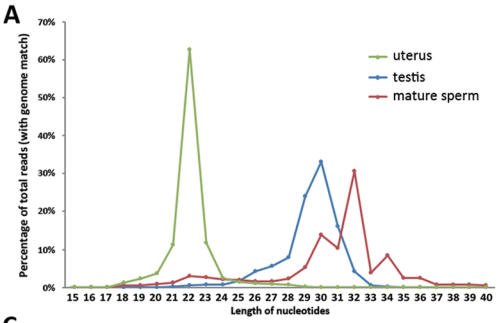
Cell Research advance online publication 9 October 2012; doi:10.1038/cr.2012.141

Dear editor,

The discovery of sperm-borne RNAs (mRNAs and small non-coding RNAs) has opened the possibility of additional paternal contributions aside from providing the DNA [1]. It has been reported that the incoming sperm can provide information for its host egg cytoplasm, which functionally influences the order of cell division [2], possibly via delivering RNAs. Indeed, the sperm-borne miRNA and mRNA have been demonstrated as active players in early embryo development [3] and transgenerational epigenetic inheritance [4]. However, given the diversity of small RNA classes (miRNA, endo-siRNA, piRNA, etc.) generated during spermatogenesis, the contents and profiles of the small RNA population carried by mature sperm remain undefined. In the present study, we isolated mature sperm from the cauda epididymis of adult male mice (Supplementary information, Data S1). The purity of sperm was > 99% as evaluated by microscopy and was confirmed by RT-PCR analyses of different biomarkers (Supplementary information, Figure S1A and S1B). The RNA extracted from mature sperm, adult testis, and uterus were processed for small RNA (< 40 nt) deep sequencing (Supplementary information, Figure S1C, S1D, S1E and Data S1). The total small RNA reads and genome-mapping statistic data (Supplementary information, Table S1) showed an abundance of small RNAs carried by mature sperm. The overall length distribution of small RNAs (Figure 1A) revealed that the dominant reads from mature sperm were at 29-34 nt, slightly different from adult testis (26-32 nt), and distinct from uterus (21-23 nt). The majority of the 26-32 nt small RNAs in mouse testis are piwi-interacting RNAs (piRNAs), which are actively involved in retrotransposon silencing that protects the integrity of the genome [5]. As it was initially suggested that piRNAs are absent in the cauda epididymis [6] and that mammalian PIWI proteins (MILI, MIWI, MIWI2) are not expressed in mature sperm [5], the abundant existence of 29-34 nt small RNAs in the mature sperm is somewhat surprising to us and suggests that they might be different from the well-known piRNA population from testis. Further analysis has revealed a distinct signature for these mature-sperm-enriched small RNAs, which represent a novel class of abundantly expressed small RNAs that can be grouped into distinct families. The small RNAs within each family showed identical 5′ sequences and only differed at their distal 3′ ends (Figure 1B and Supplementary information, Figure S2), suggesting that they are derived from the same precursor sequence. Particularly, two of these small RNA families were extremely enriched, which comprised 38.19% (family-1) and 19.14% (family-2) of all small RNA reads, respectively (Figure 1B), and together they accounted for the majority of the 30-34 nt small RNA population (Figure 1C).

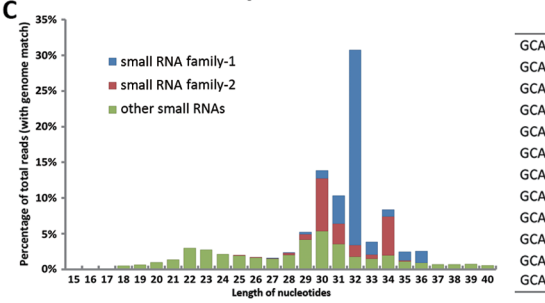
To further characterize these mature-sperm-enriched small RNAs, we performed BLAT searches for the top two abundant families against the mouse genomic databases (mm9). As shown in Figure 1D, these small RNAs are located at multiple sites on the genome, with several clusters on chromosomes 1, 8 and 13. Most strikingly, each of these genomic locations corresponds to a tRNA locus (Figure 1D). By further comparing with the genomic tRNA database, we found that each of these small RNA families unanimously matches to the 5′ half of a specific tRNA, with cleavage sites located preferentially at the anticodon loop (30-34 nt from the 5′ end), as illustrated for families 1-2 (Figure 1E and 1F) and for families 3-7 (Supplementary information, Figure S2). Their ultra-high enrichment (Figure 1G) and the preferential cleavage sites and length distributions strongly suggested that these small RNAs are not generated randomly by tRNA degradation, but under strict cleavage regulations. As these small RNAs are highly enriched in mature sperm and are derived from tRNAs, we termed them “mature-sperm-enriched tRNA-derived small RNAs” (mse-tsRNAs).

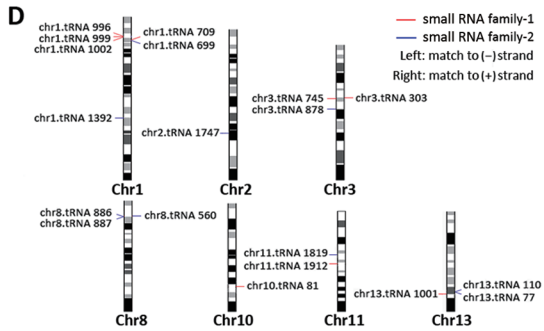
Indeed, recent evidence has demonstrated that tRNA-

