Supplementary Materials for

Biogenesis and function of tRNA fragments during sperm maturation and fertilization in mammals

Upasna Sharma, Colin C. Conine, Jeremy M. Shea, Ana Boskovic, Alan G. Derr, Xin Y. Bing, Clemence Belleannee, Alper Kucukural, Ryan W. Serra, Fengyun Sun, Lina Song, Benjamin R. Carone, Emiliano P. Ricci, Xin Z. Li, Lucas Fauquier, Melissa J. Moore, Robert Sullivan, Craig C. Mello, Manuel Garber, and Oliver J. Rando

correspondence to: oliver.rando@umassmed.edu

This PDF file includes:

Materials and Methods
Figs. S1 to S13
Captions for databases S1 to S8

Other Supplementary Materials for this manuscript includes the following:

Databases S1 to S8 as zipped archives:
Table S1. Complete small RNA dataset.
Table S2. Small RNA abundance for all tissues.
Table S3. Small RNA-Seq of B. taurus caput sperm reconstitutions.
Table S4. ES cell mRNA abundance.
Table S5. tRF-Gly-GCC effects on translation in ES cells.
Table S6. tRF-Gly-GCC effects on 4-cell stage gene expression.
Table S7. Dietary effects on preimplantation gene regulation.
Table S8. RNA effects on preimplantation gene regulation.

Materials and Methods

Mouse husbandry

Mice used in this study were primarily FVB/NJ strain background, obtained from Jackson Laboratories. All animal care and use procedures were in accordance with guidelines of the Institutional Animal Care and Use Committee. Animals were raised on one of two diets – defined Control diet (Bioserv AIN-93g) or a Low Protein diet based on AIN-93g (10% of protein rather than 19%, remaining mass made up with sucrose) – as previously described (*2*). Importantly, as we have found in natural matings that paternal dietary effects are substantially less penetrant when using females from our long term mouse colony, we restricted all experiments here to females whose parents or grandparents had been obtained from the animal vendor.

As we have previously found that siblings, whatever their diets, are more epigenomically similar than are Control animals that are not siblings, all analyses are restricted to paired siblings. In other words, all dietary effects on small RNAs in sperm, testis, etc., were assessed only for pairs of littermates split to Control or Low Protein diet, or for pools of animals split to diets in which all animals in a given pool are matched one to one with littermates in the other pool. For example, in **Fig. 1**, the eight pairs of sperm samples include seven pairs of individual sibling males, and one pair of sperm pools with two animals in each pool, with all four animals in the two matched pools being from the same litter. Similarly, paternal dietary effects on offspring metabolism or preimplantation gene regulation always utilize male siblings on different diets as the sperm donors. Finally, the majority of analyses of early embryonic gene expression were carried out using clutches oocytes from individual females (typically ~20-25 oocytes per female) that had been split into two groups for Control or Low Protein IVF (using sperm from sibling males), or for Control IVF embryos that were treated with and without tRF-Gly-GCC injection (using the same Control sperm sample in both cases for IVF), and so forth. Experimental data were then combined for multiple such paired experiments.

Testis, epididymis, and sperm collection:

Testes were dissected from 10-12 week mice fed on Control or Low Protein diet, directly frozen in liquid nitrogen and stored at -80 °C until RNA extraction. Cauda and Caput epididymides were dissected from mice and placed in Whitten's Media pH 7.4 (100 mM NaCl, 4.7 mM KCl, 1.2 mM KH_2PO_4, 1.2 mM $MgSO_4$, 5.5 mM Glucose, 1 mM Pyruvic acid, 4.8 mM Lactic acid (hemicalcium), and HEPES 20 mM) at 37 °C. To collect caput sperm two incisions were made at the distal end of caput and using a 26G needle holes were poked in the rest of the tissue to let the caput sperm ooze out. For cauda sperm collection, cauda epididymides were gently squeezed to allow the caudal fluid to ooze out. After incubation for 15 minutes at 37 °C, sperm containing media was transferred to a fresh tube where they were incubated for another 15 minutes and the sperm-free epididymis tissues were directly frozen in liquid nitrogen and stored at -80 °C. After incubation at 37 °C for total of 30 minutes, sperm were collected by centrifugation at 2000 x g for 2 minutes, followed by a 1X PBS wash, and a second wash with lysis buffer (0.1% SDS and 0.5% Triton-X) for 10 minutes on ice to eliminate somatic cell contamination, and finally washed with 1X PBS before freezing down. Unlike cauda sperm which are motile, caput sperm cannot swim up to allow a pure sperm sample collection. For the primary caput sperm dataset, sperm were obtained from 8 pairs of

caput epididymis, pelleted, washed with PBS, and washed again with lysis buffer (0.1% SDS and 0.5% Triton-X) for 10 minutes on ice to eliminate somatic cell contamination. Sperm sample purity was confirmed by microscopic examination of the samples. For sperm "reconstitutions" caput sperm were purified over a Percoll gradient as described in (*11*), washed with PBS, then either mock incubated at 37 °C or incubated with purified cauda epididymosomes at 37 °C as described in (*16*). After fusion reactions, sperm were pelleted, washed with PBS, pelleted again and snap frozen.

Testicular spermatocyte and spermatid isolation:

For each isolation, two testes were acquired from one FVB/NJ mouse at 6-10 weeks of age. Cell suspension was prepared by incubating the testes without their tunica albuginea in 5 ml elutriation buffer (100 mM NaCl, 45 mM KCl, 6 mM Na_2HPO_4, 0.6 mM KH_2PO_4, 0.23% Sodium DL-Lactate, 0.1% Glucose, 0.1% BSA, 0.011% Sodium Pyruvate, 1.2 mM $MgSO_4$ and 1.2 mM $CaCl_2$) containing 25 µg/ml liberase TM (Roche Diagnostics GmbH) for 30 min at 37 °C with gentle agitation every 5 min. The cell suspension was mixed by pipetting 20 times with a 10-ml plastic pipette. After homogenization by pipetting 10 times through a P1000 pipette, the single cell suspension was filtered twice through a 40-µm cell strainer (Fisher Scientific) on ice and centrifuged at 1500 rpm at 4 °C for 10 min, and then the pellet was resuspended with 20 ml elutriation buffer.

Separation of testis cell populations was performed by centrifugal elutriation using a JE-5.0 elutriation system and a 4-ml standard elutriation chamber (Beckman Coulter). The assembly of the system followed the manufacturer's instruction. The precise elutriation conditions are as follows: Fractions 1-3 were run at 3000 rpm and fractions 4-5 were run at 2000 rpm. The flow rate was 14, 18, 31, 23, and 40 ml/min for fractions 1-5, respectively. During elutriation, the elutriation chamber was maintained at 4 °C and the cells were collected into 50-ml conical polypropylene tubes that were packed in ice. Cells in tubes of fractions 3 to 5 were pelleted by centrifugation at 1500 rpm at 4 °C for 10 min. All pellets from the same fraction were combined and resuspended in 200 µl of elutriation buffer.

Percoll (Sigma-Aldrich) gradient (23-35%) was prepared by using a Gradient Master 108 (Biocomp) following standard program: S1/1, 2:26 (time), 82.0 (angle), 13 (rpm). After loading the cell suspension, centrifugation was carried out with SW40Ti Rotor (Beckman Coulter) at 11,000 rpm at 4 °C for 15 min. The cells were then collected in 15-ml conical polypropylene tubes and pelleted by centrifugation at 1500 rpm at 4 °C for 15 min. The cell pellets were stored at -80 °C and ready for RNA extraction. Based on cell morphology and small RNA data, Fraction 5 was determined to correspond to primary spermatocytes, with fractions 4 and 3 being relatively early and late round spermatids, respectively.

Epididymosome preparations:

Epididymosomes were prepared as previously described (*21*). In brief, gently extruded contents of the epididymis were centrifuged at 2000 × g to remove sperm, and the resulting supernatants were centrifuged at 10000 x g for 30 minutes to get rid of cellular debris. Next the supernatants were subjected to an ultracentrifugation at 120 000 × g at 4 °C for 2 h (TLA100.4 rotor; Beckman). Pellets were washed in cold PBS and

subjected to a second ultracentrifugation at 120 000 × g at 4 °C for 2 h. These pellets were then resuspended in 50 μl of PBS and used for RNA extraction. A small aliquot was used for performing electron microscopy and Nanosight analysis of the epididymosomes.

RNA Extraction:

For epididymosome RNA extraction, samples are thawed and the total volume of the sample was adjusted to 60 μl with filtered water (Macron). 33.3 μl of lysis buffer (6.4 M Guanidine HCl, 5% Tween 20, 5% Triton, 120 mM EDTA, and 120 mM Tris pH 8.0), 3.3 μl ProteinaseK (>600 mAU/ml, Qiagen 19131), and 3.3 μl water was then added to the sample. The sample was then incubated, with shaking, at 60 °C for 15 minutes on an Eppendorf thermomixer.

One volume of water (100 μl) was then added and the sample transferred to a phase lock column (5 PRIME). For phase separation 200 μl of TRI Reagent (MRC inc) and 40 μl BCP (1-bromo-2 chloropropane, MRC inc) were added. The samples were then mixed by inversion 10-15 times, followed by centrifugation at 14,000 RPM for 4 minutes (a second addition of TRI reagent and BCP can also be performed for more pure RNA). The aqueous phase was then removed and transferred to a low binding RNase/DNase free microcentrifuge tube (MSP), followed by the addition of 20 μg of glycoblue (Ambion) and 1 volume (~200 μl) of Isopropanol. The RNA was then precipitated for 30 minutes or greater at -20°C, followed by centrifugation at 14,000 RPM for 15 minutes at 4°C, and one wash with 70% cold ethanol followed by centriguation at 14,000 RPM for 5 minutes at 4°C. Finally, the RNA was reconstituted in water (10 μl for gel size selection for small RNA cloning).

For sperm RNA extraction, the same procedure was followed, apart from the addition of 3.3 μl 0.1M DTT with the lysis buffer and Proteinase K rather than the water. Also, prior to the incubation at 60 °C on the thermomixer the sperm pellet was disturbed physically using a pipette tip, and repeated pipetting. Finally, after the addition of the TRI reagent, sperm samples were vortexed for 5 minutes to ensure complete breakdown of the sperm. For testis and epididymidis samples, tissues were vortexed in 3-5 volumes TRI reagent with glass beads for 30 minutes at 4°C. The samples were then briefly spun and removed from the beads and BCP was added, followed by phase separation, and Isopropanol precipitation (as above).

Small RNA Cloning:

Small RNA cloning was carried out as in (*22*). Total RNA was combined with an equal volume of Gel Loading Buffer II (Ambion), loaded onto a 15% Polyacrylamide with 7M Urea and 1X TBE gel, and run at 15W in 1X TBE until the dye front was at the very bottom of the gel (~25 minutes for Criterion minigels). After staining with SYBR Gold (Life Technologies) for 7 minutes, and destaining in 1X TBE for 7 minutes, gel slices corresponding to 18-40 nucleotides were then cut from the gel. Gel slices were then ground (using a pipette tip or plastic pestle) and 750 μl of 0.3 M NaCl-TE pH 7.5 was added and incubated with shaking on a thermomixer overnight at room temperature. The samples were then filtered using a 0.4 μm Cellulose Acetate filter (Costar) to remove gel debris. The eluate was transferred to a new low binding microcentrifuge tube and 20 μg of glycoblue and 1 volume of Isopropanol (~700 μl) were added. Samples were precipitated for 30 or more minutes at -20 °C.

Size selection of the small RNAs was then followed by the ligation of a 3' adaptor and then a barcoded 5' adaptor as described in (*23*). The libraries were then converted to DNA using Superscript III (Invitrogen) and amplified by sequential rounds of PCR, to first add short primer tails and then longer primer tails, providing the products with the correct adaptor sequences for deep sequencing. Libraries were then sequenced by Illumina HiSeq 2000 at the Umass Deep Sequencing Core.

Normalization and data analysis:

For each small RNA library, rRNA-mapping reads (which were highly abundant in testis and epididymis samples, but rare in epididymosome and sperm samples) were removed. Remaining reads were mapped to murine tRNAs, to the unique sequences present in the 467 defined pachytene piRNA clusters (*24*), to Repeatmasker (tRNA entries from Repeatmasker were deleted to avoid duplicating tRNA-mapping reads), to miRbase, and to Refseq, yielding 27385 data points for each library – **Tables S1-S2**. Throughout the manuscript, piRNAs are defined as reads mapping to Repeatmasker along with reads mapping to the unique piRNA clusters. In this analysis pipeline, all reads mapping to a given entity – all fragments of the *Lcn5* mRNA, for example, or all the various size fragments of tRNA-Gly-GCC – are accumulated into a single data point. Importantly, in the case of multiple tRNA fragment species – 23 vs. 27 nt tRF-Gly GCC, for example – various length species from the same tRNA typically behaved concordantly in comparisons, justifying this grouping. Non rRNA-mapping reads were normalized to parts per million mapped reads for subsequent analyses.

