METHODS

Mouse strains. Either CF-1 or CD-1 wild-type mice of 4–6 weeks were purchased from Harlan or Charles River Laboratories, respectively, and used to obtain oocytes for small RNA isolation. The *Dicer*$^{flox/flox}$ and *Dicer*$^{flox/flox}$ *Zp3-cre* mice, as previously reported[14], were used to obtain Dicer-deficient oocytes.

Generation of oocyte small RNA libraries. Wild-type mice were primed with 5 IU (international unit) PMSG 48 h before being killed, and fully grown germinal vesicle oocytes were collected as previously described[29]. Total RNA was extracted using Trizol (Invitrogen) according the manufacturer's protocol, and small RNA cloning was performed as described[4].

Quantitative real-time PCR. Total RNA was extracted from fully grown oocytes from *Dicer*$^{flox/flox}$ and *Dicer*$^{flox/flox}$ *Zp3-cre* mice using the Absolutely RNA Microprep Kit (Stratagene). cDNA was prepared by reverse transcription of total RNA with Superscript II and random hexamer primers. One oocyte equivalent of the resulting cDNA was amplified using TaqMan probes and the ABI Prism Sequence Detection System 7000 (Applied Biosystems). Three replicates of 45 oocytes each were used for RNA isolation and two replicates were run for each real-time PCR reaction; a minus template served as control. Quantification was normalized to the endogenous upstream binding factor (*Ubf*) within the log-linear phase of the amplification curve obtained for each probe/primer using the comparative C_T method (ABI PRISM 7700 Sequence Detection System, User Bulletin 2, Applied Biosystems, 1997). The TaqMan gene expression assays used were: Mm00441071_m1 (*Rangap1*), Mm00835842_g1 (*Kifc1*), Mm00620601_m1 (*Oog4*), Mm00786153_s1 (*Lcp1*), Mm00728630_s1 (*Kif2c*), Mm02391771_g1 (*Hdac1*), Mm00487521_m1 (*Mad1l1*), Mm00725286_m1 (*Optn*), Mm00833431_g1 (*Hsp90ab1*), Mm00511698_m1 (*Ppp2r2b*), Mm00801709_m1 (*Emp2*), Mm00486494_m1 (*Surf6*), Mm00456972_m1 (*Ubf*). For *Bcn1* and *Ubc9*, custom TaqMan Gene Expression Assays were used that had the following primers and probes: *Bcn1* forward primer 5′-ACTGGACGCTTCAGGATTACATC-3′, *Bcn1* reverse primer 5′-GTCATGATGCTCCAGTGATCCA-3′, *Bcn1* probe 5′FAM-TTCCCAGAGGCATCCTG-3′; *Ubc9* forward primer 5′-CAGGTGAGAGCC-AAGGACAAA-3′, *Ubc9* reverse primer 5′-GGCCCACTGTACAGCTAACA-3′, *Ubc9* probe 5′FAM-CTGGCCTGCATTGATC-3′.

Bioinformatic analysis. Small RNAs were sequenced using the Illumina 1G platform. Sequencing of the upper and lower fraction libraries produced 2,785,080 reads, of which 1,037,355 (37%) could be mapped to the mouse genome (release mm9, July 2007) with no mismatches. The small RNAs are matched to a suffix array generated from the mouse genome, keeping track of exact matches to genome. Repeat masking was not used, but small RNA sequences with more that ten identical nucleotides in a row were removed from consideration. Annotation categories were assigned based on the annotation of corresponding genomic sequences extracted from the UCSC genome browser. The genome was annotated with mRNAs, non-coding RNAs and repeats. The annotations at the mapping positions (up to five mappings per small RNA) were used, along with a majority rule, to assign an annotation to each small RNA. The sequences were also re-analysed to allow 1–2 mismatches to the genome. Although the number of mapped sequences increased (from 37% to 54%), the genomic origin of repeat-associated small RNAs became ambiguous (Supplementary Fig. 1), and therefore non-informative.

To extract small RNA clusters (both piRNA and siRNA), the genome was scanned to look for regions that had more than ten uniquely mapping small RNAs, and the boundaries were defined as the location of the first/last small RNA in the cluster. To identify sequences that match the consensuses for transposable elements, the small RNAs were aligned to consensus sequences from release 11.08 of Repbase (http://www.girinst.org). The following consensuses were used: L1_MM for LINE L1 and IAPLTR1a_I_MM for the IAP retrotransposon. Matches to consensus sequences with up to three mismatches were recovered and included in the analysis. Nucleotide biases were calculated for small RNAs matching L1 and IAP consensuses as described[4]. To identify gene–pseudogene pairs, the genomic sequences of the siRNA clusters were extracted from the UCSC genome browser, and re-matched to the genome using BLAT (http://www.genome.ucsc.edu). Genomic regions with greater than 95% identity were identified, and small RNAs (both sense and antisense) mapped to these locations were extracted.

Gene ontogeny analysis of endo-siRNA targets was carried out as previously described[30] using GOBAR, which uses a hypergeometric statistic to identify nodes that are significantly enriched. A bootstrapping technique, involving repeated sampling from the reference set, is used to assign significance values to the results.

29. Schultz, R. M., Montgomery, R. R. & Belanoff, J. R. Regulation of mouse oocyte meiotic maturation: implication of a decrease in oocyte cAMP and protein dephosphorylation in commitment to resume meiosis. *Dev. Biol.* **97**, 264–273 (1983).
30. Lee, J. S., Katari, G. & Sachidanandam, R. GObar: a gene ontology based analysis and visualization tool for gene sets. *BMC Bioinformatics* **6**, 189 (2005).