MHR *Basic science of reproductive medicine*

ORIGINAL RESEARCH

MicroRNA transcriptome in the newborn mouse ovaries determined by massive parallel sequencing

Hyo Won Ahn[1], Ryan D. Morin[2], Han Zhao[3], Ronald A. Harris[4], Cristian Coarfa[4], Zi-Jiang Chen[3], Aleksandar Milosavljevic[4], Marco A. Marra[2], and Aleksandar Rajkovic[1,5,*]

[1]Department of Obstetrics, Gynecology and Reproductive Sciences, Magee-Womens Research Institute, University of Pittsburgh, Pittsburgh, PA, USA [2]Genome Sciences Centre, BC Cancer Agency, Vancouver, British Columbia, Canada [3]Center for Reproductive Medicine, Shandong Provincial Hospital, Shandong University, Shandong, People's Republic of China [4]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA [5]Department of Obstetrics and Gynecology, Baylor College of Medicine, Houston, TX, USA

*Correspondence address. Department of Obstetrics, Gynecology and Reproductive Sciences, Magee-Womens Research Institute, University of Pittsburgh, Pittsburgh, PA, USA. Tel: +1-412-641-8635; Fax: +1-412-641-6156; E-mail: rajkovic@upmc.edu

ABSTRACT: Small non-coding RNAs, such as microRNAs (miRNAs), are involved in diverse biological processes including organ development and tissue differentiation. Global disruption of miRNA biogenesis in Dicer knockout mice disrupts early embryogenesis and primordial germ cell formation. However, the role of miRNAs in early folliculogenesis is poorly understood. In order to identify a full transcriptome set of small RNAs expressed in the newborn (NB) ovary, we extracted small RNA fraction from mouse NB ovary tissues and subjected it to massive parallel sequencing using the Genome Analyzer from Illumina. Massive sequencing produced 4 655 992 reads of 33 bp each representing a total of 154 Mbp of sequence data. The Pash alignment algorithm mapped 50.13% of the reads to the mouse genome. Sequence reads were clustered based on overlapping mapping coordinates and intersected with known miRNAs, small nucleolar RNAs (snoRNAs), piwi-interacting RNA (piRNA) clusters and repetitive genomic regions; 25.2% of the reads mapped to known miRNAs, 25.5% to genomic repeats, 3.5% to piRNAs and 0.18% to snoRNAs. Three hundred and ninety-eight known miRNA species were among the sequenced small RNAs, and 118 isomiR sequences that are not in the miRBase database. Let-7 family was the most abundantly expressed miRNA, and mmu-mir-672, mmu-mir-322, mmu-mir-503 and mmu-mir-465 families are the most abundant X-linked miRNA detected. X-linked mmu-mir-503, mmu-mir-672 and mmu-mir-465 family showed preferential expression in testes and ovaries. We also identified four novel miRNAs that are preferentially expressed in gonads. Gonadal selective miRNAs may play important roles in ovarian development, folliculogenesis and female fertility.

Key words: miRNA / ovary / oocyte / microRNA / ncRNA

Introduction

Gene expression can be modified at many levels, including transcriptional, posttranscriptional and translational. Many genes, however, do not result in a functional protein and instead encode for what are known as non-coding RNAs. Such non-coding RNAs include ribosomal RNAs (rRNAs), transfer RNAs (tRNAs), small nuclear RNAs (snRNAs), small nucleolar RNAs (snoRNAs), polyadenylated non-coding RNAs, microRNAs (miRNA), piwi-interacting RNAs (piRNAs) and RNA species yet to be discovered. miRNAs have received extraordinary attention due to their likely role in widespread regulation of mRNA metabolism, at both the transcriptional and the post-transcriptional levels (Doench and Sharp, 2004; Carthew and