For scatterplots, individual tRNA genes were grouped based on the uniqueness of the 5' 30 nucleotides – of the 14 tRNA-Gly-GCC genes in the mouse genome, 11 have indistinguishable 5' ends, so data for these 11 tRNA-Gly-GCC genes were merged for all scatterplots. This results in 190 unique 5' tRF points for most analyses. For a subset of analyses such as in **Fig. 3B** or **3D** we accumulated all reads for a given anticodon for the sake of visual clarity.

In **Fig. 1E**, individual samples with values more than 3 standard deviations away from the mean abundance of that RNA across 16 sperm samples were adjusted to 3 standard deviations away from sperm mean value prior to summing read counts for the scatterplot, to minimize the visual impact of a small number of RNA species with extreme outlier values. For scatterplots such as the cauda/caput ratios in **Fig. 3A**, only RNAs with a maximal abundance of 50 ppm in either cauda or caput samples, for each tissue, are shown. For heatmaps such as **Figs.1E,S5A**, only RNAs with a mean abundance >50 ppm are shown.

Estimating sperm RNA numbers:

To estimate absolute RNA numbers delivered by sperm, we sequenced five cauda sperm samples to which had been added a cocktail of small RNAs at defined molecules per sperm. Resulting estimates of small RNA abundance per sperm span a range of ~10-fold (~100,000 to ~1,000,000) across the five samples, and provide an estimated average count of ~200,000 small RNAs (with 3' ends untreated with PNK). This number is almost certainly an underestimate, as it is likely that RNA degradation occurred during sperm counting, and we cannot determine the efficiency of sperm lysis and RNA retrieval. Nonetheless, these data indicate that a minimum of ~10,000-50,000 molecules

of various glycine tRFs, and ~500-1,000 molecules of some individual let-7 family members, are delivered to the zygote by cauda sperm.

Intact tRNA sequencing library construction:

Acidic RNA extraction and enrichment of charged, uncharged and total tRNAs was performed based on protocol by (*25*). Mouse tissues were homogenized in a tube containing 1:1 volume of lysis buffer solution (0.3 M NaOAc/HOAc (pH 4.5) and 10 mM EDTA) and acetate-saturated phenol/$CHCl_3$ (pH 4.5). Following cold centrifugation at 18 600 x g, the aqueous layers were removed and subjected to another extraction with acetate-saturated phenol/$CHCl_3$ (pH 4.5) solution. The aqueous layers were removed, and the RNA was precipitated by adding 2.7 volumes of ethanol and centrifuging at 18 600 x g. RNA pellets were resuspended in lysis buffer solution and then precipitated with ethanol a second time. Finally, precipitated RNA was resuspended in a solution containing 10 mm NaOAc/HOAc (pH 4.5) and 1 mm EDTA. tRNAs were enriched by running the total RNA on a 10% denaturing urea- acrylamide gel in 100 mM sodium acetate buffer pH 4.5 and cutting out gel containing RNA between 60-80 nts. Next the tRNA samples were divided into three equal parts. One part was treated with periodate, which selectively oxidizes uncharged tRNA and blocks the subsequent ligation to 3' sequencing adaptor. Sodium periodate oxidation was performed by incubating total RNA with 50 mm $NaIO_4$ in 100 mM NaOAc/HOAc (pH 4.5). The samples were incubated at room temperature for 30 minutes and then precipitated with ethanol. This portion (charged tRNAs) and another portion of tRNAs (total tRNAs) was then deacylated. To deacylate, samples were resuspended in 50 mM Tris-HCl (pH 9) and incubated at 37 °C for 30 min followed by ethanol precipitated. Finally, all three portions were ligated to 3' sequencing adaptor. Next 5' adaptor was ligated, followed by reverse transcription and PCR amplification of cDNA as described previously (*26*). For PCR reactions we used primers with 6 nt indexes to allow sequencing of 8 different libraries on a single sequencing lane. The libraries were sequenced on Illumina HiSeq 2000 instrument at Umass Medical School Sequencing Core facility. Resulting tRNA sequences did not span the entire tRNA, typically only covering the 3' end of the tRNA, likely as a result of tRNA modifications blocking reverse transcriptase progression. The data nonetheless reveal many expected aspects of tRNA abundance and charging (see below), and relative levels of a given tRNA in response to diet should internally control for sequence-specific effects on tRNA abundance measured here. That said, we note that any dietary effects on "tRNA abundance" could also result from dietary effects on tRNA modifications that interfere with cloning.

Applying these protocols to budding yeast grown in replete medium, or subject to histidine starvation, recapitulated expected results, with histidine-starved yeast exhibiting a ~5-fold decrease in levels of charged tRNA-His as well as a 40% increase in levels of uncharged tRNA-His (**Fig. S2B**). We next applied these protocols to testis samples obtained from mice consuming our Control or Low Protein diets. Overall tRNA profiles were consistent between biological replicates, and revealed expected features such as a positive correlation between tRNA abundance and codon usage – total tRNA abundance in testis was correlated ($r = 0.50$) with genome-wide codon usage in mouse, and was even better-correlated ($r = 0.60$) with the codon abundances obtained from the most highly-expressed transcripts in testis (**Fig. S2C**) (*27*). Charged and uncharged tRNA levels were

generally well-correlated, albeit with ~10-fold variation between tRNAs in charging ratios (**Fig. S2D**), and revealed good specificity in our charging analysis – abundant snoRNAs (some of which are similar in length to intact tRNAs) were present in our total and uncharged datasets, but were depleted ~20-100 fold in the charged tRNA libraries.

Northern blot analysis:

Total RNA from testis, cauda epididymis and caput epididymis was mixed with an equal amount of formamide loading dye (95% formamide, 0.1% TBE, bromophenol blue and xylene cyanol) and heated to 65 °C for 5 minutes to denature the RNA. The RNA was then run on a 10% acrylamide 7M urea denaturing gel, followed by transfer to a positively charged nylon membrane and UV-crosslinking. The membrane was then hybridized with ^{32}P-labelled oligos specific to the 5' end of tRNA-Gly-GCC, tRNA-Val-CAC and 5S rRNA. After washing, the membrane was exposed to a phosphorimager screen which was later visualized using Typhoon imaging systems. The Sequences of oligos used is: tRNA-Gly-GCC 5' TCT ACC ACT GAA CCA CCA AT 3', tRNA-Val-CAC 5' AAC CAC TAC ACT ACG GAA AC 3' and 5S rRNA 5' GGG TGG TAT GGC CGT AGA C 3'.

TaqMan assays:

tRF and miRNA quantification was performed using custom designed TaqMan MicroRNA Assays according to manufacturer's recommended protocols (Applied Biosystems). 10 ng of total RNA isolated from mice as described above were reverse transcribed using the TaqMan MicroRNA reverse transcription kit. q-RT-PCR was performed in 15 μL reactions using TaqMan Universal PCR Master Mix, following standard program (10 min at 95 °C, then 15 sec at 95 °C and 1 min at 60 °C for 40 cycles). Serial dilutions of template were run to confirm amplification efficiency of all TaqMan probes.

Considerations regarding epididymosome adhesion to sperm:

The astonishing concordance between epididymosome and sperm RNA populations raised the concern that our sperm preparations could be contaminated by adhering epididymosomes. Three considerations argue against the second hypothesis. First, in preliminary troubleshooting experiments we carried out deep sequencing of four sperm samples that had not been extensively washed, revealing that epididymosome-specific RNAs are removed upon washing sperm (**Fig. S6F**). Second, comparing sperm and epididymosome RNA populations reveals numerous species strongly enriched in each class – sperm carry far more abundant piRNAs and fragments of mRNAs involved in spermatogenesis (such as *Prm1*) than do epididymosomes, whereas epididymosomes carry relatively greater levels of microRNAs such as let-7 family members (**Table S2**). Finally, Peng *et al* showed that fragments of tRNA-Gly-GCC and tRNA-Glu-CTC are resistant not only to treatment with lysis buffer, but even to mechanical removal of sperm tails (*6*). We thus conclude that small RNA changes from caput to cauda sperm either result from RNA degradation/gain in the sperm cell *per se*, or by "contamination" of sperm samples with a subpopulation of unusually detergent-resistant adhesive vesicles that carry a distinct RNA payload from that of the most abundant extracellular vesicles in

the surrounding fluid (and which, given their exceptional stability, might presumably be delivered to the oocyte at fertilization).

Bull samples:

Epididymides from sexually mature Holstein bulls were obtained from a commercial slaughterhouse. No dietary perturbations were performed for bulls. Tissues were stored on ice immediately after slaughter and processed within 3 h. Bovine spermatozoa from the caput epididymis were collected by intraluminal microperfusion as previously described (*21*). After a first centrifugation at 750 x g to remove the surrounding epididymal fluid, caput spermatozoa were washed three times in 150 mM NaCl at 4 °C. About 10 x 10^7 spermatozoa were incubated for 2.5 hours at 37 °C in 150 mM NaCl, 100 mM MES-PIPES, pH 6.5 in presence or in absence of 1.2 mg cauda epididymosomes. After extensive washes, sperm pellets were directly frozen in liquid nitrogen prior to RNA extraction and sequencing.

ES cell culture and transfection:

E14 ESC lines were grown in DMEM (Gibco), and transfections were carried out in in OptiMEM in 6 well plates (*28*), with 9.5cm^2 wells of ES cells seeded at a density of 3.5X10^5 cells/mL. 1 ng of antisense LNA containing oligonucleotides (Synthesized by Exiqon) were transfected using Lipofectamine 2000 (Invitrogen) for 16 hours, then ESCs were allowed to recover for 32 hours. Controls included lipofectamine only (Mock) and anti-GFP shRNA transfections. RNA extraction was performed at the end of 48 hours using standard Trizol protocol. RNA extracted from mouse ES cells was prepared for hybridization on Mouse GeneChip 2.0 ST arrays (Affymetrix) using the GeneChIP WT PLUS kit from Affymetrix.

Interfering with tRF-Gly-GCC function:

Two distinct oligonucleotides were used in this study to interfere with tRF-Gly-GCC function. For the majority of ES transfections, we used an LNA-containing antisense oligonucleotide with the following sequence: 5'-ACCACTGAACCACCAA-3'. We also noted that an unmodified antisense oligonucleotide – 5'-GCG AGA AUU CUA CCA CUG AAC CAC CAA UGC-3' – could derepress MERVL targets in ES cells (not shown). We therefore used this antisense in zygote microinjections, as the different lengths and ends of this oligo relative to the LNA make MERVL repression unlikely to be a gain of function of both antisense oligos. In addition, preliminary microinjections (n=3 embryos each) using the LNA antisense also resulted in upregulated MERVL target expression in 4-cell embryos (not shown), giving us confidence that both approaches to tRF-Gly-GCC inhibition work in both ES cells and intact embryos.

MERV-L reporter construction and FACS:

The *2C::tdTomato* construct described in (*17*) was obtained from Addgene, and contains the MERVL LTR nt 1-730 ligated into pcDNA3 hygro tdTomato plasmid. *Sp110::tdTomato* and *MERVK::tdTomato* were created from digested *2C:tdTomato* using the same restriction sites, and the inserts were made by specific PCR amplification of the MERV-L LTR upstream of *Sp110*, and a MERVK LTR element located 5' of *Tdpoz2*, using nested PCR primers to the genomic sequences. To derive stable cell lines, E14

mESCs were transfected with 2 μg of plasmid using Lipofectamine 2000 (Invitrogen), and selected using 150 μg/mL of hygromycin for 7 days. Surviving cells were re-plated into a 100 mm dish at low density, and tdTomato+ single colonies were picked and expanded, and confirmed using PCR. Stable cell lines were transfected with either LNA antisense to tRF-Gly-GCC (Exiqon) or anti-GFP esiRNA (described above), then sorted by FACS 48-hours post-transfection at the UMass Medical School Flow Cytometry Core, using a FACSAria II Cell Sorter (BD).