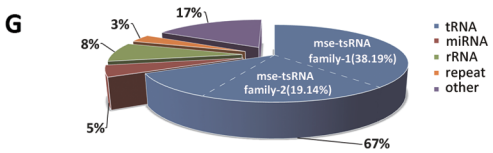


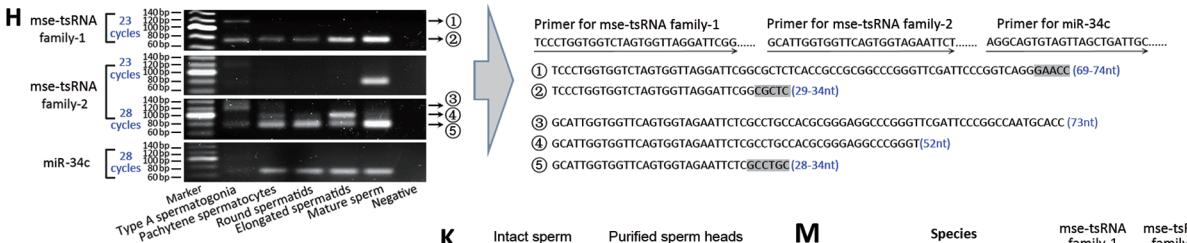
B

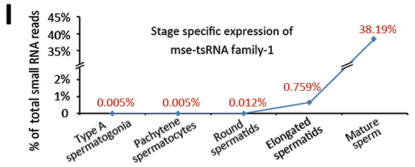
Sequence(5′-3′) small RNA family-1	Length (bp)	Reads(%) Sperm	Testis	Uterus
TCCCTGGTGGTCTAGTGGTTAGGATTC	27	197(0.00%)	26(0.00%)	18(0.00%)
TCCCTGGTGGTCTAGTGGTTAGGATTCG	28	823(0.01%)	51(0.00%)	26(0.00%)
TCCCTGGTGGTCTAGTGGTTAGGATTCGG	29	25806(0.30%)	1741(0.01%)	51(0.00%)
TCCCTGGTGGTCTAGTGGTTAGGATTCGGC	30	94632(1.09%)	105764(0.60%)	5(0.00%)
TCCCTGGTGGTCTAGTGGTTAGGATTCGGCG	31	342974(3.93%)	5677(0.03%)	4(0.00%)
TCCCTGGTGGTCTAGTGGTTAGGATTCGGCGC	32	2383402(27.33%)	2312(0.01%)	3(0.00%)
TCCCTGGTGGTCTAGTGGTTAGGATTCGGCGCT	33	152843(1.75%)	246(0.00%)	1(0.00%)
TCCCTGGTGGTCTAGTGGTTAGGATTCGGCGCTC	34	85609(0.98%)	196(0.00%)	0(0.00%)
TCCCTGGTGGTCTAGTGGTTAGGATTCGGCGCTCT	35	100978(1.16%)	27(0.00%)	0(0.00%)
TCCCTGGTGGTCTAGTGGTTAGGATTCGGCGCTCTC	36	142427(1.63%)	11(0.00%)	0(0.00%)
TCCCTGGTGGTCTAGTGGTTAGGATTCGGCGCTCTCA	37	350(0.00%)	0(0.00%)	0(0.00%)
TCCCTGGTGGTCTAGTGGTTAGGATTCGGCGCTCTCAC	38	249(0.00%)	0(0.00%)	0(0.00%)
		3330290(38.19%)	116051(0.66%)	108(0.00%)
small RNA family-2				
GCATTGGTGGTTCAGTGGTAGAATT	25	240(0.00%)	1(0.00%)	13(0.00%)
GCATTGGTGGTTCAGTGGTAGAATTC	26	464(0.01%)	54(0.00%)	42(0.00%)
GCATTGGTGGTTCAGTGGTAGAATTCT	27	2733(0.03%)	60.00%)	7(0.00%)
GCATTGGTGGTTCAGTGGTAGAATTCTC	28	25877(0.30%)	16(0.00%)	2(0.00%)
GCATTGGTGGTTCAGTGGTAGAATTCTCG	29	67765(0.78%)	11(0.00%)	0(0.00%)
GCATTGGTGGTTCAGTGGTAGAATTCTCGC	30	643057(7.37%)	15(0.00%)	0(0.00%)
GCATTGGTGGTTCAGTGGTAGAATTCTCGCC	31	250784(2.88%)	9(0.00%)	0(0.00%)
GCATTGGTGGTTCAGTGGTAGAATTCTCGCCT	32	138846(1.59%)	4(0.00%)	1(0.00%)
GCATTGGTGGTTCAGTGGTAGAATTCTCGCCTG	33	51980(0.60%)	11(0.00%)	0(0.00%)
GCATTGGTGGTTCAGTGGTAGAATTCTCGCCTGC	34	472739(5.42%)	104(0.00%)	0(0.00%)
GCATTGGTGGTTCAGTGGTAGAATTCTCGCCTGCC	35	13972(0.16%)	38(0.00%)	0(0.00%)
GCATTGGTGGTTCAGTGGTAGAATTCTCGCCTGCCA	36	320(0.00%)	0(0.00%)	0(0.00%)
		1668777(19.14%)	269(0.00%)	65(0.00%)

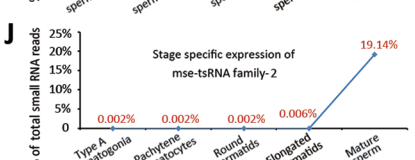


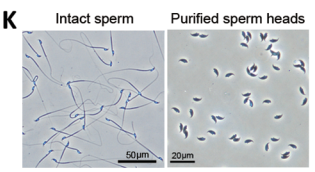


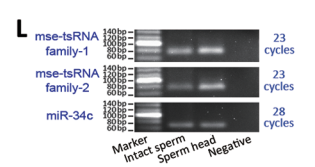


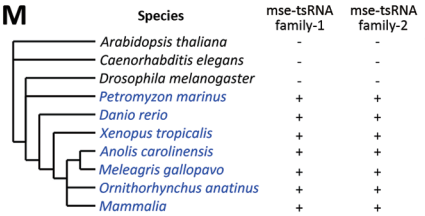


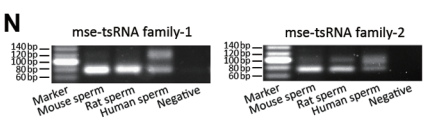












derived RNA fragments are biologically functional [7], and their production could be induced under various stress conditions (physical or chemical stress) by specific RNase [8]. The mature sperm is produced from testicular spermatogenesis followed by maturation during passage through the epididymis. The physiological condition and specific enzyme governing mse-tsRNAs production and accumulation in mature sperm are currently unknown.

To monitor the biogenesis of mse-tsRNAs during sperm formation, we next analyzed purified mouse testicular spermatogenic cells (type A spermatogonia, pachytene spermatocytes, round spermatids and elongated spermatids) and mature sperm. We found that the levels of mse-tsRNA family-1 and -2 were relatively constant during the early stages of spermatogenesis, and were substantially increased at late- (family-1) or post-spermatogenesis (family-2) (Figure 1H). The RT-PCR results were further confirmed by analyzing small RNA deep-sequencing data obtained from purified mouse testicular spermatogenic cells [9] and mature sperm (Figure 1I and 1J). By analyzing the PCR product size followed by product sequencing, we could find both mse-tsRNAs and their tRNA precursors, supporting the hypothesis that the mse-tsRNAs are derived from tRNA cleavage (Figure 1H). Interestingly, we consistently observed a sequence (52 nt) in mse-tsRNA family-2 PCR products, which is shorter than the expected length of intact tRNA and mapped to its 5′ portion (Figure 1H), suggesting the involvement of a two-step cleavage of tRNA in generating mse-tsRNA family-2. The increase of mse-tsRNA family-1 and -2 seems not to correlate with the expression of mammalian PIWI proteins (MILI, MIWI, MIWI2) [5], suggesting that they might not be closely related to the piRNAs, and their ultra-high enrichment in mature sperm might be due to specific tRNA cleavage and/or selective accumulation of cleavage products at late- or post-spermatogenesis (such as during epididymal transition). The underlying mechanisms are currently unknown.

