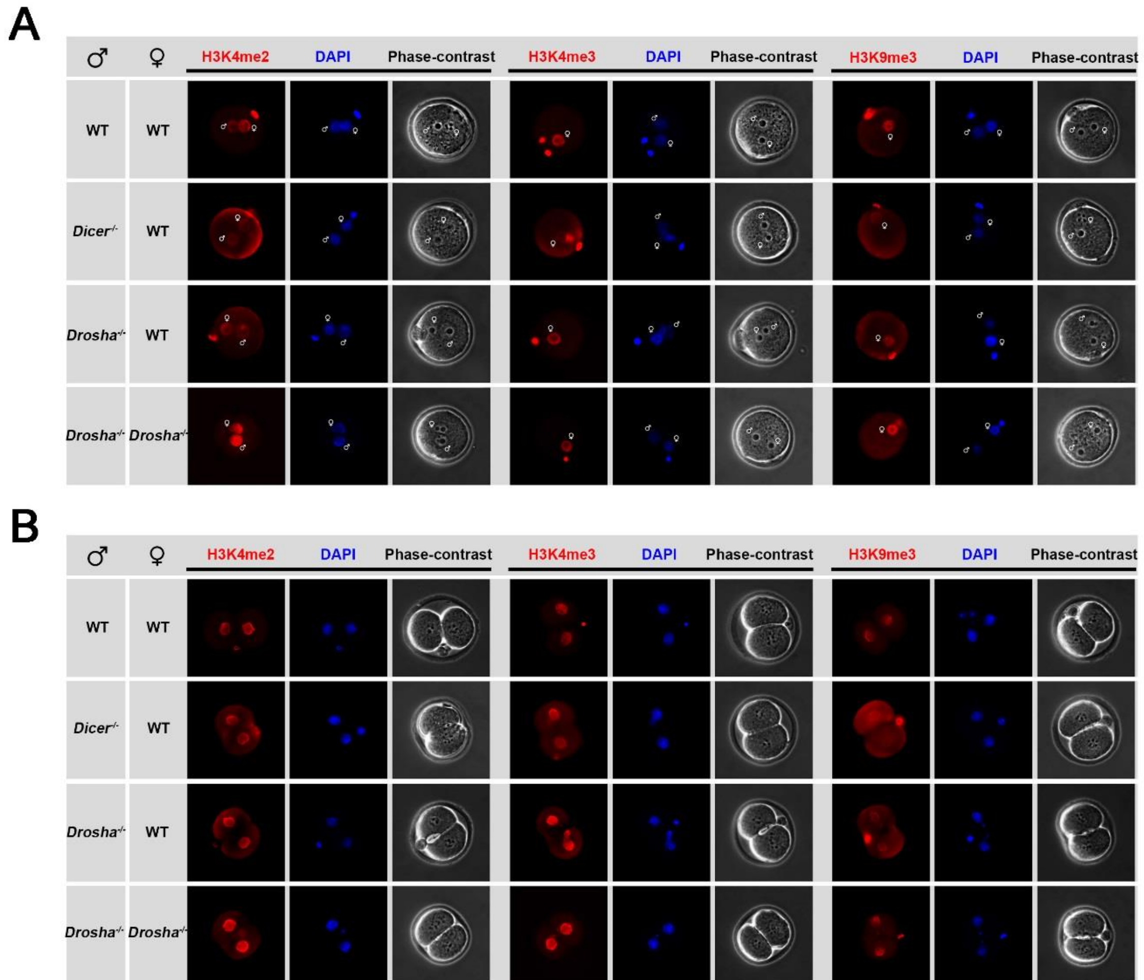

Fig. S1. Immunohistochemical analyses of H3K4me2, H3K4me3, and H3K9me3 in 2PN and 2-cell embryos derived from ICSI using WT, Dicer, and Drosha cKO sperm. (A) H3K4me2, H3K4me3 and H3K9me3 immunostaining patterns in 2PN embryos. (B) Immunostaining of H3K4me2, H3K4me3 and H3K9me3 in 2-cell embryos. Four types of ICSI were analyzed, including WT sperm + WT oocyte, Dicer cKO sperm + WT oocyte, Drosha cKO sperm + WT oocyte, and Drosha cKO sperm + Drosha cKO oocyte.

Development • Supplementary information

Table S1. miRNAs that are up-regulated in Dicer cKO sperm, as determined by sncRNA-Seq. miRNAs with a log_2 expression (relative to the WT control) value above 1 and adjusted p-values (p_{adj}) <0.05 were considered to be significantly up-regulated. Both the log_2 expression and p_{adj} values were determined by DESeq2.