Pol2 chromatin immunoprecipitation:

Pol-II antibodies used: sc-899X (Santa Cruz, n=3 replicates) and ab5095 (Abcam, n=2 replicates). 2 μg of antibodies were conjugated to 50 μL of Dynadeads M-280 overnight with 0.1% BSA in PBS. $1X10^7$ mESCs transfected with anti-tRF-GG LNA-oligos and anti-GFP esiRNA were fixed in 1% formaldehyde for 5 mins at RT. Cells were lysed in 650 μL of nuclei lysis buffer (50 mM Tris-HCl pH 8, 10 mM EDTA pH 8, 1% SDS, 1X protease inhibitor cocktail) for 10 mins on ice. Chromatin was sheared on a Covaris S220 using the following settings: peak power 105, Duty Factor 2, 200 cycles, 300 seconds per 130 μL microtube. 130 μL of chromatin (~$2X10^6$ cells) was diluted with 900 μL of ChIP dilution buffer (50 mM Tris-HCl pH 8, 0.167 M NaCl, 1.1% Triton-X100, 0.11% sodium deoxycholate) and 500 μL RIPA-150 (50 mM Tris-HCl pH8, 0.15 M NaCl, 1 mM EDTA pH 8, 0.1% SDS, 1% Triton-X100, 0.1% sodium deoxycholate). An appropriate volume was also saved for input DNA, and extracted as below. Chromatin was added to Pol-II antibody conjugated Dynabeads and gently rotated for 4 hrs at 4 °C. Beads were washed once in RIPA-150, twice in RIPA-500 (50 mM Tris-HCl pH 8, 0.5 M NaCl, 1 mM EDTA pH 8, 0.1% SDS, 1% Triton-X100, 0.1% sodium deoxycholate), twice in RIPA-$LiCl_2$ (50 mM Tris-HCl pH 8, 1 mM EDTA pH 8, 1% NP40, 0.7% sodium deoxycholate, 0.5 M $LiCl_2$), and twice with TE buffer pH 8, 5 mins at 4 °C on rotator per wash. Beads were eluted with 200 μL elution buffer (10 mM Tris-HCl pH 8, 0.3 M NaCl, 5 mM EDTA pH 8, 0.5% SDS) with 1 μL RNase A (Qiagen) at 65 °C for 4hrs. Supernatant was removed from beads then incubated overnight at 55 °C with 1 μL Proteinase K (20 mg/mL). Finally, DNA was extracted using phenol-chloroform isoamyl alcohol, and precipitated with isopropanol and glycogen. DNA concentrations were determined using Qubit (Thermo-Fisher), then 0.5 ng of DNA was loaded for each qPCR reaction. Data in **Fig. S9C** were normalized to IgG control ChIPs, then the ratio was calculated for tRF-GG inhibition vs. GFP controls.

In vitro fertilization:

In vitro fertilization was performed according to "Manipulating the Mouse Embryo" Second Edition (*29*). FVB/NJ mice were used as egg donors and sperm was isolated from males fed dietary regimes as above. Fertilization took place in 250 μL HTF media covered in mineral oil, pre-gassed in 5% CO_2 at 37 °C.

Embryo RNA microinjection experiments:

Zygotes for microinjection studies were generated by IVF using sperm from mice fed Control diet. After four hours of IVF, the zygotes were washed three times in HTF medium and placed in a drop of KSOM medium at 37 °C in 5% CO_2 5% O_2 for 2 hours. Embryos were then transferred to FHM medium containing 0.1% PVA, and subjected to

micromanipulation. Embryos were microinjected with either H3.3-GFP mRNA alone (*30*) (control group) or H3.3-GFP mRNA plus one of several synthetic tRNA fragments (experimental group), or with H3.3-GFP mRNA plus gel-purified small RNAs (18-40 nts) isolated from Control vs. Low Protein sperm. RNA injections were carried out using a Femtojet (Eppendorf) microinjector at 100 hPa pressure for 0.2 seconds, with 7 hPa compensation pressure.

RNAs used for microinjections and their concentrations were: 100 ng/μl of H3.3-GFP mRNA, 200 ng/μl of tRF-Gly-GCC antisense RNA (5'GCG AGA AUU CUA CCA CUG AAC CAC CAA UGC 3'), 200 ng/μl of tRF-Glu-CTC sense RNA (5' UCC CUG GUG GUC UAG UGG UUA GGA UUC GG 3'), and 200 ng/μl of tRF-Gly-GCC sense RNA with modified residues (5' GCA JUL GUG GUU CAG UGG DAG AAU UCU CGC 3' where J=2'-O-methyluridine, L=N2-methylguanosine, D=dihydrouridine). For sperm small RNA microinjections, total RNA was extracted from either Control or Low Protein sperm, 18-40 nucleotides RNA was size selected on a gel and purified for microinjections, used at either 0.5 ng/μl (**Fig. 5D**) or 2 ng/μl concentration (**Fig. S11A**). After the microinjections, embryos were placed back into culture and H3.3-GFP fluorescence was verified at the 2-cell stage. GFP-positive injected embryos were cultured until the late 2-cell stage (32 hours post IVF for most injections, 28 hours for the tRF-Gly-GCC injections in **Fig. S11B**) or 4-cell stage (tRF-Gly-GCC antisense RNA microinjections, **Fig. 4E-F**), at which point embryos were collected and processed for single-embryo RNA sequencing.

Intracytoplasmic Sperm Injection (ICSI):

The testes and epididymis of 8-12 week old FVB/NJ mice were dissected into PBS. For rete sperm isolation, the efferent duct leading from the rete testis to the caput epididymis was located and the testis was removed under a stereo microscope and placed into a fresh dish containing PBS. After 2 washes with PBS, several cuts in the Rete testis were made, releasing its contents including the rete sperm. Released rete testis contents were transferred into an eppendorf tube, spun at 14,000 RPM for 2 minutes, and washed twice with modified nuclear isolation medium (NIM) with 1% polyvinyl alcohol. Sperm were finally resuspended in 500 μl NIM 1% PVA for use in ICSI.

For cauda sperm isolation, the cauda epididymis was isolated and placed in PBS. The cauda sperm were then released by making an incision in the cauda epididymis followed by squeezing to release the epididymal contents. The cauda sperm were then spun and washed as the Rete sperm.

For sperm heads (both rete and cauda), after collecting the sperm in PBS in an eppendorf tube, the sperm were spun at 14,000 RPM and washed once with PBS. The sperm were then resuspended in 500 μl PBS and then drawn through a 26G needle on a 1 ml syringe between 20-30 times. The shearing force from being drawn in and out of the needle removes the sperm head from tail for the majority of sperm. The sperm were then washed twice in NIM 1% PVA and finally resupended in 100-500 μl NIM 1% PVA for use in ICSI.

Females were superovulated by an intraperitoneal (i.p) injection of pregnant mare's serum gonadotropin (PMSG; 5 IU) followed by an i.p. injection of human chorionic gonadotropin (hCG; 5 IU) 48 hours later. Eggs were then collected from the oviducts of the females 13-16 hours later by placing the dissected ampulla of the oviduct

into KSOM containing 3 mg/ml hyaluranidase to digest the cumulus cells away from the eggs. After several minutes in hyaluranidase the eggs were washed 4-5 times in KSOM, finally being placed in KSOM in a 37 °C incubator until injected.

For ICSI, plates were made with drops of NIM 1% PVA for washing the injection needle, drops of NIM 1% PVA with sperm, drops of FHM with 0.1% PVA for the eggs to be added to for injection, and finally covered with mineral oil. 10-15 eggs at a time were placed into the FHM + 0.1% PVA drops on the injection plate for subsequent injection. Sperm (heads or whole) were then picked and injected into the eggs. After completion of the 10-15 injections, the injected eggs were maintained at room temperature for 5 minutes, washed 4 times in KSOM, and then placed in 50 μl KSOM in a 37 °C 5% O_2 incubator for development. The process was repeated for a total of 80-120 injections per day.

28 hours post-injection, 2-cell embryos were collected into 5 μl TCL buffer with 1% βME and then stored at -80°C for processing into single embryo RNA-sequencing libraries.

Single embryo RNA-Seq:

Single embryo RNA-Seq was carried out using the SMART-Seq protocol, as described in (*18, 19*). Data were mapped using RSEM, and were normalized to parts per million mapped reads after removing microRNA, snoRNA, and rRNA-mapping reads. Embryos with fewer than 10,000 detectable transcripts were removed from the dataset. For a given condition (Control, Low Protein, testicular sperm ICSI, etc.) we calculated the geometric mean of the mRNA abundance across all relevant embryos to minimize the influence of outliers, and filtered for genes with a geometric mean abundance >5 ppm. The cumulative distribution of the log2 fold change between different conditions is then plotted for various comparisons in **Figs. 5C-F**, **S11**, **S12**. For analysis of MERVL target gene expression, we used all genes changing at least 2-fold in our ES cell Affymetrix and RNA-Seq data (**Tables S4-S5**) for tRF-Gly-GCC targets, and we used the union of tables S2 and S6 from MacFarlan *et al* for remaining MERVL targets.

Immunofluorescence experiments and analysis of embryos:

To determine the effects of different diets on cell fate in the mouse embryo, IVF experiments were performed as described above, using sperm from sibling males either on Control or Low Protein diets. 4 hours after IVF, embryos were washed in HTF medium and placed in KSOM drops for long-term culture, in a 5% CO_2 5% O_2 incubator.

Embryos were collected at embryonic day 4 (E4) at mid-blastocyst stage and fixed in 4% PFA for Cdx2 staining. Immunofluorescence staining was performed as described in (*31*). Primary antibody used was Anti-CDX2 (BioGenex) at 1:100. Secondary antibody used was AlexaFluor 488 goat anti-mouse IgG (Molecular Probes) at 1:500. After the final wash, stained blastocysts were mounted in a gradient of Vectashield mounting medium with DAPI, in drops to retain the three-dimensional structure of the blastocyst. Microscopy was performed on a Zeiss Axiovert 200 inverted microscope with Orca-ER camera (Hamamatsu, NJ), using a 40x/1.4 NA oil objective. Z-sections were taken every 3 μm through the entire embryo. Analysis and cell counting was performed using Zeiss AxioVision4.9.1 software.

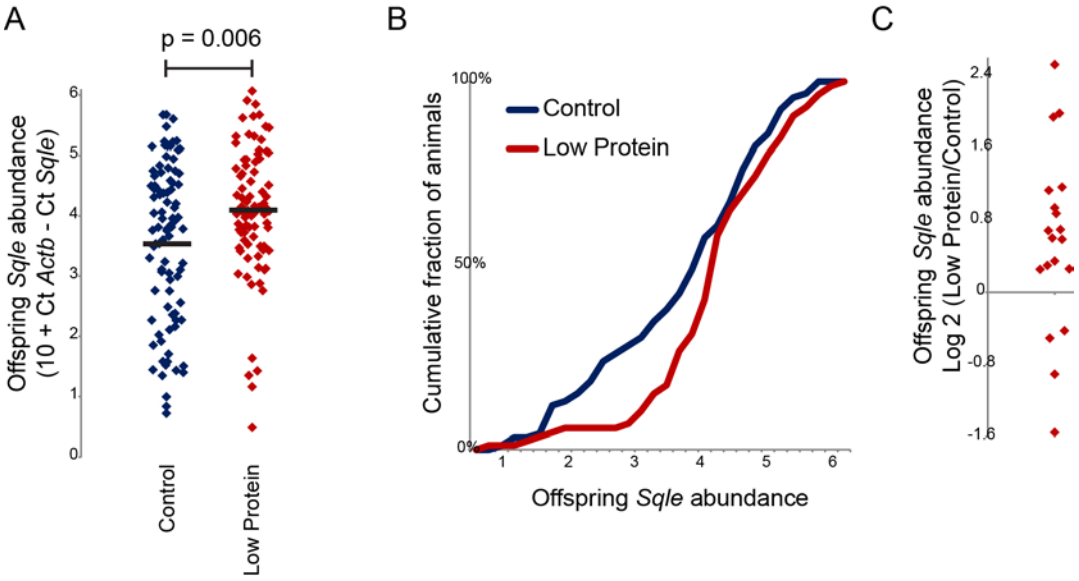

Fig. S1. Dietary information is carried in sperm.

(**A**) Sperm from males consuming Control or Low Protein diet was used to fertilize oocytes gathered from Control females. 2-cell stage embryos were then implanted into pseudopregnant females and allowed to develop to birth. At 3 weeks of age offspring were sacrificed (n=92 for Control, n=86 for Low Protein), and livers were harvested for analysis of *Sqle*, a gene previously shown to be upregulated in offspring of Low Protein males relative to Control males (*2*). *Sqle* levels (normalized to *Actb*) are shown for all offspring as individual points, with horizontal lines showing mean expression.