We also analyzed the relative expression of miR-34c by RT-PCR as a quality control, as its expression in spermatogenic cells and mature sperm have been previously reported [3, 10]. As shown in Figure 1H, our results were consistent with previous reports that miR-34c was almost absent in spermatogonia, but was highly expressed from pachytene spermatocytes and continued to be highly expressed in spermatids [10] and mature sperm [3]. Note that miR-34c expression in mature sperm is much less than that of mse-tsRNA family-1 and -2, as shown by the RT-PCR results and by the reads number from our miRNA profiling database (Supplementary information, Figure S3).

It is important to determine whether the sperm-borne mse-tsRNAs are located in sperm head, which could indicate their potential delivery into oocytes at fertilization. Using established methods to isolate purified sperm heads, we demonstrated that mse-tsRNAs are abundantly localized in the purified sperm head (Figure 1K and 1L), suggesting that they could be delivered into oocytes at fertilization. We next analyzed the expression levels of mse-tsRNA family-1 in oocytes, zygotes and parthenogenetically activated oocytes. Surprisingly, quantitative RT-PCR analysis revealed that the level of mse-tsRNA family-1 in zygotes is significantly lower than that in the oocytes and parthenogenetically activated oocytes (Supplementary information, Figure S4). These results might suggest a fertilization-triggered usage/consumption of mse-tsRNAs, which may reflect a functional role for mse-tsRNAs in early embryo events.

As the mse-tsRNAs are derived from their tRNA templates, their sequence conservation could simply reflect the evolutionary conservation of their tRNA precursors. As shown in the evolutionary conservation analysis (Figure 1M and Supplementary information, Table S2), the tRNAs generating mse-tsRNA families 1-7 are highly

Figure 1 Identification and characterization of tRNA-derived small RNAs in mature mouse sperm. **(A)** Length distributions of small RNAs in uterus, testis and mature sperm. **(B)** Alignments and statistics of small RNA family-1 and -2 in mature sperm, testis and uterus; the top three enriched RNAs for each family were highlighted. **(C)** Small RNA family-1 and -2 comprised the majority of 30-34 nt small RNA population. **(D)** Chromosome locations of small RNA family-1 and -2. **(E**, **F)** Illustrations showing that mse-tsRNA family-1 **(E)** and mse-tsRNA family-2 **(F)** are derived from 5′ halves of $tRNA^{Glu}$ and $tRNA^{Gly}$ respectively; top three cleavage sites for each tRNA were marked by arrow heads. **(G)** Catalogue of small RNA populations in mature mouse sperm. **(H)** RT-PCR analyses of mse-tsRNA family-1, -2 and miR-34c in purified spermatogenic cells and mature sperm, followed by product sequencing. The variable nucleotides were marked by shade. Samples for RT-PCR were equally loaded by calibration of RNA concentration. Similar results were obtained in 3 independent experiments. **(I**, **J)** Percentages of mse-tsRNA family-1 **(I)** and -2 **(J)** reads from deep-sequencing datasets of purified spermatogenic cells and mature sperm. **(K)** Representative photos of intact sperm and purified sperm heads. **(L)** RT-PCR analyses of mse-tsRNA family-1, -2 and miR-34c in intact sperm and purified sperm heads. **(M)** Evolutionary tree showing that the tRNA precursors of mse-tsRNA family-1 and -2 are highly conserved in vertebrate species. **(N)** RT-PCR analyses of mse-tsRNA family-1 and -2 in mature sperm from mouse, rat and human.

conserved in vertebrate species, from fish to mammals, but absent from the worms, flies and plants. Indeed, besides our reported data for mouse sperm, existing small RNA datasets have shown that the mse-tsRNA family-1 is among the most highly expressed small RNA sequences in zebra fish testis (NCBI GEO Datasets: GSM830247) [11] and human sperm (NCBI GEO Datasets: GSM530235) [12], supporting the spermatozoal expression of mse-tsRNAs in a wide range of species. The expression of mse-tsRNA family-1 and -2 in mature sperm from mouse, rat and human were further analyzed using RT-PCR and confirmed by product sequencing (Figure 1N). These results suggest that mse-tsRNAs might serve as an ancient paternal element with evolutionarily conserved functions.

Taken together, the present study revealed a previously hidden layer of sperm-borne small RNAs, identifying a novel class of tRNA-derived mse-tsRNAs with ultra-high accumulation in mature sperm. The biogenesis and function of these mse-tsRNAs are interesting topics that warrant future investigations.

Acknowledgments

This research was supported by National Basic Research Program of China (2011CB944401 and 2011CB710905), Strategic Priority Research Program of the Chinese Academy of Sciences (XDA 01010202), National Natural Science Foundation of China (31200879), and Sciences Knowledge Innovation Program of the Chinese Academy of Sciences (KSCX2-EW-R-06).