Click here to Download Table S1

Table S2. miRNAs that are down-regulated in Dicer cKO sperm, as determined by sncRNA-Seq. miRNAs with a log_2 expression (relative to the WT control) value between -1 and -3.322 (i.e. a fold change of 10), and adjusted p-values (p_{adj}) <0.05 were considered to be significantly down-regulated. Both the log_2 expression and p_{adj} values were determined by DESeq2.

Click here to Download Table S2

Table S3. miRNAs that are absent in Dicer cKO sperm, as determined by sncRNA-Seq. Significantly down-regulated miRNAs with a fold change greater than 10 (log_2 value < -3.322) were considered as absent miRNAs. Significantly down-regulated miRNA are defined as miRNA with a log_2 expression (relative to the WT control) below -1 and adjusted p-values (p_{adj}) less than 0.05 (both determined by DESeq2).

Click here to Download Table S3

Development • Supplementary information

Table S4. miRNA with no significant changes in Dicer cKO sperm, as determined by sncRNA-Seq. miRNAs that did not have both a log_2 expression (relative to the WT control) value above 1 or below -1, and adjusted p-values (p_{adj}) <0.05 were not considered to be significantly changed in the Dicer-cKO sperm. Both the log_2 expression and p_{adj} values were determined by DESeq2.

Click here to Download Table S4

Table S5. miRNAs that are up-regulated in Drosha cKO sperm, as determined by sncRNA-Seq. miRNAs with a log_2 expression (relative to the WT control) value above 1 and adjusted p-values (p_{adj}) < 0.05 were considered to be significantly up-regulated. Both the log_2 expression and p_{adj} values were determined by DESeq2.

Click here to Download Table S5

Table S6. miRNAs that are down-regulated in Drosha cKO sperm, as determined by sncRNA-Seq. miRNAs with a log_2 expression (relative to the WT control) value between -1 and -3.322 (i.e., a fold change of 10), and adjusted p-values (p_{adj}) < 0.05 were considered to be significantly down-regulated. Both the log_2 expression and p_{adj} values were determined by DESeq2.

Click here to Download Table S6

Development • Supplementary information

Table S7. miRNA that are absent in Drosha cKO sperm, as determined by sncRNA-Seq. Significantly down-regulated miRNAs with a fold change greater than 10 (log_2 value < -3.322) were considered as absent miRNAs. Significantly down-regulated miRNAs are defined as miRNA with a log_2 expression (relative to the WT control) below -1 and adjusted p-values (p_{adj}) < 0.05 (both determined by DESeq2).

Click here to Download Table S7

Table S8. miRNA with no significant changes in Drosha cKO sperm, as determined by sncRNA-Seq. miRNAs that did not have both a log_2 expression (relative to the WT control) value above 1 or below -1, and adjusted p-values (p_{adj}) < 0.05 were not considered to be significantly changed in the Drosha-cKO sperm. Both the log_2 expression and p_{adj} values were determined by DESeq2.

Click here to Download Table S8

Table S9. Novel sperm-borne endo-siRNAs and their putative targets. A. The nucleotide sequences of novel endo-siRNAs identified from sncRNA-Seq analyses. B. Putative targets for each novel sperm endo-siRNA. The list of transcripts used in this study was comprised of cDNA from ENSEMBL (release 75; downloaded July 18th, 2014), refMrna from UCSC Genome Browser (mm10; downloaded July 18th, 2014), and repetitive elements from Repbase (v19.06; downloaded July 18th, 2014).

Click here to Download Table S9

Development • Supplementary information

Table S10. endo-siRNA expression in Dicer and Drosha cKO sperm, as determined by sncRNA-Seq. endo-siRNA expression levels (relative to the WT control) are provided as $\log_2$ values. endo-siRNA that had $\log_2$ values above 1 or below -1, and adjusted p-values (p_{adj}) < 0.05, were considered significantly up or down-regulated, respectively. Significantly down-regulated endo-siRNA with a fold change greater than 10 ($\log_2$ value < -3.322) were considered as absent endo-siRNA. Absent and up-regulated endo-siRNA are highlighted in yellow and red, respectively. Both the $\log_2$ and p_{adj} values were determined using DESeq2. endo-siRNAs with less than 1 normalized read in at least one sample type were removed prior to DESeq2 analysis, and are not shown above.

Click here to Download Table S10

Table S11. Preimplantation development of embryos derived from ICSI using Drosha cKO oocytes and Drosha cKO spermatozoa with or without sperm RNA supplementation.