Note that the ability of paternal diet to alter offspring metabolism even in IVF-derived offspring is a conservative test of our hypothesis, as a single day of embryo culture has been reported to alter metabolism of offspring generated using cultured embryos (*32*), which could plausibly have overwhelmed any dietary effects in this system. The ability of sperm from Low Protein males to generate offspring with altered *Sqle* expression thus represents a conservative test of the hypothesis that sperm carry information capable of impacting offspring metabolism.

(**B**) Cumulative distribution of *Sqle* expression for all offspring generated using Control or Low Protein sperm, as indicated.

(**C**) Consistent litter effects. Here, *Sqle* levels were averaged for all offspring of a given litter. As sperm samples were always obtained from male siblings split to different diets, litter pairs resulting from paired fathers were compared, with each dot representing the ratio of *Sqle* expression between appropriately paired litters.

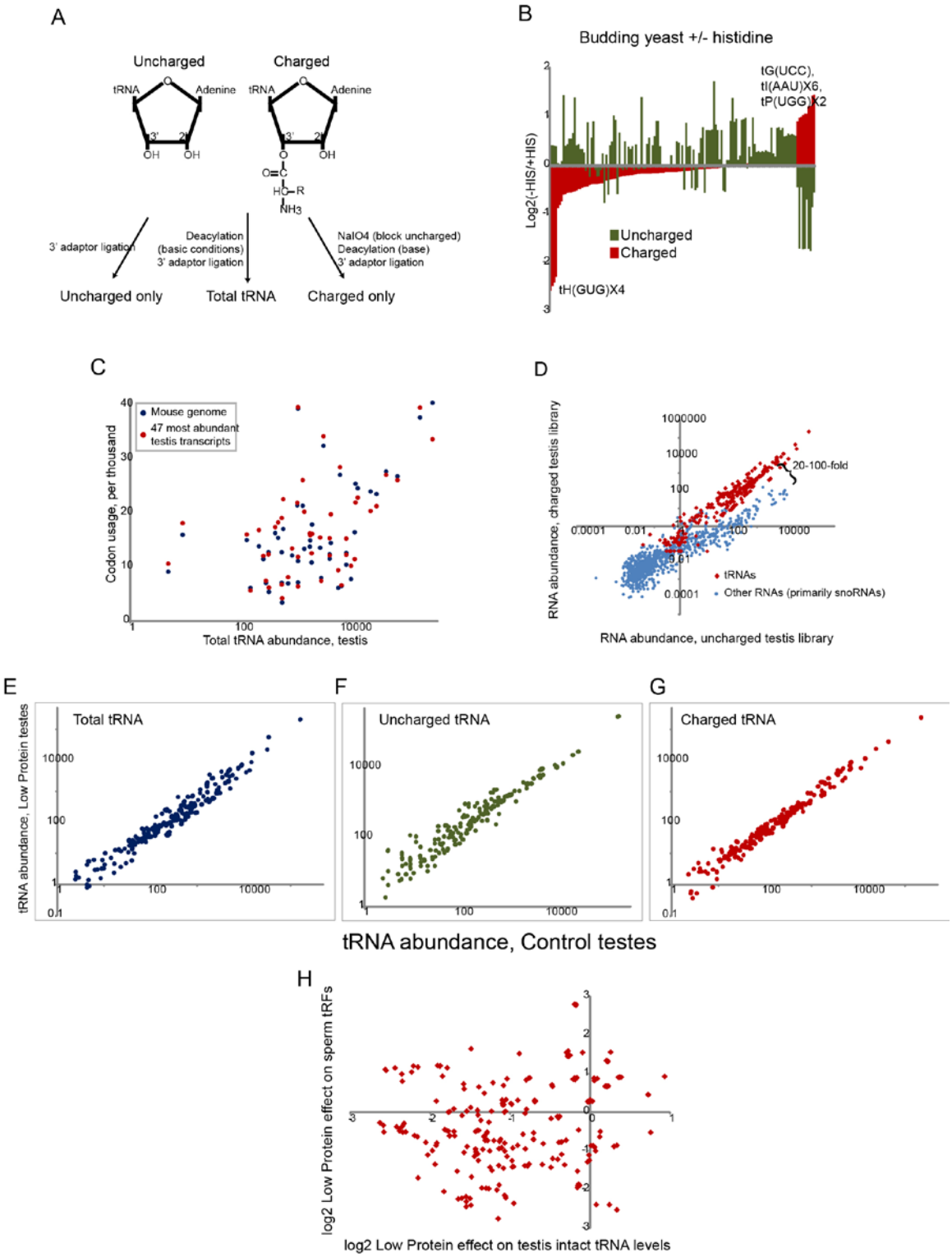
A
Uncharged
Charged
tRNA
Adenine
3'
2'
OH
NH3
3' adaptor ligation
Deacylation
(basic conditions)
3' adaptor ligation
NaIO4 (block uncharged)
Deacylation (base)
3' adaptor ligation
Uncharged only
Total tRNA
Charged only
B
Budding yeast +/- histidine
tG(UCC),
tI(AAU)X6,
tP(UGG)X2
Log2(-HIS/+HIS)
Uncharged
Charged
tH(GUG)X4
C
Mouse genome
47 most abundant
testis transcripts
Codon usage, per thousand
Total tRNA abundance, testis
D
RNA abundance, charged testis library
20-100-fold
tRNAs
Other RNAs (primarily snoRNAs)
RNA abundance, uncharged testis library
E
Total tRNA
tRNA abundance, Low Protein testes
F
Uncharged tRNA
G
Charged tRNA
tRNA abundance, Control testes
H
log2 Low Protein effect on sperm tRFs
log2 Low Protein effect on testis intact tRNA levels

Fig. S2 Dietary effects on tRNAs in testes.

(**A**) Schematic illustrating assay for tRNA charging analysis. RNA purified from a given tissue is isolated under acidic conditions to preserve charged tRNAs, and subject to the three treatments shown to enable deep sequencing characterization of charged, uncharged, and total tRNA levels.
(**B**) Validation of tRNA charging protocol. Budding yeast grown in the presence (+HIS) or absence (-HIS) of histidine were subjected to the tRNA analysis shown in (A). Changes in tRNA abundance for charged and uncharged tRNAs are shown on the y axis, sorted by the change in charged tRNA abundance. As expected charged tRNA-His levels drop dramatically after two hours of histidine starvation, while levels of uncharged tRNA-His increase.
(**C**) Testicular tRNA abundance correlates with codon bias in mouse. X axis shows intact tRNA abundance in testis (total tRNA is shown here but similar results hold for uncharged or charged tRNA datasets) in log scale, y axis shows the corresponding codon abundance (in codon frequency/1000) in all murine mRNAs, or in the 47 most-highly expressed mRNAs in testis. Data for testis mRNA abundance is from (*2*).
(**D**) Validation of tRNA charging analysis. Scatterplot shows abundance of ~60-80 nt RNAs in the total RNA protocol (x axis, log scale) compared to abundance of RNAs in the charged tRNA protocol (y axis, log scale). While tRNA levels are broadly consistent between the two protocols (charged/uncharged ratios do vary up to ~10-fold between individual tRNAs), other RNAs captured in the total RNA protocol, mostly snoRNAs (some of which are of similar size to tRNAs), are ~20-100 fold less abundant in the charged tRNA library, as expected.
(**E-G**) Low Protein vs. Control effects on tRNA levels for total (E), uncharged (F), and charged (G) tRNA levels in testis.
(**H**) Dietary effects on sperm tRFs are not explained by effects on intact tRNA abundance in testes. Log ratio between Control and Low Protein males is shown for total tRNA levels in testis (x axis) compared to tRNA fragment levels in cauda sperm (y axis).

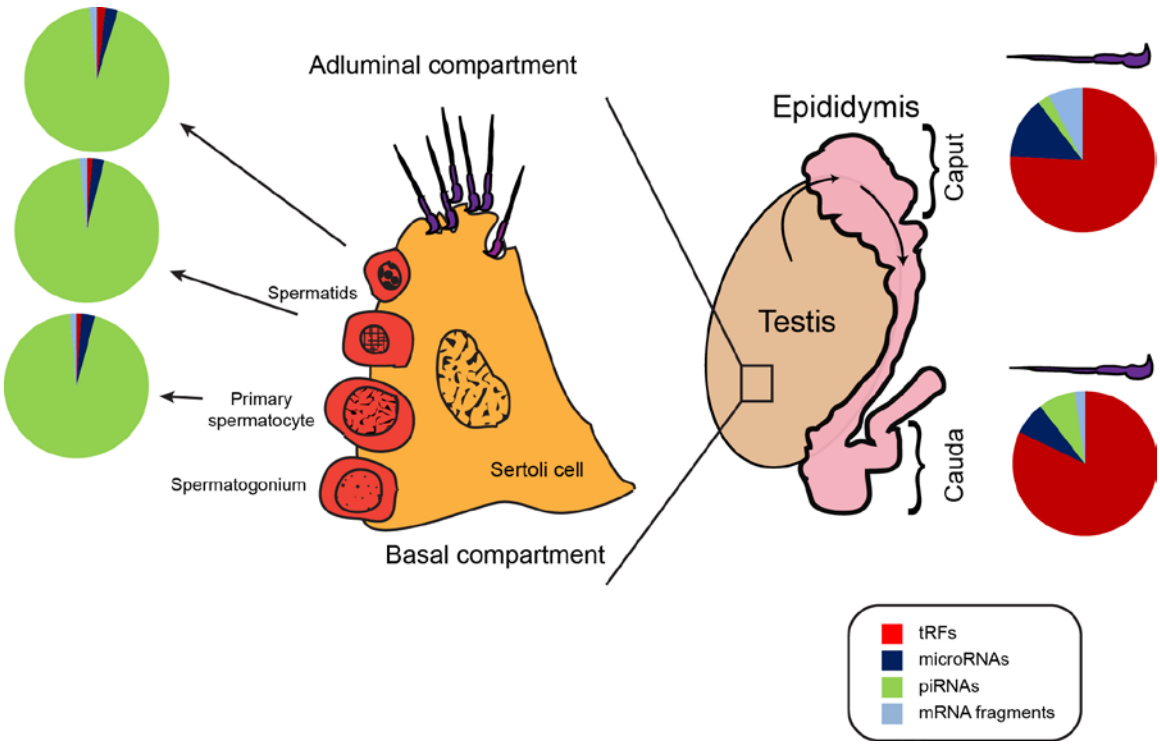

Fig. S3 RNA populations in testicular sperm populations.

Cartoon showing testicular spermatogenesis and post-testicular maturation in the epididymis. For each purified gamete population, pie charts show the relative abundance of tRNA fragments, microRNAs, piRNAs (defined as reads mapping to either repeatmasker consensus sequences or to unique piRNA clusters), and to Refseq (mRNA fragments), as indicated. Data are shown for three separate fractions of purified testicular germ cells, and for sperm isolated from caput and cauda epididymis, as indicated. Consistent with the low levels of tRNA fragments found in intact testes, we find that spermatocytes and two populations of post-meiotic spermatids carry extremely low levels of tRNA fragments, indicating that the absence of tRNA fragments in intact testis is not a result of contamination by testicular somatic cells. Data for caput sperm are described in more detail later in the manuscript, but are included here for comparison to cauda sperm and spermatid data.