Hongying Peng[1, *], Junchao Shi[1, *], Ying Zhang[1], He Zhang[1], Shangying Liao[1], Wei Li[1, 2], Li Lei[1, 2], Chunsheng Han[1], Lina Ning[1], Yujing Cao[1], Qi Zhou[1], Qi Chen[1], Enkui Duan[1]

[1]*State Key Laboratory of Reproductive Biology, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China;* [2]*Graduate Universty of Chinese Academy of Sciences, Beijing 100049, China*

*These two authors contributed equally to this work.
Correspondence: Enkui Duan[a], Qi Chen[b]
[a]Tel: +86-10-64807182;
E-mail: duane@ioz.ac.cn

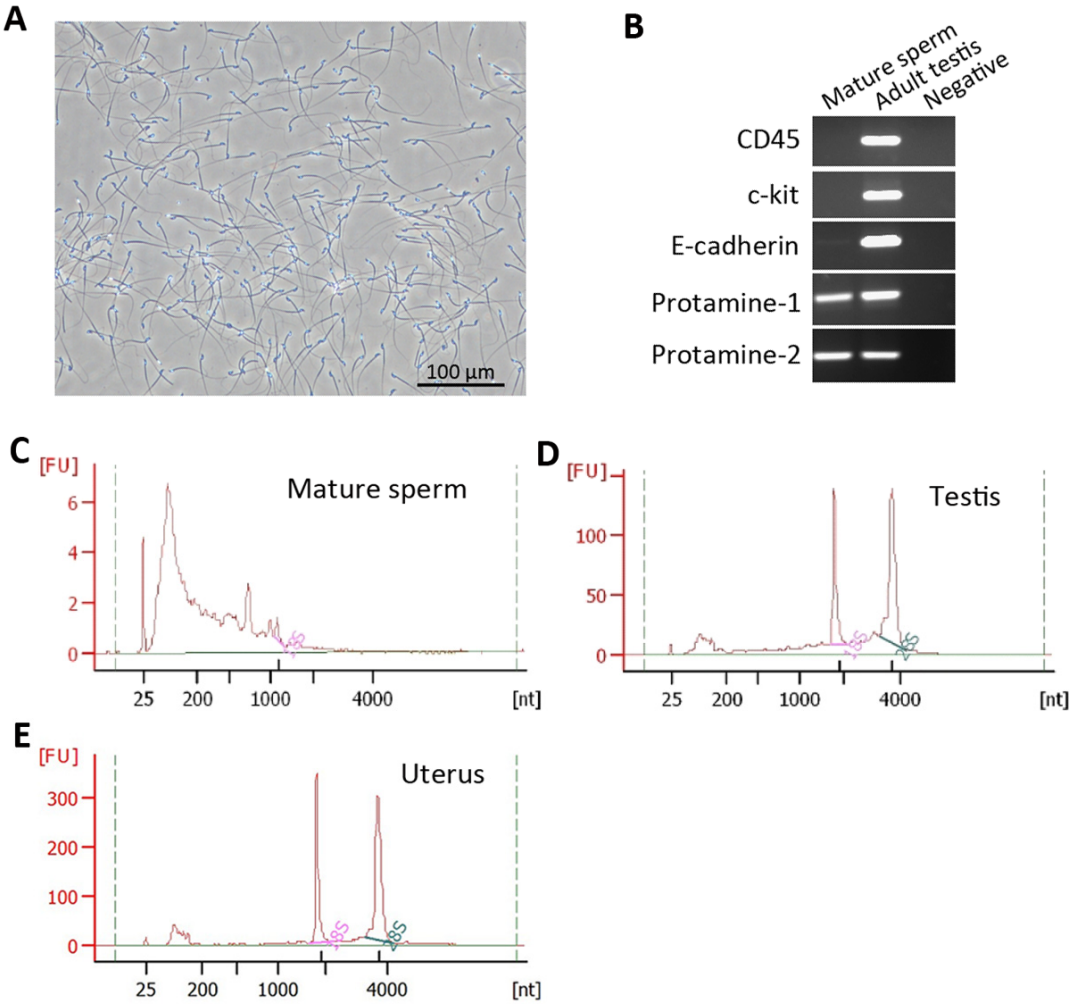
[b]Tel: +86-10-64807182
E-mail: chenqi@ioz.ac.cn

References

1 Krawetz SA. Paternal contribution: new insights and future challenges. *Nat Rev Genet* 2005; **6**:633-642.
2 Piotrowska K, Zernicka-Goetz M. Role for sperm in spatial patterning of the early mouse embryo. *Nature* 2001; **409**:517-521
3 Liu WM, Pang RT, Chiu PC, *et al.* Sperm-borne microRNA-34c is required for the first cleavage division in mouse. *Proc Natl Acad Sci USA* 2012; **109**:490-494.
4 Rassoulzadegan M, Grandjean V, Gounon P, *et al*. RNA-mediated non-mendelian inheritance of an epigenetic change in the mouse. *Nature* 2006; **441**:469-474.
5 Thomson T, Lin H. The biogenesis and function of PIWI proteins and piRNAs: progress and prospect. *Annu Rev Cell Dev Biol* 2009; **25**:355-376. .
6 Grivna ST, Beyret E, Wang Z, Lin H. A novel class of small RNAs in mouse spermatogenic cells. *Genes Dev* 2006; **20**:1709-1714.
7 Lee YS, Shibata Y, Malhotra A, Dutta A. A novel class of small RNAs: tRNA-derived RNA fragments (tRFs). *Genes Dev* 2009; **23**:2639-2649.
8 Yamasaki S, Ivanov P, Hu GF, Anderson P. Angiogenin cleaves tRNA and promotes stress-induced translational repression. *J Cell Biol* 2009; **185**:35-42.
9 Gan H, Lin X, Zhang Z, *et al.* piRNA profiling during specific stages of mouse spermatogenesis. *RNA* 2011; **17**:1191-1203.
10 Liang X, Zhou D, Wei C, *et al.* MicroRNA-34c enhances murine male germ cell apoptosis through targeting ATF1. *PLoS One* 2012; **7**:e33861.
11 Kamminga LM, Luteijn MJ, den Broeder MJ, *et al*. Hen1 is required for oocyte development and piRNA stability in zebrafish. *Embo J* 2010; **29**:3688-3700.
12 Krawetz SA, Kruger A, Lalancette C, *et al*. A survey of small RNAs in human sperm. *Hum Reprod* 2011; **26**:3401-3412.

(**Supplementary information** is linked to the online version of the paper on the *Cell Research* website.)

Supplementary Figure S1-S4 and Table S1-2



Supplementary figure S1 Confirmation of sperm cell purity and electrophoretic size distribution of extracted RNAs. **(A)** Morphology of collected sperm under phase contrast microscopy. **(B)** Biomarkers of leukocytes (CD45), testicular germ cells (c-kit) and epithelial cells (E-cadherin) were unable to be amplified from the RNA extracted from our sperm sample, while positive markers of sperm (protamine-1 and-2) were easily detected, indicating the purity of our sperm sample. **(C-E)** Electrophoretic size distribution of RNAs in mouse mature sperm **(C)**, testis **(D)** and uterus **(E)** analyzed by Agilent Bioanalyzer. Note that sperm RNAs are enriched in smaller size fractions and lack the 28s, 18s peaks that are observed in other tissue types, which is similar to previous reports for mature human sperm [1].