Injected content	Total No. of surviving oocytes (no. of experiments)	No. of oocytes with 2PN (% of $\frac{2PN}{Total}$)	No. of 2-cell embryos (% of $\frac{2\text{-cell}}{2PN}$)	No. of 4-cell embryos (% of $\frac{4\text{-cell}}{2\text{-cell}}$)	No. of morula (% of $\frac{morula}{2\text{-cell}}$)	No. of Blastocyst (% of $\frac{Blastocyst}{2\text{-cell}}$
Drosha cKO	142 (9)	100 (70.42)	72 (72.00)	24 (33.33)	16 (22.22)	9 (12.50)
Drosha cKO + Small RNAs	106 (6)	81 (76.42)	48 (59.25)	26 (54.17)*	19 (39.58)*	9 (18.75)
Drosha cKO + Total RNAs	111 (6)	91 (81.98)*	59 (64.84)	28 (47.45)	16 (28.81)	9 (15.25)

Note: exp., experiment; Statistic analyses were conducted using χ^2 test, * p<0.05, ** p<0.01, as compared to Drosha cKO sperm.

Development • Supplementary information

Table S12. sncRNA-Seq data on miRNA expression in WT oocytes and 2PN embryos derived from ICSI using WT, Dicer cKO and Drosha cKO sperm.

Click here to Download Table S12

Table S13. High throughput real-time quantitative PCR (HT-qPCR) data on expression levels of 96 genes in single oocytes and preimplantation embryos at 2PN, 2-cell, 4-cell, morula and blastocyst stages, which were derived from ICSI using WT, Dicer and Drosha cKO sperm.

Click here to Download Table S13

Table S14. Fold changes of dysregulated genes (highlighted in red, green or grey, 39 out of 96) in Dicer cKO sperm-derived 2PN and 2-cell embryos and fold changes of miRNAs that target these dysregulated genes in Dicer cKO sperm.

Click here to Download Table S14

Table S15. Fold changes of dysregulated genes (highlighted in red, green or grey, 33 out of 96) in Drosha cKO sperm-derived 2PN and 2-cell embryos and fold changes of miRNAs that target these dysregulated genes in Drosha cKO sperm.

Click here to Download Table S15

Table S16. Sequences of primers used for high throughput qPCR analyses.

Gene symbol	Forward (5' → 3')	Reverse (5' →3')
Actb	GACGGCCAGGTCATCACTAT	ATGCCACAGGATTCCATACC
Ahcy	CCCTACAAAGTCGCGGACATC	GAGGCTGAGTACATCTCCCG
Aqp3	GCTTTTGGCTTCGCTGTCAC	TAGATGGGCAGCTTGATCCAG
Atp12a	ATGCGCCGGAAAACAGAAATC	CCTCCTCCTGACTCTTGTTGG
Axin1	CTCCAAGCAGAGGACAAAATCA	GGATGGGTTCCCCACAGAAATA
Bmp4	AGGAGGAGGAGGAAGAGCAG	CCTGGGATGTTCTCCAGATG
Bmpr1a	CTGTACATAAGGCATTGACAGG	TAATGGTAAACTTCATAAGGCACC
Bmpr2	CTGCTTTAACCTCCTGTCAGC	GGAGACCACTTTGGATAAGCAC
Cdc25c	CCCTGAATCTCCGAAAGACA	CGAAAGGTCAAGGCAACATT
Cdkn2d	CTTCTTCACCGGGAGCTG	TGCTGGACTTCCAAACATCA
Cdx2	TCACCATCAGGAGGAAAAGTG	CTGCGGTTCTGAAACCAAAT
Cebpa	CAAGAACAGCAACGAGTACCG	GTCACTGGTCAACTCCAGCAC
Creb3	AAGGCTCCGCTGGACTTAGA	TGTGGAAGGGAGTAGTTGTGA
Creb3l2	CATGTACCACACGCACTTCTC	CCACCTCCATTGACTCGCTC