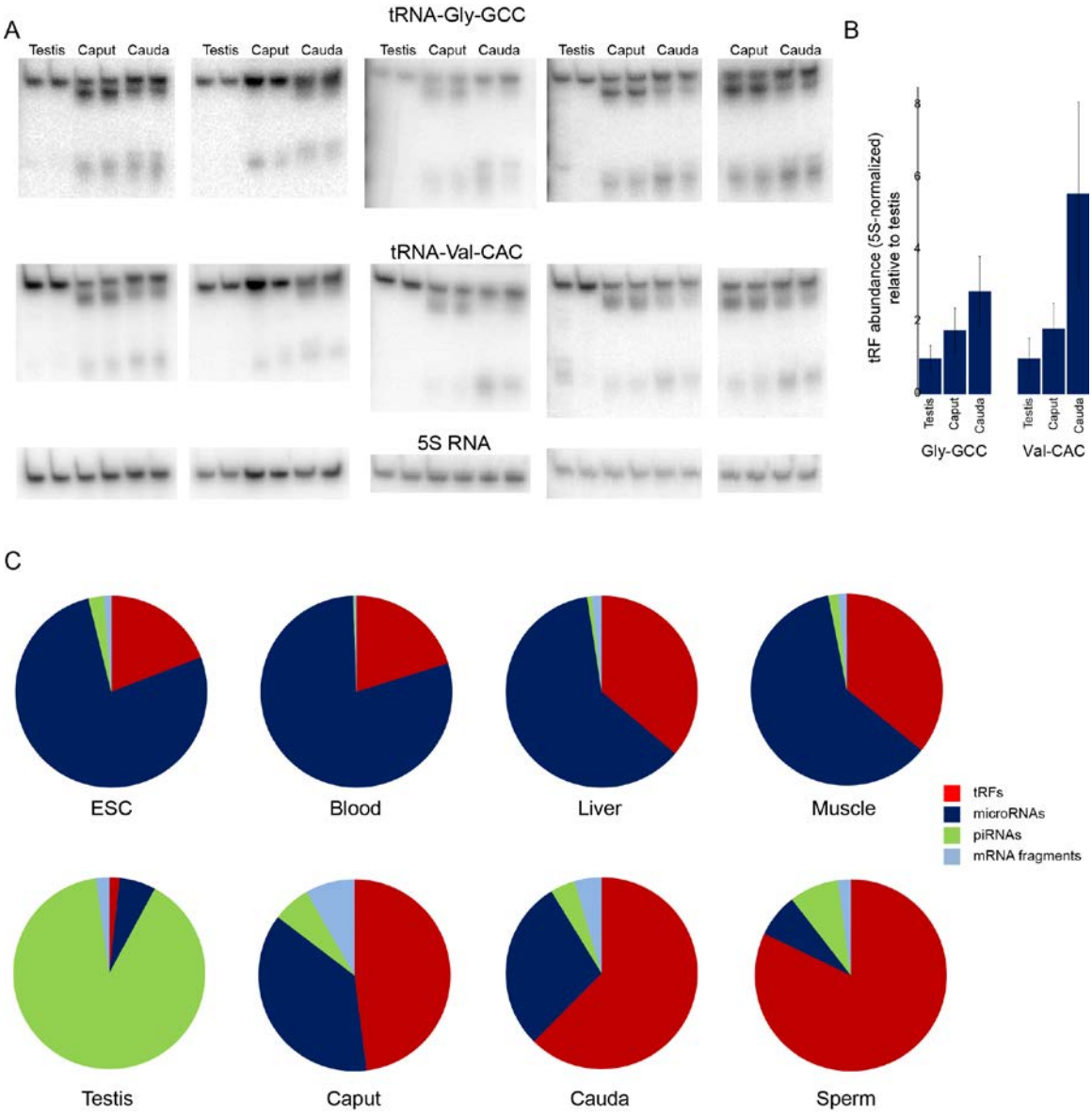

Fig. S4 tRNA fragments are abundant in the epididymis.

(**A**) Northern blot analysis of total RNA isolated from testis, cauda epididymis, and caput epididymis, as indicated. Each panel (except the rightmost panel, missing testis samples) shows paired samples in which each of the three tissues was obtained from the same animal. For each pair of samples, the left sample was isolated from a Control animal and right sample was isolated from a littermate consuming Low Protein diet. Note that the ~55-60 nt band observed in epididymis samples for tRNA-Gly-GCC and tRNA-Val-CAC varied somewhat in abundance between samples. This band almost certainly represents a T loop tRNA cleavage product rather than an intact tRNA differing in size from the ~75 nt tRNA by virtue of amino acid charging, as 1) it migrates at ~55-60 nt, which is too short to be an intact tRNA, 2) we observed this band both when using RNAs isolated

under acidic conditions and using RNAs isolated under more basic deacylating conditions, and 3) it is absent in testis samples.
(**B**) Quantitation of Northern blot data. For the indicated tRNAs, levels of the ~30 nt tRNA fragment were quantitated and normalized to 5S RNA abundance. Bars show levels of tRFs in testis, caput epididymis, and cauda epididymis, normalized to testis levels. Error bars show s.e.m.
(**C**) Pie charts showing the percentage of small RNAs mapping to the indicated features, for each tissue. rRNA-mapping reads are excluded. Here, piRNAs refer to all small RNAs mapping either to Repeatmasker or to unique piRNA clusters (*24*). Charts include those shown in **Fig. 2D**, along with charts for additional tissues. The increase in tRF abundance from caput to cauda epididymis is statistically significant (p=0.01, t-test).

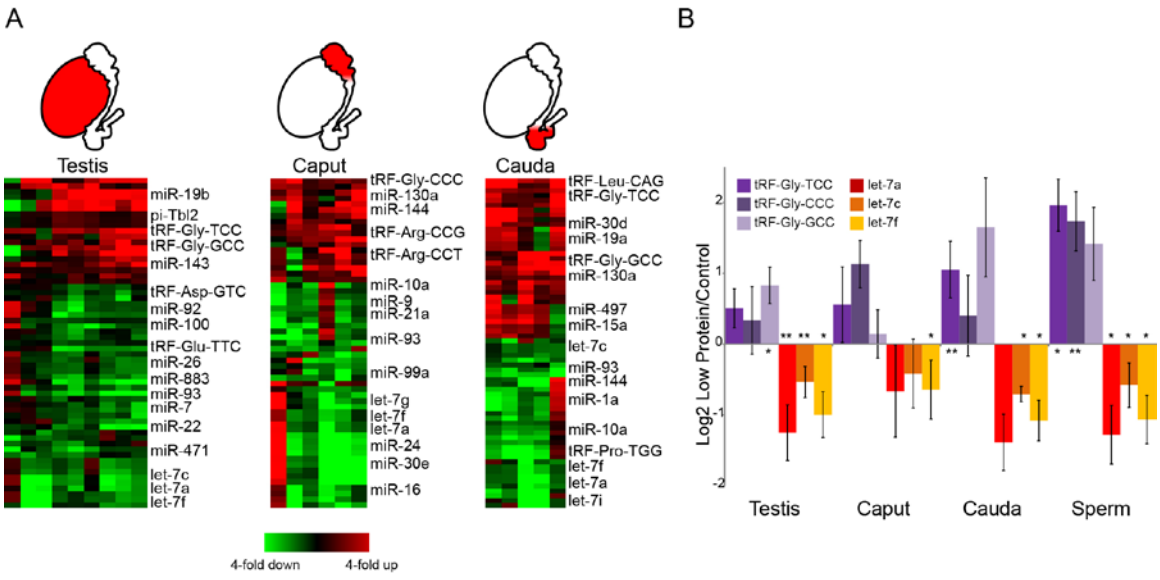

Fig. S5 Consistent dietary effects throughout the reproductive tract

(**A**) Dietary effects on small RNA abundance in testis and caput and cauda epididymis samples. Each heatmap shows log2 of Low Protein/Control RNA abundance for a pair of samples, showing RNAs (rows) that exhibit consistent dietary effects across >75% of samples. Low Protein diet caused significant changes in abundance of small RNAs throughout the male reproductive tract, with different tissues responding to dietary conditions with overlapping, but nonidentical sets of diet-regulated small RNAs.
(**B**) Coherent dietary effects on tRF-Gly and let-7 family members throughout the male reproductive tract. For each RNA, bars show average and standard error of the mean Low Protein effect on the RNA species in the indicated tissue. * and ** indicate paired t-test p values of <.05 and <0.01, respectively. Testis and cauda epididymis both showed significant dietary effects on various glycine tRFs and let-7 family members, and a similar trend was seen in caput epididymis, although caput samples showed greater sample-to-sample variation resulting from heterogeneity in dissections of this tissue.

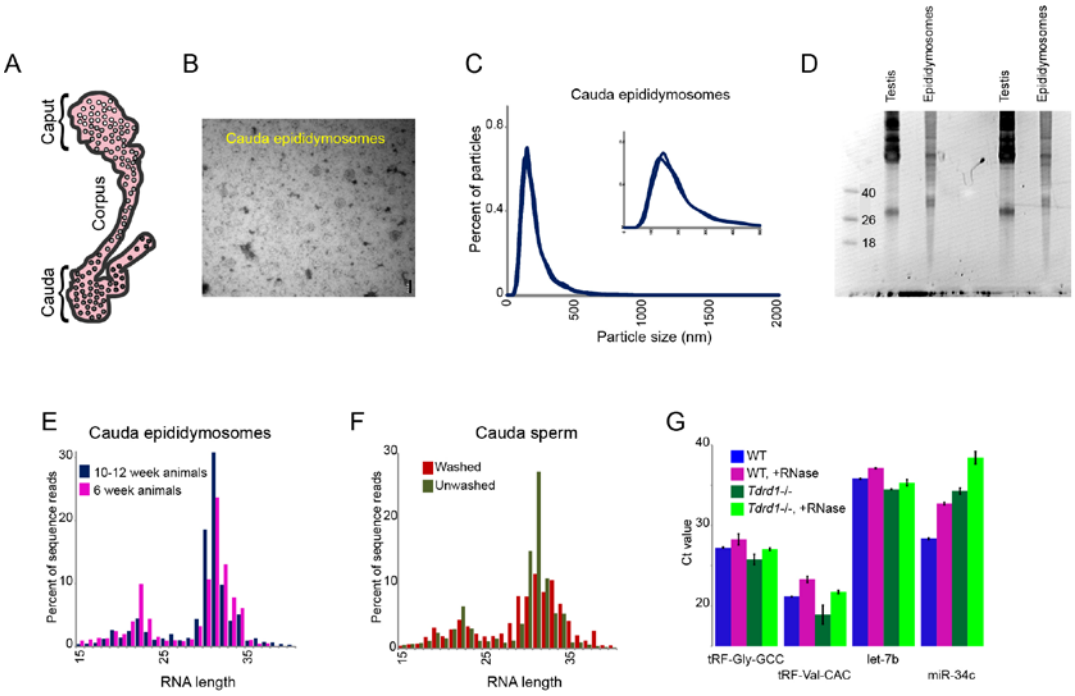

Fig. S6 Characterization of cauda epididymosomes.

(**A**) Schematic of murine epididymis. Circles represent epididymosomes.
(**B**) Transmission electron micrograph of purified cauda epididymosomes, showing abundant vesicles of ~120-150 nm.
(**C**) Epididymosome size distributions. Nanosight sizing data for two independent cauda epididymosome preps. Data for 0-2000 nm are shown in main panel, while inset shows 0-500 nm zoom-in.
(**D**) Gel electrophoresis (15% polyacrylamide 7 M urea gel) of total RNA isolated from whole testes, and from cauda epididymosomes, as indicated. Abundant ~26-30 nt piRNAs in testis are absent in epididymosomes, which instead carry highly abundant ~30-34 nt tRNA fragments.
(**E**) MicroRNAs decrease in abundance in cauda epididymosomes as animals age. Size distributions for deep sequencing libraries are shown for averaged data from 4 epididymosome preparations from 6 week old animals, and from 11 preparations from 10-12 week old animals, as indicated. Younger animals tend to carry slightly higher levels of microRNAs in cauda epididymosomes.
(**F**) Unwashed sperm are contaminated with RNAs that are abundant in epididymosomes. Deep sequencing RNA size distributions are shown for 9 averaged datasets obtained using 3X-washed cauda sperm, and from 2 averaged datasets using sperm subject to swim-up but flash-frozen after first pelleting. In addition to the more homogeneous size of tRFs observed in unwashed sperm, we also find much higher levels of epididymosome-enriched RNAs, such as let-7 family members, in the unwashed dataset. These data show that washing sperm with detergent is essential for removing contaminating epididymosomes prior to analysis of the sperm RNA payload.
(**G**) Characterization of epididymosomal RNAs. For each of the 4 small RNAs indicated, TaqMan was used to quantitate the level of each RNA from 1 μg of RNA from purified

cauda epididymosomes. Epididymosomes were isolated either from wild type animals or from *Tdrd1*-/- mutants that lack mature sperm, and were either mock treated or treated with RNaseA prior to RNA extraction, as indicated. RNaseA treatment completely eliminated detectable small RNAs when carried out *after* RNA extraction (not shown). Y axis shows Ct value for the average of three TaqMan assays, with error bar showing standard deviation. All RNAs were present at similar levels in both genetic backgrounds. For tRF-Gly-GCC, tRF-Val-CAC, and let-7b, RNaseA treatment typically lead to a ~2-fold loss of the RNA of interest, indicating that at least half of the RNA is protected from nuclease treatment, presumably by virtue of being inside a vesicle. Curiously, mir-34c was substantially more RNase-sensitive, suggesting that this RNA co-purifies with epididymosomes but is located outside of the vesicle.