1. Johnson GD, Sendler E, Lalancette C, Hauser R, Diamond MP, Krawetz SA. Cleavage of rRNA ensures translational cessation in sperm at fertilization. *Mol Hum Reprod*. **17** (12):721-726.

Supplementary Table S1

Statistics of small RNA reads and genome mapping data for Mature sperm, Adult testis and Uterus

Libraries	Mature sperm		Adult testis		Uterus	
	Reads	Sequences	Reads	Sequences	Reads	Sequences
After adapters being removed	15690722	1302963	23783785	2512006	15876066	526203
Mapping to genome	8720312	631580	16493791	1286548	12466462	196535

A

Sequence(5'-3') mse-tsRNA family-3	length(bp)	reads		
		sperm	testis	uterus
GCATGGGTGGTTCAGTGGTAGAATTCT	27	447	2	0
GCATGGGTGGTTCAGTGGTAGAATTCTC	28	1926	1	0
GCATGGGTGGTTCAGTGGTAGAATTCTCG	29	4829	0	0
GCATGGGTGGTTCAGTGGTAGAATTCTCGC	30	31269	1	0
GCATGGGTGGTTCAGTGGTAGAATTCTCGCC	31	13135	0	0
GCATGGGTGGTTCAGTGGTAGAATTCTCGCCT	32	11204	1	0
GCATGGGTGGTTCAGTGGTAGAATTCTCGCCTG	33	966	1	0
GCATGGGTGGTTCAGTGGTAGAATTCTCGCCTGC	34	15821	7	0
GCATGGGTGGTTCAGTGGTAGAATTCTCGCCTGCC	35	389	1	0
		79986	14	0

B

mse-tsRNA family-4	length(bp)	reads		
		sperm	testis	uterus
TCCCACATGGTCTAGCGGTTAGGATT	26	1166	14	35
TCCCACATGGTCTAGCGGTTAGGATTC	27	186	102	11
TCCCACATGGTCTAGCGGTTAGGATTCC	28	344	94	7
TCCCACATGGTCTAGCGGTTAGGATTCCT	29	215	278	3
TCCCACATGGTCTAGCGGTTAGGATTCCTG	30	2821	1389	0
TCCCACATGGTCTAGCGGTTAGGATTCCTGG	31	30068	358	0
TCCCACATGGTCTAGCGGTTAGGATTCCTGGT	32	7447	54	0
TCCCACATGGTCTAGCGGTTAGGATTCCTGGTT	33	6472	357	0
TCCCACATGGTCTAGCGGTTAGGATTCCTGGTTT	34	10270	12	0
TCCCACATGGTCTAGCGGTTAGGATTCCTGGTTTT	35	5106	5	0
TCCCACATGGTCTAGCGGTTAGGATTCCTGGTTTTC	36	13301	1	0
		77396	2664	56

C

mse-tsRNA family-5	length(bp)	reads		
		sperm	testis	uterus
GTTTCCGTAGTGTAGTGGTTATCAC	25	180	3	1
GTTTCCGTAGTGTAGTGGTTATCACG	26	1647	3	2
GTTTCCGTAGTGTAGTGGTTATCACGT	27	585	13	6
GTTTCCGTAGTGTAGTGGTTATCACGTT	28	484	4	2
GTTTCCGTAGTGTAGTGGTTATCACGTTC	29	361	9	2
GTTTCCGTAGTGTAGTGGTTATCACGTTCG	30	795	5	0
GTTTCCGTAGTGTAGTGGTTATCACGTTCGC	31	6075	9	0
GTTTCCGTAGTGTAGTGGTTATCACGTTCGCC	32	5866	28	0
GTTTCCGTAGTGTAGTGGTTATCACGTTCGCCT	33	26772	40	0
GTTTCCGTAGTGTAGTGGTTATCACGTTCGCCTC	34	28030	29	0
GTTTCCGTAGTGTAGTGGTTATCACGTTCGCCTCA	35	1499	1	0
GTTTCCGTAGTGTAGTGGTTATCACGTTCGCCTCAC	36	1635	0	0
		73929	144	13

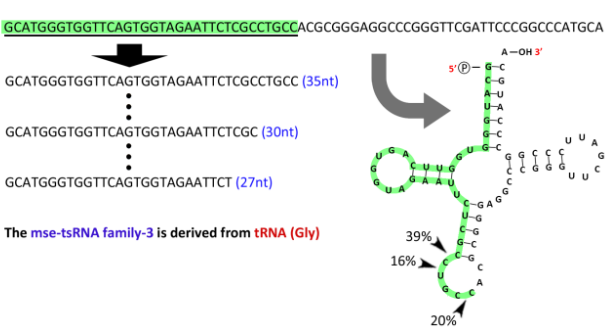
D

mse-tsRNA family-6	length(bp)	reads		
		sperm	testis	uterus
TCCCATATGGTCTAGCGGTTAGGATT	26	676	17	9
TCCCATATGGTCTAGCGGTTAGGATTC	27	92	114	6
TCCCATATGGTCTAGCGGTTAGGATTCC	28	190	109	5
TCCCATATGGTCTAGCGGTTAGGATTCCT	29	111	441	0
TCCCATATGGTCTAGCGGTTAGGATTCCTG	30	1144	1954	0
TCCCATATGGTCTAGCGGTTAGGATTCCTGG	31	12381	666	0
TCCCATATGGTCTAGCGGTTAGGATTCCTGGT	32	3099	69	0
TCCCATATGGTCTAGCGGTTAGGATTCCTGGTT	33	2988	508	0
TCCCATATGGTCTAGCGGTTAGGATTCCTGGTTT	34	4687	5	0
TCCCATATGGTCTAGCGGTTAGGATTCCTGGTTTT	35	2485	4	0
TCCCATATGGTCTAGCGGTTAGGATTCCTGGTTTTC	36	5827	1	0
		33680	3888	20