Development • Supplementary information

Dazl	ATACCTCCGGCTTATACAACTGT	GACTTCTTTTGCGGGCCATTT
Dicer1	GAGCTACAGGATCGCCAAGTC	GTGGAGCTGTGGTTCTGGTC
Dnmt1	GCCAGTTGTGTGACTTGGAA	AGTGTGTGTTCCGTTCTCCA
Dnmt3a	GACTCGCGTGCAATAACCTTAG	GGTCACTTTCCCTCACTCTGG
Dnmt3b	CTCGCAAGGTGTGGGCTTTTGTAAC	CTGGGCATCTGTCATCTTTGCACC
Dppa1	ATGATGTCCCTTCAAGTCCTCA	TGTGTTGGGATCACTTCAGTGT
Dppa2	TCAACGAGAACCAATCTGAGGA	GCGTAGCGTAGTCTGTGTTTG
Dppa3	AGGCTCGAAGGAAATGAGTTTG	TCCTAATTCTTCCCGATTTTCG
Dppa5a	AAGGAGTGCTGAAGCTGGAGG	CAGCTTAACCTGCATCCAGGTC
Drosha	GGACCATCACGAAGGACACT	CACGGGTCTCTTGGTTTTGT
Eed	AAGGGCACATAGAGCATTTAGAC	CGAGCAGGAAGACAGTACAAAG
Ehmt2	CTTCTTCAGCTCCAGGGACATC	GAATGCTTGCACTTCTCAGAGC
Ago2	TGCTAGGTCTCCACAGTGTCTC	CAATCTCTCGGACTTGTCACTG
Elavl2	ACACAGCCAATGGTCCAACC	TTCCCGGAGTCAACTGGTGA
Eomes	CCCTATGGCTCAAATTCCAC	TGGGGTTGAGTCCGTTTATG
Eras	GTAGCTGTGGCTGCTCTGTAG	GATGTCTGTGGTAACTTGGTCG
Esrrb	GGCGTTCTTCAAGAGAACCA	CCCACTTTGAGGCATTTCAT
Etv5	TCAGTCTGATAACTTGGTGCTTC	GGCTTCCTATCGTAGGCACAA
Ezh2	TCAGGAACCTTGAGTACTGTGG	CTTTGCAGCTGGTGAGAAGGC
Fgf4	CGAGGGACAGTCTTCTGGAG	TCGGTAAAGAAAGGCACACC
Fgfr2	AATCTCCCAACCAGAAGCGTA	CTCCCCAATAAGCACTGTCCT
Fn1	GGAGGGAAAACAGAGCTTGA	GGGTCTACTCCACCGAACAA
Foxh1	GACCTGCTCTGTGATCTAGAC	ATGCTGTACCAGGAAAGGAGC

Gapdh	AACTTTGGCATTGTGGAAGG	GGATGCAGGGATGATGTTCT
Gata3	CTTATCAAGCCCAAGCGAAG	CATTAGCGTTCCTCCTCCAG
Gata4	CTACCTGTGCAATGCCTGTG	GGTGGTGGTAGTCTGGCAGT
Gata6	TTGCTCCGGTAACAGCAGTG	GTGGTCGCTTGTGTAGAAGGA
Grhl1	CAGTTCAGCTCTGGCACTCA	GAGGGGAAGAAAACCTCCTG
Grhl2	AAGCCCAGTGCAACAACTCC	TGTCCGGTCCTCTGTAGGTTT
Hand1	TCAAAAAGACGGATGGTGGT	GCGCCCTTTAATCCTCTTCT
Hnf4a	CACGCGGAGGTCAAGCTAC	CCCAGAGATGGGAGAGGTGAT
Hprt	TCAGTCAACGGGGGACATAAA	GGGGCTGTACTGCTTAACCAG
Id2	CCAGAGACCTGGACAGAACC	ATTCAGATGCCTGCAAGGAC
Ifitm3	TGGTCCTCAGCATCCTGATGG	AGGGTGAAGCACTTCAGGACC
Irf8	CATGCAAACAGAATCCTTGGAG	GAACAGGTTGTCCTGGCTTCG
Klf2	CTCAGCGAGCCTATCTTGCC	CACGTTGTTTAGGTCCTCATCC
Klf4	GCGAGTCTGACATGGCTGT	AGAGAGTTCCTCACGCCAAC
Klf5	CTCCGGAGACGATCTGAAAC	TCTTGTCTGGACCAGCTCCT
Krt8	GCACTCAGGAGAAGGAGCAG	ATGTTGCTCCTCGACGTCTT
Lcp1	TCCGTGTCTGACGAAGAAATG	GCGGCCTTGAACAAGTCAT
Mbnl3	GTCCGGTTCTGATTTCTGGA	CTGAAATTCACGGCAAACCT
Ybx2	CTTCTATCGAAGGCGGTTTG	GCTGTTGATCACCCTCCAAT
Nanog	ATGCCTGCAGTTTTTCATCC	GAGGCAGGTCTTCAGAGGAA
Nes	CCCTGAAGTCGAGGAGCTG	CTGCTGCACCTCTAAGCGA
Neurog3	AGTGCTCAGTTCCAATTCCAC	CGGCTTCTTCGCTTTTTGCTG
Nodal	AGCCACTGTCCAGTTCTCCAG	GTGTCTGCCAAGCATACATCTC