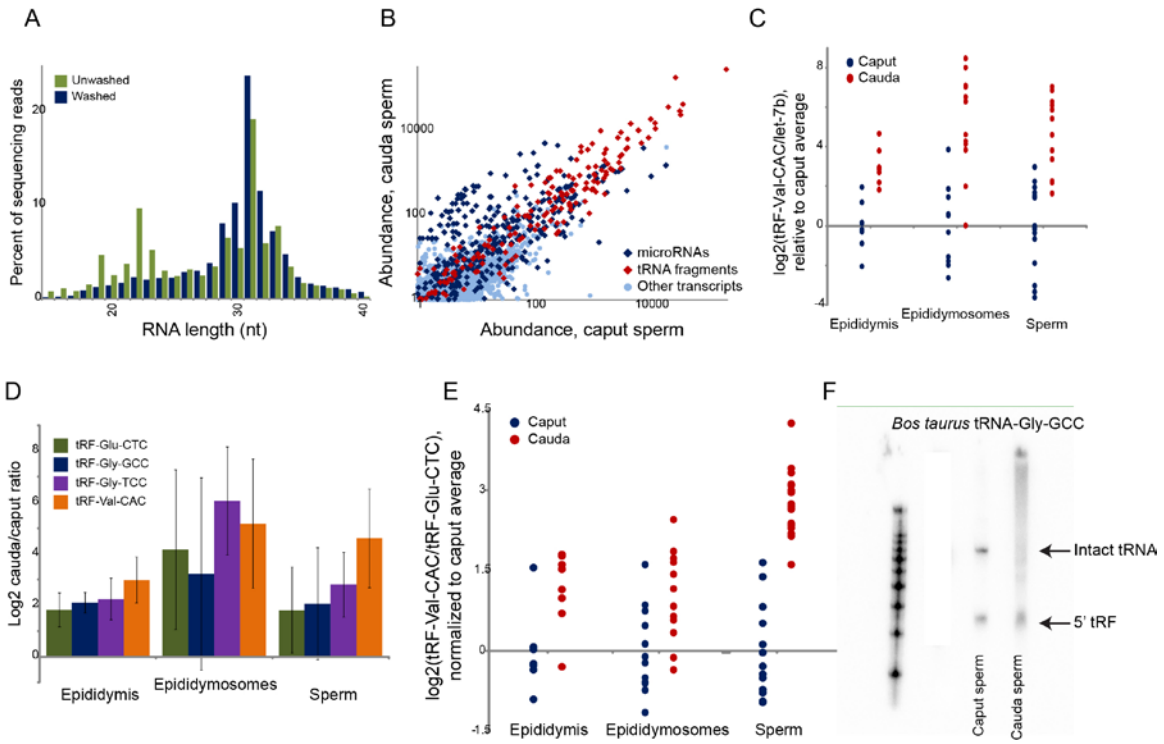

Fig. S7 RNA populations in caput sperm.

(**A**) Unwashed caput sperm are contaminated with RNAs abundant in epididymosomes. As in **Fig. S6F**, but for caput sperm isolated with and without washing with epithelial cell lysis buffer. RNA isolated from unwashed caput sperm included numerous microRNAs that were most abundant in epididymosomes, and which were lost in washed caput sperm, demonstrating that our lysis buffer wash effectively removes contaminating epididymosomes from sperm. Thus, the correlations between the caput/cauda ratios for small RNA abundance in epididymis, and the corresponding caput/cauda ratios in sperm, are unlikely to reflect artifactual contamination of sperm preparations by epididymosomes. The abundant tRFs present in caput sperm indicate that the dramatic increase in tRF abundance in sperm relative to spermatids (**Fig. S3**) occurs either during the first few days of epididymal transit, or very late during testicular spermatogenesis.
(**B**) Comparison of small RNA payloads of cauda vs. caput sperm for all RNA species with an abundance of at least 1 ppm in both sperm populations. These changes in RNA abundance could result from extant RNAs from caput sperm being degraded during further transit through the epididymis, or from small RNAs being gained via processing or trafficking during post-testicular maturation. Many aspects of these data are consistent with recent reports (*14*), as for example cauda sperm exhibit apparent gain of microRNAs such as the X-linked miR-880 cluster, and apparent decrease in let-7 levels.
(**C**) tRF-Val-CAC levels increase in the distal reproductive tract. tRF-Val-CAC and let-7b levels were measured by TaqMan, and tRF-Val-CAC level was normalized to let-7b. For each tissue type, the average ΔΔCt value for all caput samples was subtracted from each individual caput or cauda sample, revealing a general gain of tRF-Val-CAC (or loss

of let-7b, or both) in cauda relative to caput samples (t-test $p < 10^{-4}$ for all three tissue types).
(**D**) Gain in four tRFs from caput to cauda. Taqman data for the four tRFs indicated was normalized to let-7b as in (C), and here the average cauda/caput difference for each tissue is shown plus/minus the standard deviation. Similar results are obtained using miR-21 as a normalized control. In addition, nearly-identical results were obtained using TaqMan assays for the 23, 27, or 29 nt variants of tRF-Gly-GCC (only data for 27 nt is shown). These data are consistent either with a general gain of tRFs from caput to cauda samples of all three tissue types – epididymis, epididymosomes, and sperm – or loss of let-7 or miR-21.
(**E**) tRF-Val-CAC is strongly cauda-enriched. Data from (D) are shown with tRF-Val-CAC normalized to tRF-Glu-CTC rather than to let-7 (which decreases in relative abundance from caput to cauda). This shows that the apparent gain of tRF-Val-CAC in cauda samples is not a normalization artifact due to the distal decrease in let-7 levels.
(**F**) Caput sperm carry intact tRNAs. Northern blots were performed against the 5' end of tRNA-Gly-GCC for samples of bull caput sperm and bull cauda sperm. Note that a gel lane not pertinent to this study has been hidden using a white box. The presence of intact tRNAs in caput sperm therefore does not rule out the null hypothesis for tRF biogenesis in sperm – that tRNA fragments in mature cauda sperm are generated from tRNAs synthesized during testicular spermatogenesis – and thus the tissue of origin (epididymis, spermatogonial stem cell, etc.) for tRNA fragments found in cauda sperm can only be definitively assigned using transgenic approaches using tissue-specific expression of "tracer" tRNA molecules, or similar approaches.

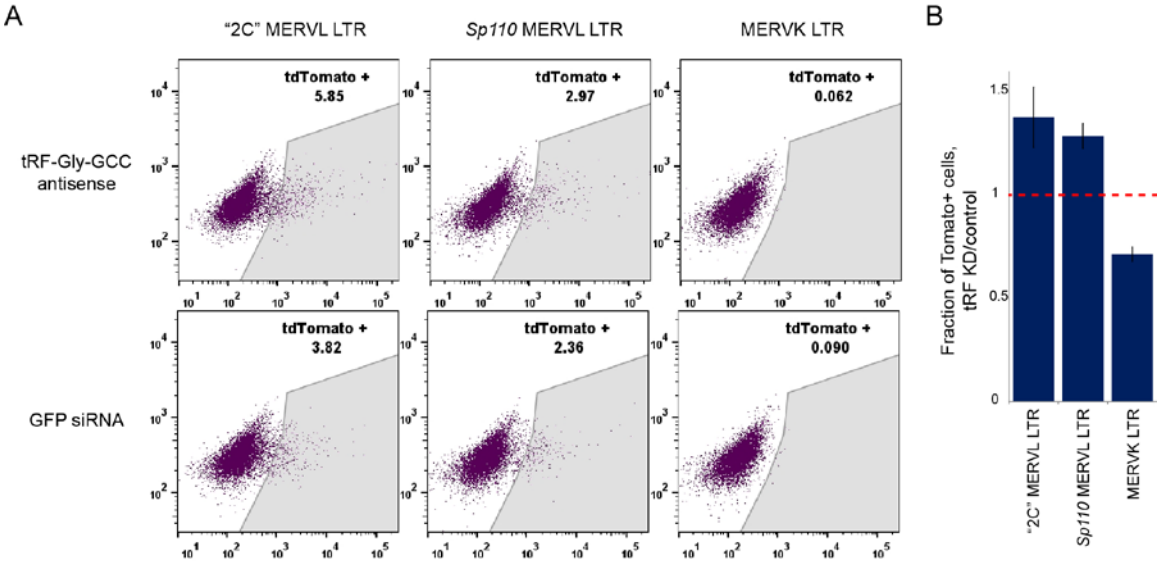

Fig. S8 tRF-Gly-GCC regulates a MERVL reporter construct.

(**A**) Fluorescent reporter data for stable ES cell lines carrying tdTomato driven by the "2C" MERVL LTR reported in Macfarlan *et al*, by the MERVL LTR located upstream of *Sp110*, or by a MERVK LTR. Data for one replicate transfection experiment is shown, with gated reporter-positive cell counts indicated above each plot.
(**B**) Average change in percentage of Tomato-positive cells between GFP KD and tRF-Gly-GCC inhibition is shown for the three reporters. Each bar shows mean and standard deviation from three replicate experiments.

Fig. S9

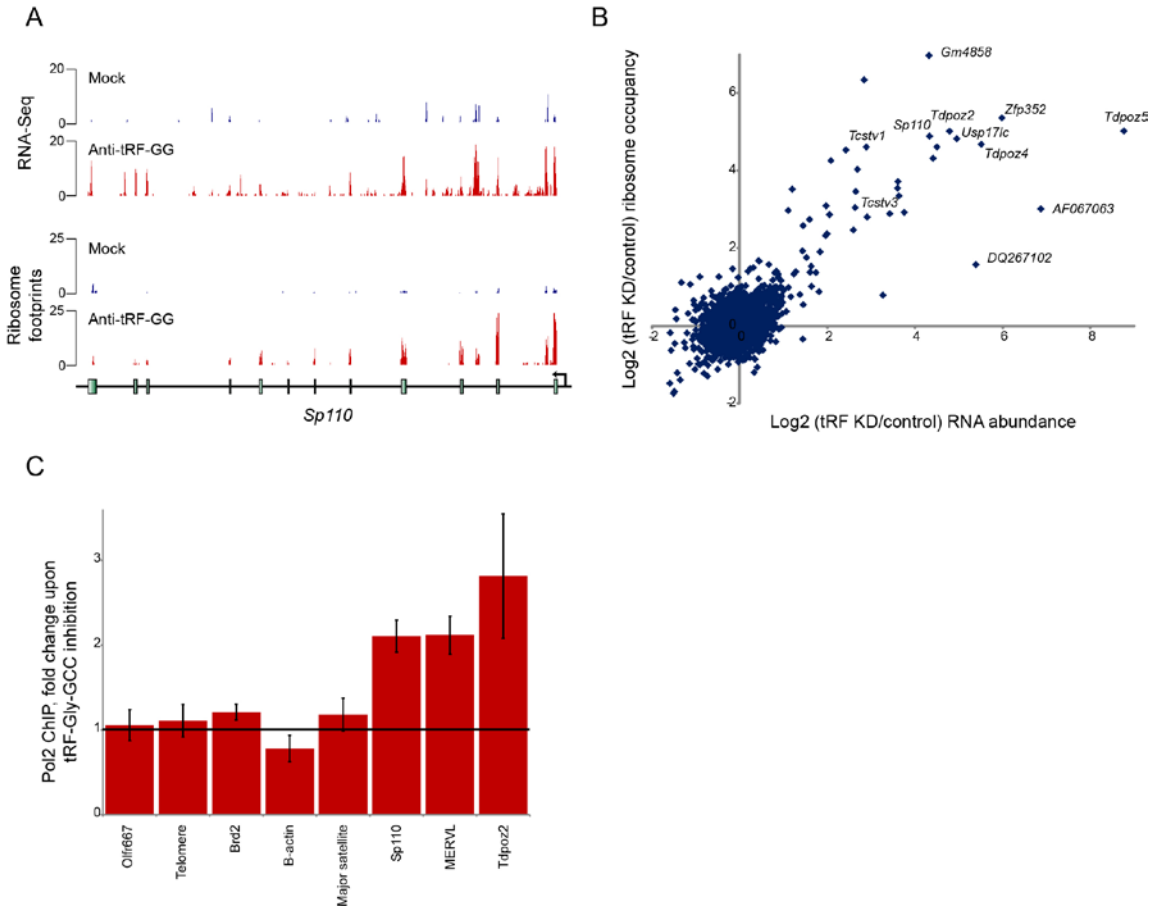

Fig. S9 Mechanistic basis for tRF-Gly-GCC regulation of MERVL targets.