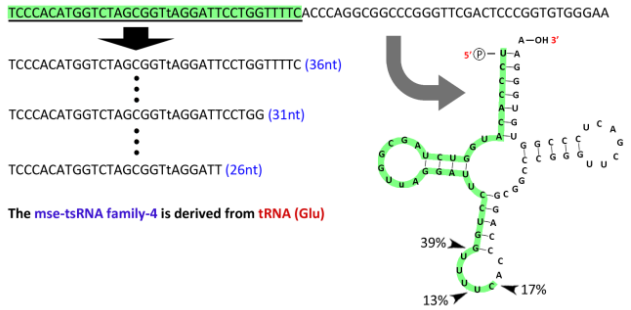
E

mse-tsRNA family-7	length(bp)	reads		
		sperm	testis	uterus
GGGGGTATAGCTCAGTGGTAGAGCA	25	514	0	5
GGGGGTATAGCTCAGTGGTAGAGCAT	26	37	0	2
GGGGGTATAGCTCAGTGGTAGAGCATT	27	22	0	8
GGGGGTATAGCTCAGTGGTAGAGCATTT	28	18	0	0
GGGGGTATAGCTCAGTGGTAGAGCATTTG	29	113	0	0
GGGGGTATAGCTCAGTGGTAGAGCATTTGA	30	385	11	1
GGGGGTATAGCTCAGTGGTAGAGCATTTGAC	31	85	10	0
GGGGGTATAGCTCAGTGGTAGAGCATTTGACT	32	5401	18	0
GGGGGTATAGCTCAGTGGTAGAGCATTTGACTG	33	2557	18	0
GGGGGTATAGCTCAGTGGTAGAGCATTTGACTGC	34	21985	30	0
GGGGGTATAGCTCAGTGGTAGAGCATTTGACTGCA	35	109	9	0
		31226	96	16