Pdgfa	TGTGCCCATTCGCAGGAAG	GAGGTATCTCGTAAATGACCGTC
Pdgfra	TCCATGCTAGACTCAGAAGTCA	TCCCGGTGGACACAATTTTTC
Pecam1	ACGCTGGTGCTCTATGCAAG	TCAGTTGCTGCCCATTCATCA
Pou5f1	AGAGGATCACCTTGGGGTACA	CGAAGCGACAGATGGTGGTC
Prdm1	AGCATGACCTGACATTGACACC	CTCAACACTCTCATGTAAGAGGC
Prdm4	ACTCCCTGTGAGCAGGCACAC	CAGCAGAGTCCTCTACTCAAGCC
Prdm5	GGCAGAGTGCCATTTCTGCCATAAG	CACATGGTGTGGTTCCTTGCCG
Runx1	GCCATGAAGAACCAGGTAGC	GACGGTGATGGTCAGAGTGA
Sall1	CTCAACATTTCCAATCCGACCC	GGCATCCTTGCTCTTAGTGGG
Sall4	AGTGTCACCTGCCAATAGCC	GGCTGTGCTCGGATAAATGT
Smad1	CTTCTGTTCGCAAATCAACTGG	AATGAACAGAGTTACCAGGTTTGG
Smad3	GTGGACTCTTGTTTAGGAGGC	GGACACGGCTCTTTAACAATGG
Smad4	GTATCTTAGGGCAAGACTGCAGAC	GTCTGTGGTACAGTCAATGTGTC
Smad5	TACAGAGCGAGTGCTTGTGTTC	CATACACAGAAGGAGGCAGACC
Snai1	CACACGCTGCCTTGTGTCT	GGTCAGCAAAAGCACGGTT
Sox13	CTCATCAGCCTGGACTCCTC	TCTTGATGTGGCTGCTGTTC
Sox17	CTTTATGGTGTGGGCCAAAG	GCTTCTCTGCCAAGGTCAAC
Sox2	GCGGAGTGGAAACTTTTGTCC	CGGGAAGCGTGTACTTATCCTT
Stat1	ACCTTACACCCCGAGTGCAG	CTTGGTCATTTCATTTTGGCATGG
Stat2	ATGGCACTGCTTCTTTCTGCTC	ACGAAGGACTTCAGGACAACAG
Stat5a	GGTGAAGGCGACCATCATCAG	GTGGTACTCCATGACGCAACAG
Stat5b	GAAGTGGCTATCTCCTACAACG	GTGGCTGCCTCAATTTATACAC
T	ATCAGAGTCCTTTGCTAGGTAG	GTTACAATCTTCTGGCTATGC

Tcf23	GAGTCGCATCAACAGGACAAG	GTCTTCACCCGAGTCCGTT
Tfap2a	GAAGACTGCGAGCACGTAGAA	GTCCTTGTTAATAGGGATGGCG
Tfap2c	CAGAGGGCTCGAAGGATTC	AGAGATGTTGTCCCGGAGAG
Terf1	GAGACACTGGAGGCTGGATG	CTCTTTCTCTTCCCCCTCCTC
Tert	GGAGATGAAAAAGGACATCGAG	CTGCTCCAGTAGGAAAATCCTG
Trim43c	AGGCTCAAAATCCAGCAGAA	CTTCACTGAGGGGTGGACAT
Tspan8	TCTGGGTATGTGGTACACTGAT	AGGGGTTCGTGCTAGAGTCTC
Uchl1	TCACGGCCCAGCATGAAAA	ACAGGAGTTTCCGATGGTCTG
Vim	CGGCTGCGAGAGAAATTGC	CCACTTTCCGTTCAAGGTCAAG
Wnt5a	TATCCATTCATGGGCTTGGTAG	TTGCAATTCATTGTAGCTGAGG
Wnt5b	CCTGATAACTGGGATGGTAGCC	GGTTGGAGTGACTTGAGACCAC
Yy1	GCCAGATGCTGATGTTCAGTG	TTGTTGCCCTTTCTGTTACACG
Zfp42	TCCATGGCATAGTTCCAACAG	TAACTGATTTTCTGCCGTATGC