(**A**) tRF-Gly-GCC effects on MERVL targets are unlikely to result from secondary effects of translational regulation of MERVL regulators. RNA-Seq and ribosome footprinting (*33*) were carried out for 4 replicate ES cell cultures either mock transfected or transfected with the LNA antisense oligonucleotide targeting tRF-Gly-GCC, as in **Fig. 4**. As all four replicates were well-correlated, we show merged data for all replicates here. Merged data for RNA-Seq and ribosome footprinting are shown for tRF-Gly-GCC target *Sp110*, showing 1) that MERVL targets are translated in ES cells, and 2) that ribosome footprinting data lack intronic and UTR reads, as expected.
(**B**) Scatterplot of effects of tRF-Gly-GCC inhibition on mRNA abundance and on ribosome footprints. For each measurement, the log2 of the fold change upon tRF-Gly-GCC inhibition is plotted, as indicated. In general, upregulated mRNAs are also associated with ribosomes, as expected. More importantly, we find no substantial effects of tRF-Gly-GCC inhibition specifically on ribosome footprinting data, arguing that MERVL target mRNA changes are not secondary effects of tRF-Gly-GCC regulation of translation of some MERVL regulator such as the CAF complex or Rex1.

(**C**) tRF-Gly-GCC inhibition affects MERVL target genes at the level of transcription. RNA Pol2 ChIP was carried out in ES cells subject to transfection with siRNAs targeting GFP, or with the LNA antisense to tRF-Gly-GCC. Q-PCR was carried out using the indicated primers, and the effect of tRF-Gly-GCC inhibition on Pol2 ChIP signal for each locus is shown. Data show average and s.e.m. for five biological replicates, revealing increased Pol2 signal specifically over the three MERVL targets (*Sp110*, MERVL consensus, and *Tdpoz2*).

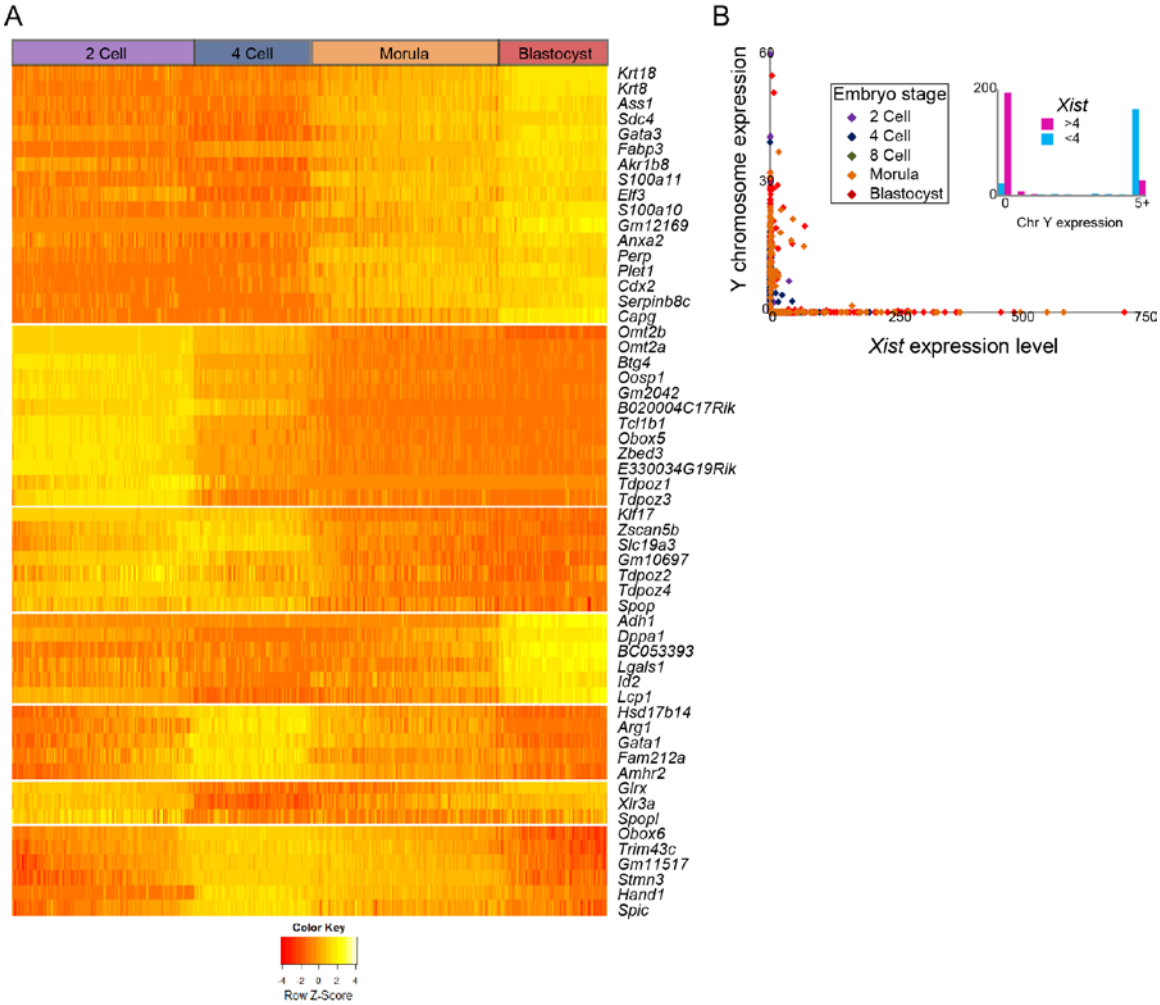

Fig. S10 Gene expression dynamics during preimplantation development.

(**A**) Heatmap of the top 53 genes that contribute to Principal Components 1 and 2 of single embryo RNA-Seq dataset. Embryos are ordered left to right from early development according to radial angle in the 2-dimensional PCA space.
(**B**) Anticorrelation between *Xist* expression and expression of Y-linked genes allows robust sexing of embryos. Data for all 606 embryos is shown as a scatterplot, with embryo stages colored as indicated. Inset shows data for embryos (excluding 2-cell stage embryos, many of which have not yet activated *Xist* expression) separated into low *Xist* expression (TPM < 4) or high *Xist* expression, showing average Y-linked gene expression for these classes.

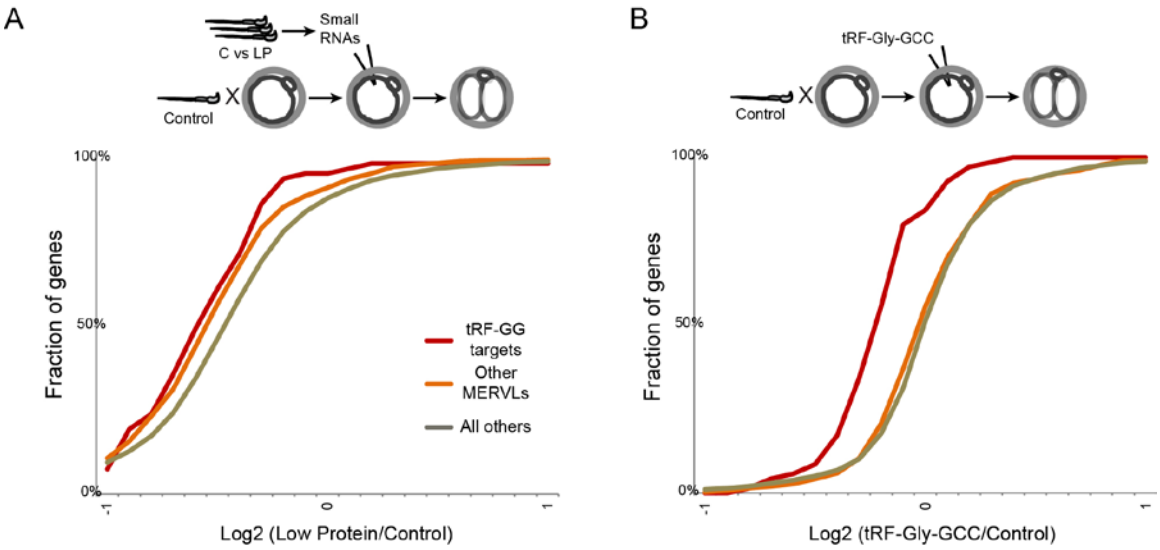

Fig. S11 Effects of small RNAs on 2-cell stage gene expression.

(**A**) Cumulative distribution plot for ES cell MERVL targets, additional MERVL targets defined in (*17*), and all remaining genes (filtered for those with abundance > 5 ppm). X axis shows the log2 fold change between 2-cell embryos which had been injected with 2 ng/µl small RNAs purified from either Control or Low Protein sperm samples (n=23 C, 21 LP embryos). As in **Fig. 5D**, but for an independent experiment using 2 ng/µl rather than 0.5 ng/µl. Here again MERVL targets are significantly repressed in LP-injected embryos relative to C-injected embryos – the lack of increased repressive activity relative to the lower concentration injections likely reflects experimental noise but also could result from the use of different Control/Low Protein sperm samples, effects of oocyte donors, etc.
(**B**) As in (A), but comparing embryos injected with synthetic tRF-Gly-GCC vs. GFP only controls (n=19 and 26 embryos, respectively). This is an independent replicate of **Fig. 5E**, but here embryos were collected 28 hours after injection rather than 32 hours. tRF-Gly-GCC inhibition of MERVL expression is similar at both time points.

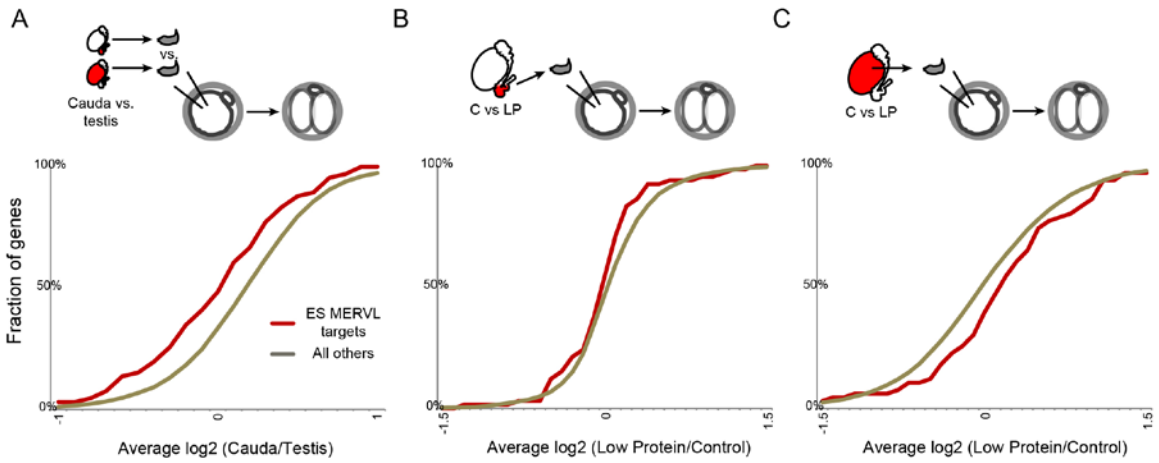

Fig. S12 Analysis of ICSI embryos.

(**A-C**) Cumulative distribution plots for MERVL target genes and remaining genes in 2-cell stage embryos generated via intracytoplasmic sperm injections (ICSI) using sperm heads from either testicular sperm or cauda sperm. (A) compares embryos generated from testicular vs. cauda sperm heads (n=45, 37), (B) compares embryos generated using cauda sperm heads from either Control or Low Protein animals (n= 20, 17), and (C) compares testicular sperm heads from Control or Low Protein animals (n= 22, 23). Comparison between cauda and testicular sperm reveals a modest repression of MERVL targets in cauda-derived embryos (p = 0.02, KS test), consistent with the expectation that sperm gain tRF-Gly-GCC during epididymal transit (although we have not sequenced small RNAs from mature testicular sperm so it is plausible though unlikely that sperm gain tRFs between the round spermatid stage and their arrival at the rete testis). However, there were surprisingly modest gene expression differences between cauda and testis ICSI embryos, given the major differences in sperm RNA payload (**Fig. S3**). Moreover, the effect of paternal diet on MERVL was not significantly recapitulated through ICSI using cauda sperm heads (B). One potential explanation for this would be if ICSI disrupts the ability of sperm-borne small RNAs to function in the zygote. Interestingly, the sperm midpiece is one of the major sites of epididymosomal fusion in vitro (*34*), and is also the location of the remnants of the chromatoid body, which plays key roles in RNA processing and storage in early sperm. This motivated us to carry out ICSI using intact sperm (**Fig. 5F**), although this proved exceptionally technically challenging (21 of ~500 injections survived to the 2-cell stage) and of course is an extraordinarily nonphysiological means of fertilization. Nonetheless, embryos generated using intact cauda and testicular sperm showed many more changes in mRNA abundance (**Fig. 5F** – note the widened x axis) than the equivalent comparison for sperm heads (A), including more robust downregulation of MERVL targets in cauda-derived embryos (KS test p=0.001 and 1.5 X 10^{-12} for tRF-Gly-GCC targets and all other MERVL targets, respectively). MERVL targets were also downregulated in intact sperm ICSI embryos relative to sperm head ICSI embryos (not shown), which is consistent with the hypothesis that intact sperm carry more tRF-Gly-GCC than isolated sperm heads.