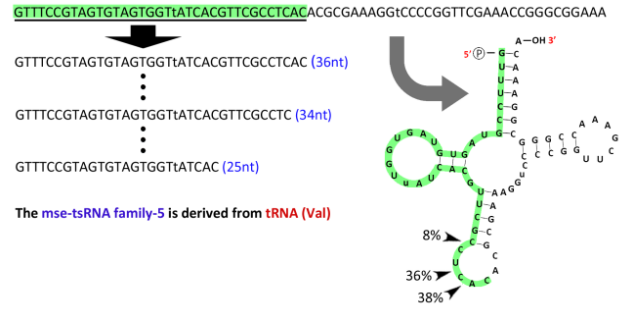
F



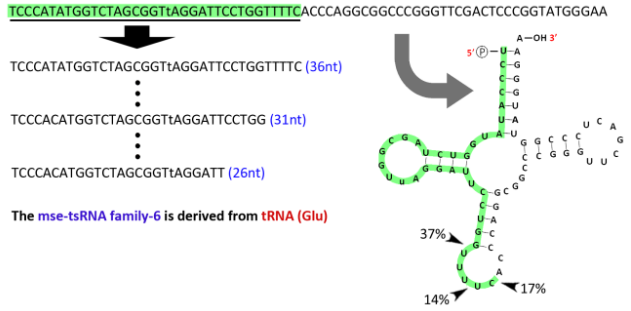
G



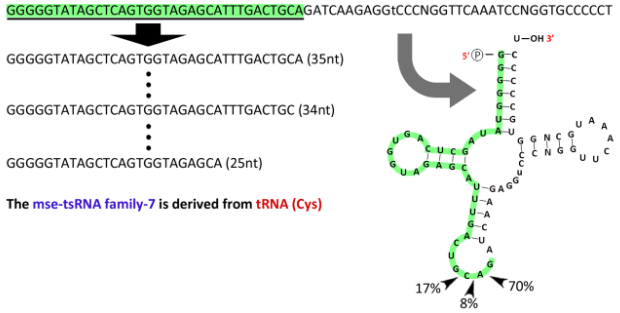
H



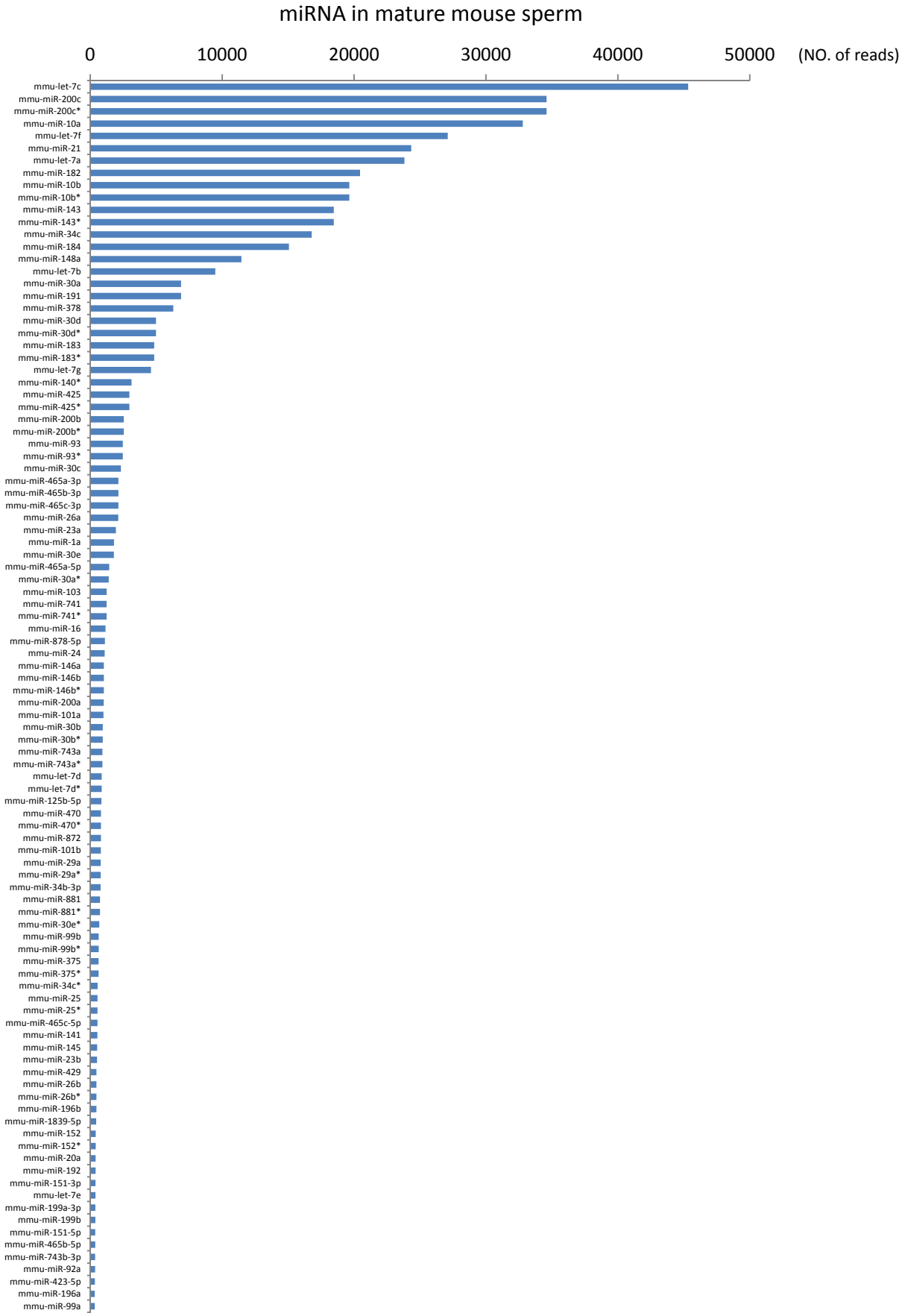
I



J



Supplementary Figure S2 Alignments, statistics and tRNA origin of mse-tsRNA family 3-7 in mature sperm, testis and uterus. (**A-E**) Alignments and statistics of mse-tsRNA family 3-7 in mature sperm, testis and uterus. The top three enriched length for each family were highlighted. (**F-J**) Illustrations showing tRNA origin of mse-tsRNA family 3-7, all these mse-tsRNA families are derived from 5' halves of tRNA. The top three cleavage sites were marked by arrow heads. The preferenced length distribution at 30-34nt and cleavage sites suggested specific nuclease involved in the cleavage, possibly by recognizing the anticodon loops.



Supplementary Figure S3 miRNA profiles in mature mouse sperm. The top 100 expressed miRNAs are listed, which take up 97.24% of all miRNA reads in mature mouse sperm.

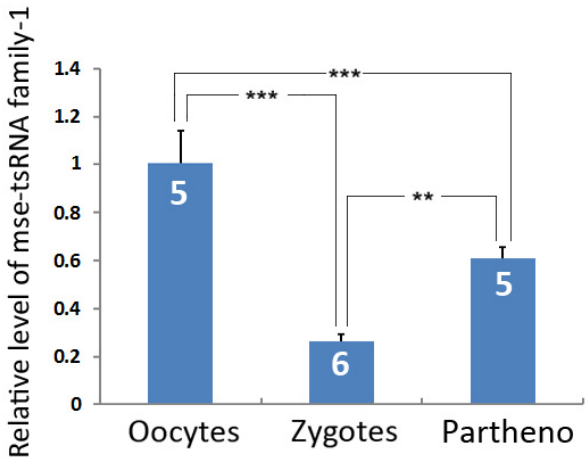
A

Quantitative Analysis of mse-tsRNA family-1 by Real Time RT-PCR

Samples	Ct value	Relative amount
Oocytes	22.77±0.22	1.00
Zygotes	24.69±0.16	0.26
Partheno	23.46±0.11	0.61

Data of relative amount are normalized to the Ct of oocytes and calculated by the $2^{-\Delta C_t}$ method

B



Supplementary Figure S4 Quantitative Analysis of mse-tsRNA family-1 in oocytes, zygotes and parthenogenetically activated oocytes. For the quantitative examination, 100 embryos were pooled for RNA extraction as an independent sample. We collected 5 samples for oocytes (100 x 5), 6 samples for zygotes (100 x 6) and 5 samples for parthenogenetically activated oocytes (100 x 5). Then the RNAs were extracted for each sample followed by quantitative RT-PCR for mse-tsRNA family-1. **(A, B)** Ct value and relative expression level of mse-tsRNA family-1 (***p<0.001, **p<0.01, Error bar: SEM). Numbers within the columns represented the independent samples used for each group.

Supplementary Table S2

Evolutionary conservation of tRNAs that harboring mse-tsRNA family 1-7

Species	Mse-tsRNA Family-1	Mse-tsRNA Family-2	Mse-tsRNA Family-3	Mse-tsRNA Family-4	Mse-tsRNA Family-5	Mse-tsRNA Family-6	Mse-tsRNA Family-7
Homo sapiens	+	+	+	+	+	+	+
Pan troglodytes	+	+	+	+	+	+	+
Gorilla gorilla gorilla	+	+	+	+	+	+	+
Pongo pygmaeus abelii	+	+	+	+	+	+	+
Nomascus leucogenys	+	+	+	+	+	+	+
Macaca mulatta	+	+	+	+	+	+	+
Callithrix jacchus	+	+	+	+	+	+	+
Tarsier syrichta	+	+	-	+	+	+	+
Microcebus murinus	+	+	+	+	+	+	+
Mus musculus	+	+	+	+	+	+	+
Rattus norvegicus	+	+	+	+	+	+	+
Cavia porcellus	+	+	+	+	+	-	+
Heterocephalus glaber	+	+	+	+	+	+	+
Oryctolagus cuniculus	+	+	+	+	+	+	+
Felis catus	+	+	+	+	-	+	+
Canis familiaris	+	+	+	+	+	+	+
Ailuropoda melanoleuca	-	+	-	+	+	+	+
Equus caballus	+	+	-	+	+	+	+
Bos taurus	+	+	+	+	+	+	+
Ovis aries	+	+	-	+	+	-	+
Sus scrofa	+	+	-	+	+	+	+
Loxodonta africana	+	+	+	+	+	+	+
Monodelphis domestica	+	+	+	-	+	+	+
Ornithorhynchus anatinus	+	+	-	-	+	+	+
Taeniopygia guttata	+	+	-	+	-	-	+
Gallus gallus	+	+	-	+	-	+	+
Meleagris gallopavo	+	+	-	-	-	+	+
Anolis carolinensis	-	+	-	-	-	+	+
Xenopus tropicalis	+	+	-	-	+	+	+
Oryzias latipes	+	+	-	-	+	+	+
Gasterosteus aculeatus	+	+	-	-	+	+	+
Takifugu rubripes	+	+	-	-	+	+	+
Tetraodon nigroviridis	+	+	-	-	+	+	+
Danio rerio	+	+	-	+	+	+	+
Petromyzon marinus	+	+	-	-	+	-	+
Drosophila melanogaster	-	-	-	-	-	-	-
Caenorhabditis elegans	-	-	-	-	-	-	-
Arabidopsis thaliana	-	-	-	-	-	-	-

Supplementary information, Data S1

RNA extraction, small RNA library preparation and sequencing

Mature sperm were isolated from cauda epididymis of adult male mice (CD1 background, 10-12 weeks), as previously performed [1]. Total RNA was isolated using the TRIzol Reagent (Invitrogen) according to the manufacturer's instructions. Before the standard protocol been performed, the sperm was treated with somatic cell lysis buffer (0.1% SDS, 0.5% Triton X in DEPC H_2O) for 20 min on ice to eliminate somatic cell contamination. RNA extraction for adult testis, uterus and elongated spermatids were performed with standard protocols[1]. The prepared RNA for Mature sperm, testis, uterus and elongated spermatids were shipped to the Beijing Genomics Institute (BGI), Shenzhen, China in dry ice for small RNA library construction and Solexa high-throughput sequencing followed their standard protocols[2]. Briefly, 18–40nt fraction was excised and purified, the obtained small RNA was ligated sequentially to 5′ - and 3′ -adapters, followed by RT-PCR to produce the sequencing library. The PCR products were purified and sequenced by Illumina Hi-Seq 2000.