Supplementary Materials and Methods

Intracytoplasmic sperm injection (ICSI)

ICSI was carried out as described (Kimura and Yanagimachi, 1995; Stein and Schultz, 2010), with some medications. Briefly, WT and cKO sperm were suspended in 500µl HEPES-CZB medium followed by centrifugation at 700g for 5min. The sperm pellet was resuspended with 200µl NIM/PVA medium (Stein and Schultz, 2010), and the suspension was sonicated four times for 15 second each. After sonication, ~50% of the sperm heads were separate from the tails; an aliquot of 1-2µl sperm suspension was diluted with 50µl NIM/PVA medium. Single sperm head was picked up from the sperm NIM/PVA suspension, and injected it into WT or *Zp3-Cre;Drosha*$^{lox/del}$ oocytes using a Piezo drill under the control of an electric micromanipulator (TransferMan NK2, Eppendorf) at room temperature. Injection of ~15 oocytes was completed within 5 min. Injected oocytes were then transferred to KSOM+AA medium covered by mineral oil followed by culture at 37°C in an incubator with humidified atmosphere and 5% CO_2. Each ICSI process was completed within 2h after oocyte collection.

Single cell high-throughput qPCR (HT qPCR)

Single oocytes and embryos were harvested individually into tubes with one oocyte or one embryo in each. 5µl CellDirect 2X Reaction Mix lysis buffer (Invitrogen, One-Step qRT-PCR Kit, Grand Island, NY) was added into each tube containing one single oocyte or embryo before snap freezing in liquid nitrogen followed by storage at -80°C. Single cell high-throughput qPCR (HT qPCR) was performed in two steps. The first included reverse transcription and specific target amplification (STA). In brief, total RNA from single MII oocytes and single embryos was reverse-transcribed to cDNAs, and the cDNAs were then pre-amplified for 18~20 cycles using a mixture of 96 pairs of primers designed to amplify 96 genes of interest (Table S16). The pre-amplified amplicons were cleaned up using ExoSAP-IT (Affymetrix) to eliminate unincorporated primers. The second step was the HT qPCR, which was performed in the 96 X 96 chips using the Biomark™ HD system (Fluidigm, South San Francisco, CA), according to the protocols provided by the manufacturer. The final data were analyzed using the Biomark real-time PCR analysis software (Fluidigm, Version 3.1.2). *Gapdh* was used as an internal control for data normalization. A heat-map was generated based on normalized values and the original HT qPCR dataset can be found in Table S13.

Small noncoding RNA deep sequencing (sncRNA-Seq)

The lysates were processed using the Ion Total RNA-Seq Kit v2 with modifications. In brief, 100% alcohol was used in the purification step to minimize the sample loss, and 30 amplification cycles were used to ensure the capture of low-copy small RNAs. The pre-amplified libraries were then subjected to small RNA enrichment and removal of amplification noise by size selection using a Pippin Prep™ Kit 3010. Barcodes were added using the Ion Xpress RNA-Seq Barcode 1-16 kit (Invitrogen). The final sncRNA cDNA libraries were sequenced on the Ion Proton sequencer (Life Technologies) using the P1 chips in Nevada Genomics Center. In the sperm sncRNA-Seq experiments, one replicate from the WT group and one from the Dicer cKO group were not annotated

Development • Supplementary information

due to poor sequencing quality (R^2 values between other replicates > 0.1), leaving duplicates for WT and Dicer cKO, and triplicate sequencing data for Drosha cKO samples for annotation.

SncRNA-Seq data analysis processing

The oocyte and 2PN embryo sncRNA-Seq data was annotated as follows: Reads shorter than 15nt were discarded. The remaining reads were matched to known murine sncRNA, consisting of miRNA (miRBase, release 21), tRNA (Genomic tRNA Database, mm10), piRNA (piRNABank), rRNA (ENSEMBL, release 76), snoRNA (ENSEMBL, release 76), snRNA (ENSEMBL, release 76), and mitochondrial RNA (ENSEMBL, release 76) using Sequery (up to 2 mismatches allowed). Unmatched reads were matched to mouse testis(Song et al., 2011) and putative sperm endo-siRNAs with Sequery (0 mismatches allowed). The remaining unmatched reads were aligned to the mouse genome (mm10) via Bowtie (settings -n 2 -k 3 --best -S --al -q) (Langmead et al., 2009). Aligned reads were matched to the genomic coordinates of known murine miRNA, tRNA, rRNA, snoRNA, snRNA, and mitochondrial RNA (same databases used previously). Read counts were obtained by in-house Python scripts. Raw reads were normalized to the number of total aligned reads (via Bowtie, n = 2) and are provided as reads per million aligned reads (RPM). Genes with fewer than 3 RPM in every oocyte and 2PN sample were not included in subsequent analyses. To compare the miRNA present in the sperm, oocyte, and 2PN data, miRNA that had at least 3RPM were defined as being 'present' in the sample. After applying this cutoff, we determined which miRNA were 'present' in WT 2PN and sperm, but not WT oocyte. The list of WT 2PN and sperm specific miRNA was then compared to the miRNA that were 'present' in the Drosha cKO and Dicer cKO sperm data.