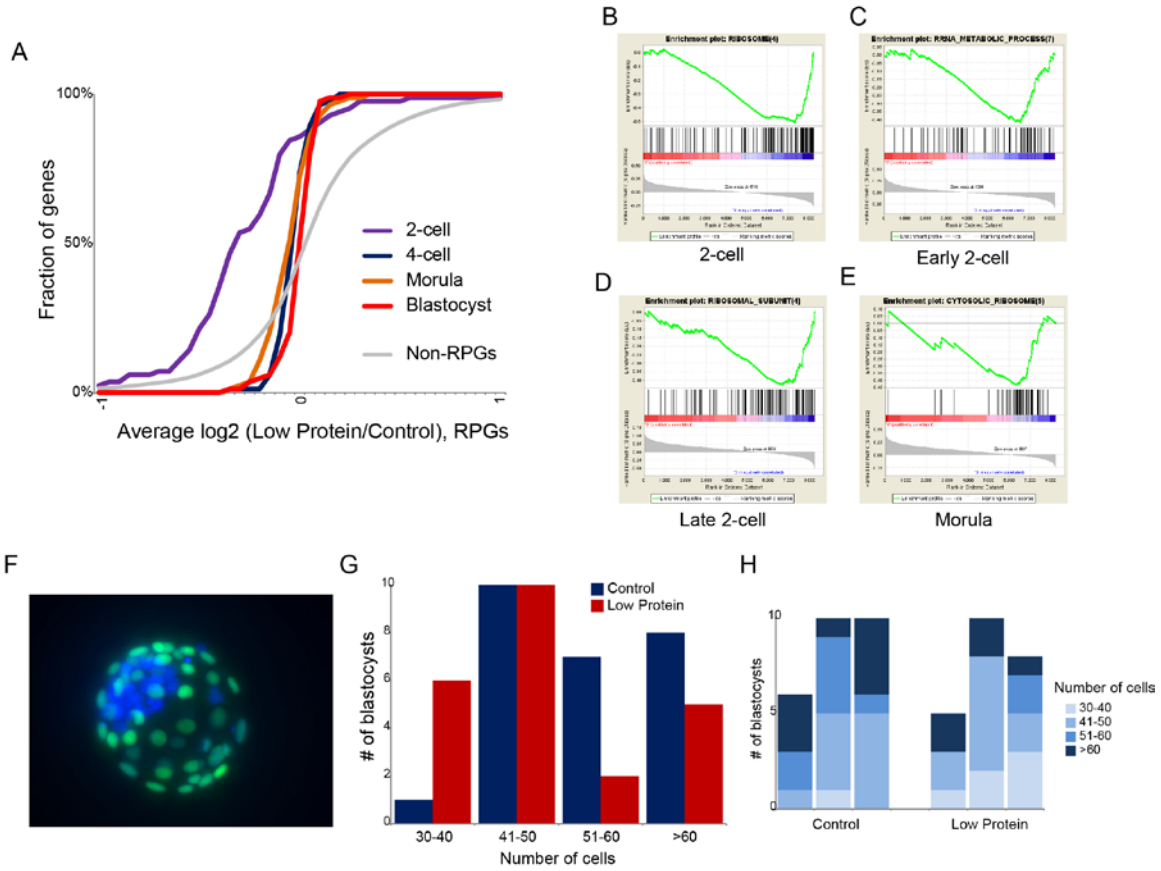

Fig. S13 Dietary effects on preimplantation developmental temp.

(**A**) Dietary effects on ribosomal protein gene expression at four embryonic stages. Cumulative distribution for Low Protein effect on all ribosomal protein genes at the indicated stages. Grey line shows distribution of dietary effects on all non-RPG genes, for all four stages.
(**B-E**) Geneset enrichment plots for Control vs. Low Protein embryo RNA-Seq. Genesets shown are ribosome (B), rRNA metabolic process (C), ribosomal subunits (D), and cytosolic ribosome (E). Embryo stages are indicated – early and late 2-cell embryos were defined based on the second principal component (**Fig. 5B**). Repression of ribosomal biogenesis genes at both stages suggests that ribosomal protein regulation differences between Control and Low Protein embryos is unlikely to simply be a consequence of delayed zygotic genome activation.
(**F**) Blastocyst stained with DAPI (blue) and anti-CDX2 (green) to image total cell number and trophectoderm cells.
(**G**) Low Protein diet reproducibly alters developmental tempo. For embryos generated via IVF using Control or Low Protein sperm, plot shows the number of blastocysts with the indicated number of cells. Dietary effects on blastocyst size have been previously observed in other dietary paradigms (*20*, *35*).
(**H**) Individual replicates of the cell counting data in (G). Reduced cell numbers were observed in Low Protein embryos in three independent IVF replicates, as shown.

Additional Data table S1 (separate file) Complete small RNA dataset.

All murine small RNA-Seq datasets. Rows give read counts mapping to rRNAs, tRNAs, repeatmasker consensus sequences, unique piRNA clusters, and microRNAs are all provided here. Reads mapping to Refseq are excluded in the interest of space, but are provided for tissue averages in **Table S2**.

Additional Data table S2 (separate file) Small RNA abundance for all tissues.

Small RNA data for all tissues are shown here normalized to ppm (excluding rRNAs). For each tissue, total reads were obtained by summing reads from each relevant library.

Additional Data table S3 (separate file) Small RNA-Seq of *B. taurus* caput sperm reconstitutions.

Data for four replicate sperm reconstitutions are shown, with counts of RNAs mapping to rRNAs, tRNAs, or microRNAs shown as indicated. Samples include purified caput sperm, mock-treated caput sperm (incubated at 37 °C), caput sperm incubated with cauda epididymosomes, cauda epididymosomes alone, and a single sample of cauda sperm.

Additional Data table S4 (separate file) ES cell mRNA abundance.

Affymetrix microarray data are shown for ES cells transfected with GFP knockdown siRNAs, or with an LNA antisense targeting the 5' end of tRNA-Gly-GCC.

Additional Data table S5 (separate file) tRF-Gly-GCC effects on translation in ES cells.

RNA-Seq data and ribosome footprinting for ES cells subject to mock tranfection, GFP knockdown, or tRF-Gly-GCC inhibition (4 replicates each).

Additional Data table S6 (separate file) tRF-Gly-GCC effects on 4-cell stage gene expression.

Single embryo RNA-Seq data for 4-cell stage embryos. Zygotes generated via IVF were microinjected with H3.3-GFP mRNA, with or without an antisense oligo targeting tRF-Gly-GCC, then allowed to develop to the 4-cell stage and subject to single embryo RNA-Seq.

Additional Data table S7 (separate file) Dietary effects on preimplantation gene regulation.

Single-embryo RNA-Seq data for embryos at varying stages of development, generated using IVF with sperm obtained from male mice consuming the indicated diets.

Additional Data table S8 (separate file) RNA effects on preimplantation gene regulation.

Single-embryo RNA-Seq data for ICSI experiments, sperm small RNA injections, and tRF-Gly-GCC injections. All data are for late 2-cell stage embryos, and data are only shown for transcripts expressed at greater than 5 parts per million in a given experiment.

21. C. Belleannee, E. Calvo, J. Caballero, R. Sullivan, Epididymosomes convey different repertoires of microRNAs throughout the bovine epididymis. *Biology of reproduction* **89**, 30 (2013)
22. W. Gu, H. C. Lee, D. Chaves, E. M. Youngman, G. J. Pazour, D. Conte, Jr., C. C. Mello, CapSeq and CIP-TAP identify Pol II start sites and reveal capped small RNAs as C. elegans piRNA precursors. *Cell* **151**, 1488-1500 (2013)
23. W. Gu, M. Shirayama, D. Conte, Jr., J. Vasale, P. J. Batista, J. M. Claycomb, J. J. Moresco, E. M. Youngman, J. Keys, M. J. Stoltz, C. C. Chen, D. A. Chaves, S. Duan, K. D. Kasschau, N. Fahlgren, J. R. Yates, 3rd, S. Mitani, J. C. Carrington, C. C. Mello, Distinct argonaute-mediated 22G-RNA pathways direct genome surveillance in the C. elegans germline. *Molecular cell* **36**, 231-244 (2009); published online EpubOct 23 (10.1016/j.molcel.2009.09.020).
24. X. Z. Li, C. K. Roy, X. Dong, E. Bolcun-Filas, J. Wang, B. W. Han, J. Xu, M. J. Moore, J. C. Schimenti, Z. Weng, P. D. Zamore, An ancient transcription factor initiates the burst of piRNA production during early meiosis in mouse testes. *Molecular cell* **50**, 67-81 (2013)
25. J. M. Zaborske, J. Narasimhan, L. Jiang, S. A. Wek, K. A. Dittmar, F. Freimoser, T. Pan, R. C. Wek, Genome-wide analysis of tRNA charging and activation of the eIF2 kinase Gcn2p. *The Journal of biological chemistry* **284**, 25254-25267 (2009)
26. M. Ghildiyal, H. Seitz, M. D. Horwich, C. Li, T. Du, S. Lee, J. Xu, E. L. Kittler, M. L. Zapp, Z. Weng, P. D. Zamore, Endogenous siRNAs derived from transposons and mRNAs in Drosophila somatic cells. *Science (New York, N.Y* **320**, 1077-1081 (2008)
27. K. A. Dittmar, J. M. Goodenbour, T. Pan, Tissue-specific differences in human transfer RNA expression. *PLoS genetics* **2**, e221 (2006)
28. T. G. Fazzio, J. T. Huff, B. Panning, An RNAi screen of chromatin proteins identifies Tip60-p400 as a regulator of embryonic stem cell identity. *Cell* **134**, 162-174 (2008)
29. A. Nagy, *Manipulating the mouse embryo : a laboratory manual*. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., ed. 3rd, 2003), pp. x, 764 p.
30. A. Santenard, C. Ziegler-Birling, M. Koch, L. Tora, A. J. Bannister, M. E. Torres-Padilla, Heterochromatin formation in the mouse embryo requires critical residues of the histone variant H3.3. *Nature cell biology* **12**, 853-862 (2010); published online EpubSep (10.1038/ncb2089).
31. M. E. Torres-Padilla, A. J. Bannister, P. J. Hurd, T. Kouzarides, M. Zernicka-Goetz, Dynamic distribution of the replacement histone variant H3.3 in the mouse oocyte and preimplantation embryos. *The International journal of developmental biology* **50**, 455-461 (2006)10.1387/ijdb.052073mt).
32. S. K. Feuer, X. Liu, A. Donjacour, W. Lin, R. K. Simbulan, G. Giritharan, L. D. Piane, K. Kolahi, K. Ameri, E. Maltepe, P. F. Rinaudo, Use of a mouse in vitro fertilization model to understand the developmental origins of health and disease hypothesis. *Endocrinology* **155**, 1956-1969 (2014); published online EpubMay (10.1210/en.2013-2081).
33. N. T. Ingolia, S. Ghaemmaghami, J. R. Newman, J. S. Weissman, Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling.

Science (New York, N.Y **324**, 218-223 (2009); published online EpubApr 10 (10.1126/science.1168978).

34. G. S. Griffiths, D. S. Galileo, K. Reese, P. A. Martin-Deleon, Investigating the role of murine epididymosomes and uterosomes in GPI-linked protein transfer to sperm using SPAM1 as a model. *Molecular reproduction and development* **75**, 1627-1636 (2008); published online EpubNov (10.1002/mrd.20907).

35. N. O. McPherson, H. W. Bakos, J. A. Owens, B. P. Setchell, M. Lane, Improving metabolic health in obese male mice via diet and exercise restores embryo development and fetal growth. *PloS one* **8**, e71459 (2013)10.1371/journal.pone.0071459).