Data processing and analysis

Sequence tags from the Solexa sequencing went through data cleaning by BGI standard protocols, which included filtering out low quality tags and several contaminants, using software developed by BGI. The small RNA clean reads were mapped with mouse genome (mm9) by SOAP (Short Oligonucleotide Analysis Package, developed by BGI) to analyze their expression and distribution on the genome. The small RNA reads that 100% match to genome were used for further analysis. Small RNA annotation was performed by Alignment to Rfam (http://rfam.sanger.ac.uk/) and GenBank (http://www.ncbi.nlm.nih.gov/genbank/). The genome locations of mature-sperm-enriched small RNA families described in present study were searched by BLAT (https://genome.ucsc.edu/cgi-bin/hgBlat) by mapping to NCBI37/mm9 database. The sequences of mse-tsRNAs were matched with **Genomic tRNA Database** (http://gtrnadb.ucsc.edu/) to obtain their position information from tRNAs' secondary structure and potential cleavage sites. Evolutional conservation analysis of mse-tsRNA family 1-7 and their precursor tRNAs were performed by matching to all species included in **Genomic tRNA Database.** Mse-tsRNA family 1-7 were searched in piRNA database (http://pirnabank.ibab.ac.in/index.shtml)[3] to confirmed their existence in different species. The reads number and percentage of mse-tsRNA family-1 and -2 in purified testicular spermatogenic cells types were performed by analyzing datasets for each cell type, the datasets for type A spermatogonia, pachytene spermatocytes and round spermatids have been published previously (GSM610965, GSM610966, GSM610967)

(http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE24822). The small RNA datasets performed in present study were deposited at NCBI: for Mature mouse sperm, testis and uterus (GSE38702), for elongated spermatids (GSE40397).

Purification of testicular spermatogenic cells

Purification of testicular spermatogenic cells (type A spermatogonia, pachytene spermatocytes, round spermatids and elongated spermatids) were performed as we previously described[4].

Purification of sperm heads

Purification of sperm heads from intact sperm were performed as previously described[5].

Collection of oocytes, zygotes and parthenogenetically activated oocytes for quantitative RT-PCR

To collect fertilized eggs (zygotes), CD1 female mice were superovulated by injecting (I.P) 10 IU of PMSG, followed by 10 IU of hCG after 48h, and were then mated with male mice. The fertilized eggs were collected 17h after hCG injection from vaginal plug positive females. The unfertilized oocytes were collected by the same protocols but without mating with male mice. For parthenogenetically activation of oocytes, the oocytes were first collected 13h after hCG injection, then cultured in the parthenogenetically activation media (10 mM $SrCl_2$ CZB) for 5h. For the quantitative examination, 100 embryos were pooled for RNA extraction (Trizol reagent) as an independent sample. We collected 5 samples for oocytes (100 x 5), 6 samples for zygotes (100 x 6) and 5 samples for parthenogenetically activated oocytes (100 x 5). Then the RNAs were extracted for each sample followed by quantitative RT-PCR for mse-tsRNA family-1.

RT-PCR for small RNAs

Reverse Transcription for small RNAs (Testicular spermatogenic cells, mature sperm from mice, rat and human) were performed using QuantiMir RT kit (System Biosciences) following the manufacturer's instruction. Briefly, 2 μg of RNA from each sample was polyadenylated, and then converted to cDNAs with a unique adaptor in the presence of reverse transcriptase, and the cDNAs were amplified with specific small RNA primer in combination with the universal adaptor to examine the expression of a particular small RNA. PCR mixtures contained 12.5 μL of 2× *GoTaq® Green Master Mix,* 0.5μL of forward primer (10 μM) and 0.5μL of universal primer (10 μM), 1 μL of cDNA and 10.5 μL of double-distilled water. For equal loading of each sample, the RNA concentration were measured and calibrated by a MICRO UV-VIS · FLUORESCENCE SPECTROPHOTOMETER (e-spect, Malcom, Japan). The conditions used for PCR were as follows:

For mse-tsRNAs, 95 °C for 10 min, followed by 23 or 28 cycles at 95 °C for 15 s, 60 °C for 15 s and 72 °C for 15 s. For miR-34c, the PCR conditions were 95 °C for 10min, followed by 28 cycles at 95 °C for 15 s, and 60 °C for 1 min. Primers and product sizes were shown in Figure 1H. The amplified products were analyzed by electrophoresis on 4% agarose gels stained with ethidium bromide. PCR products were purified with TaKaRa MiniBEST DNA Fragment Purification Kit and subcloned into pGEM-T easy vector (Promega) for sequencing. SYBR Green based Quantitative PCR examinations of mse-tsRNA family-1 for oocytes, zygotes and parthenogenetically activated oocytes were performed on Roche LightCycler 480 II.

References:

1. Chen Q, Peng H, Lei L *et al.*. Aquaporin3 is a sperm water channel essential for postcopulatory sperm osmoadaptation and migration. *Cell Res* 2011; **21** (6):922-933.
2. Hafner M, Landgraf P, Ludwig J *et al.*. Identification of microRNAs and other small regulatory RNAs using cDNA library sequencing. *Methods* 2008; **44** (1):3-12.
3. Sai LS, Agrawal S. piRNABank: a web resource on classified and clustered Piwi-interacting RNAs. *Nucleic Acids Res* 2008; **36** (Database issue):D173-D177.
4. Gan H, Lin X, Zhang Z *et al.*. piRNA profiling during specific stages of mouse spermatogenesis. *Rna* 2011; **17** (7):1191-1203.
5. Villegas J, Zárraga AM, Muller I *et al.* A novel chimeric mitochondrial RNA localized in the nucleus of mouse sperm. *DNA Cell Biol* 2000; **19** (9):579-588.