For sperm sncRNA-Seq data, prior to novel endo-siRNA identification, known sncRNAs and novel miRNAs were removed using a previously described method, with slight modifications (Ortogero N, 2015). Known murine sncRNA populations were acquired from miRBase, the genomic tRNA Data Base, piRNABank, ENSEMBL (rRNA, snoRNA, snRNA, mito-tRNA, mito-rRNA) using the most up-to-date versions at the time of download (6/24/2014) (Sai Lakshmi and Agrawal, 2008; Chan and Lowe, 2009; Flicek et al., 2014; Kozomara and Griffiths-Jones, 2014). Endo-siRNAs from a previous study in testis were also included (Song et al., 2011). Novel miRNA identification was performed using miRDeep* (An et al., 2013). Novel endo-siRNA analysis was performed using a previously described method (Schuster A, 2015). The murine transcriptome used consisted of sequences from ENSEMBL (cDNA sequences), UCSC (refMrna), and Repbase (repetitive sequences, e.g. retrotransposons), using the most up-to-date versions at the time of download (7/18/2014) (Jurka et al., 2005; Flicek et al., 2014; Karolchik et al., 2014). Endo-siRNA and miRNA expression levels were determined via Sequery (v1.0) using the newly identified sperm endo-siRNA and mature miRNA (miRBase, release 20) sequences as the reference, respectively (Ortogero et al., 2013; Kozomara and Griffiths-Jones, 2014). Up to 2 mismatches were allowed for miRNA matching; 0 mismatches were allowed for endo-siRNA matching. Raw reads were normalized to account for sequencing depth by dividing each sncRNA read count by the respective library's total number of aligned reads (aligned with Bowtie, k = 1, n = 2) and multiplying by a factor of a million resulting in reads per million (RPM) values (Langmead et al., 2009). Differential expression was measured by DESeq2 under default parameters, using raw read counts lacking all sncRNA species that had fewer than 3 normalized reads in all sample types (Drosha-cKO, Dicer-cKO, and WT). Differential expression analysis for miRNA and endo-siRNA was performed separately. Significant dysregulation was defined as having an adjusted p-value ≥ 0.05 and a $\log_2$ fold-change ≥ 1 (up-regulated) or ≤ -1(down-regulated) (Michael I Love, 2014).

Development • Supplementary information

RNA-Seq raw data analysis

Multiple programs (i.e., Bowtie and Sequery) were used to process our sncRNA-Seq data in order to account for confounding factors unique to sncRNA annotation. Many sncRNA databases (e.g., miRBase) use consensus sequences and oftentimes, slight variations are present in the 5' and 3' boundaries of sncRNA genes. To ensure that we did not identify any previously known sncRNAs as endo-siRNAs in our subsequent analysis, it was critical that we annotated our sncRNA-Seq data as thoroughly as possible. The programs we selected allowed us to use very up-to-date sncRNA references and to incorporate lesser-known sncRNAs (e.g., small nucleolar RNA, snoRNA). The first program we used in our annotation pipeline, Sequery, matches the individual sequences (reads) in the sncRNA-Seq data against the consensus sequences of known mouse sncRNA (reference), supplied by the user. The number of mismatches between the reads and reference sequences that was allowed was two for all the sncRNAs except endo-siRNAs; because endo-siRNAs bind to their targets with complete complementarity, their sequences are not as variable as other sncRNAs (e.g., miRNAs), so zero mismatches were allowed for endo-siRNA-to-read matching. Next, the reads which did not match to any sncRNA sequences in the reference were aligned to the mouse genome using Bowtie. Our decision to use Bowtie instead the newer version, Bowtie2, was based on the former being faster and more sensitive when aligning short reads (>50nt). The genomic coordinates (e.g., Chr1: 1000-1020) for each aligned read were matched against the coordinates of known mouse sncRNA genes, using Python scripts we developed in-house. Therefore, using coordinates allowed us to account for aforementioned boundary variations that are sometimes an issue for sncRNA genes (e.g., with the criteria we used, a read aligned at Chr1:1000-1020 would be matched to a gene at Chr1:1003-1023). By running our sncRNA-Seq data through the two programs, we accounted for sncRNA variations and ensured that the remaining reads did not represent currently known sncRNA.

Bioinformatic analysis

TargetScan was used for predicting the identified miRNA families to target the 96 genes. Fold changes of the identified miRNAs in Dicer and Drosha cKO sperm were then determined based on our sncRNA-Seq data (Tables S1-S8). For a particular dysregulated mRNA, the absolute values of fold changes of all its targeting miRNAs were summed, and the total fold changes ≥2 were regarded as miRNAs with significant effects on their target mRNAs. Log_2 fold changes of dysregulated mRNAs at 2PN and 2-cell stages were plotted, and the dysregulated mRNAs with significant changes in their targeting miRNAs in cKO sperm (total fold changes ≥2) were indicated by red fonts. Those dysregulated mRNAs without significant changes in their targeting miRNAs in cKO sperm (total fold changes <2) were indicated by green fonts. The transcripts without known targeting miRNAs in the databases were marked with black font.

Development • Supplementary information

References

An, J., Lai, J., Lehman, M. L. and Nelson, C. C. (2013) 'miRDeep*: an integrated application tool for miRNA identification from RNA sequencing data', *Nucleic Acids Res* 41(2): 727-37.

Chan, P. P. and Lowe, T. M. (2009) 'GtRNAdb: a database of transfer RNA genes detected in genomic sequence', *Nucleic Acids Res* 37(Database issue): D93-7.

Flicek, P., Amode, M. R., Barrell, D., Beal, K., Billis, K., Brent, S., Carvalho-Silva, D., Clapham, P., Coates, G., Fitzgerald, S. et al. (2014) 'Ensembl 2014', *Nucleic Acids Res* 42(Database issue): D749-55.

Jurka, J., Kapitonov, V. V., Pavlicek, A., Klonowski, P., Kohany, O. and Walichiewicz, J. (2005) 'Repbase Update, a database of eukaryotic repetitive elements', *Cytogenet Genome Res* 110(1-4): 462-7.

Karolchik, D., Barber, G. P., Casper, J., Clawson, H., Cline, M. S., Diekhans, M., Dreszer, T. R., Fujita, P. A., Guruvadoo, L., Haeussler, M. et al. (2014) 'The UCSC Genome Browser database: 2014 update', *Nucleic Acids Res* 42(Database issue): D764-70.

Kimura, Y. and Yanagimachi, R. (1995) 'Intracytoplasmic sperm injection in the mouse', *Biol Reprod* 52(4): 709-20.

Kozomara, A. and Griffiths-Jones, S. (2014) 'miRBase: annotating high confidence microRNAs using deep sequencing data', *Nucleic Acids Res* 42(D1): D68-D73.

Langmead, B., Trapnell, C., Pop, M. and Salzberg, S. L. (2009) 'Ultrafast and memory-efficient alignment of short DNA sequences to the human genome', *Genome Biol* 10(3): R25.

Michael I Love, W. H., Simon Anders (2014) 'Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2', *BioRxiv*: 1-50.

Ortogero, N., Hennig, G. W., Langille, C., Ro, S., McCarrey, J. R. and Yan, W. (2013) 'Computer-assisted annotation of murine Sertoli cell small RNA transcriptome', *Biol Reprod* 88(1): 3.

Ortogero N, H. G., Luong D, Yan W (2015) Computer-assisted annotation of small RNA transcriptomes. in M. Sioud (ed.) *RNA Interference*, vol. 1218: Humana Press.

Sai Lakshmi, S. and Agrawal, S. (2008) 'piRNABank: a web resource on classified and clustered Piwi-interacting RNAs', *Nucleic Acids Res* 36(Database issue): D173-7.

Schuster A, H. G., Ortogero N, Luong D, Yan W (2015) In silico identification of novel endo-siRNA. in M. Sioud (ed.) *RNA Interference*, vol. 1218.

Song, R., Hennig, G. W., Wu, Q., Jose, C., Zheng, H. and Yan, W. (2011) 'Male germ cells express abundant endogenous siRNAs', *Proc Natl Acad Sci U S A* 108(32): 13159-64.

Stein, P. and Schultz, R. M. (2010) 'ICSI in the mouse', *Methods Enzymol* 476: 251-62.