Sontheimer, 2009). There are currently 579 known mouse miRNAs (miRBase v14.0) and 721 human miRNAs (miRBase v14.0; Griffiths-Jones et al., 2006). miRNAs are derived from hairpin containing pre-miRNAs after endoribonuclease processing by Drosha (also known as Rnasen) and Dicer enzymes (Meister and Tuschl, 2004). miRNAs are thought to regulate gene expression by either repressing or blocking translational mechanism by base-paring with the target mRNA, usually in the 3′-untranslated (UTR) region (Olsen and Ambros, 1999), although translational activation has also been reported (Filipowicz et al., 2008).

miRNA expression profiles differ between tissues and developmental stages. To better understand roles of miRNAs in the ovary, several studies have assessed expression profile of miRNAs in the ovary using

different technologies. Ro et al. (2007) discovered 122 miRNAs, including 15 novel miRNAs, in the adult mouse ovary by sequencing 800 cDNAs corresponding to the small RNA fraction. Mishima et al. (2008) sequenced 11 744 cDNAs derived from small RNAs in the adult ovary and found that 154 known miRNAs are expressed in the adult ovary and discovered 1 novel miRNA. In a separate study, the miRNA microarray analysis of periovulatory granulosa cells identified a total of 212 mature miRNAs (Fiedler et al., 2008). mir-132 and 212 were highly up-regulated in the granulosa cells following LH/hCG induction. All three of the previous studies focused on the adult ovary, where most of the cell types are somatic in origin, with various stages of folliculogenesis throughout the ovary. The only study to assess miRNA expression in the developing ovary, used miRNA microarrays to assess the effect of germ-cell-specific transcription factor *Nobox* deficiency on miRNA expression in the newborn (NB) ovary (Choi et al., 2007). A total of 177 known miRNAs were shown to be expressed in the NB ovary, with little apparent effect of *Nobox* deficiency on their expression (Choi et al., 2007). Microarray analysis is limited to known miRNAs, with little information gained regarding other small RNAs or novel RNAs.

A critical event in the mammalian female gonadal development is the development of the finite pool of primordial follicles that will be used for reproduction and hormonal homeostasis. Mouse female germ cells, following mitotic proliferation in the gonadal ridge, exist as germ cell clusters and enter meiosis I circa embryonic day 13.5, and by the time of birth, arrest in the diplotene stage of the first meiotic division (Wilhelm et al., 2007). Germ cell clusters formed in the embryonic gonad break down shortly after birth with resulting large loss of oocytes and formation of primordial follicles (Pepling and Spradling, 2001; Choi and Rajkovic, 2006). The primordial follicle endowment, as well as the subsequent rate of primordial ovarian follicle loss, likely determines reproductive life span (Pangas and Rajkovic, 2006). We and others have previously shown that multiple oocyte-specific genes such as *Nobox*, *Figla*, *Sohlh1*, *Sohlh2*, *Lhx8* and *Mater* are expressed in the NB ovaries and that these genes are essential at various stages of ovarian development and early embryogenesis (Soyal et al., 2000; Ballow et al., 2006; Choi et al., 2008). NB ovary is an important developmental time point in the mouse, when germ cell clusters breakdown and primordial follicles form. We therefore hypothesized that novel miRNAs preferentially expressed in the NB ovary exist and may have important roles in ovarian development and early folliculogenesis. We utilized massive parallel sequencing on the Genome Analyzer Platform from Illumina (San Diego, CA, USA) to determine small RNA NB ovary transcriptome (Morozova et al., 2009).

Materials and Methods

Small RNA library construction and sequencing

Mouse C57BL/6/129S6/SvEv strain was used to harvest tissues. All experimental and surgical procedures complied with the Guide for the Care and Use of Laboratory Animals and were approved by the Institutional Agricultural Animal Care and Use of Committee of Baylor College of Medicine. Ovaries from 54 wild-type female NB (within 24 h of delivery) mice were isolated and stored in RNAlater solution (Ambion, Austin, TX, USA) for RNA extraction. Total RNA was extracted using mirVana™ miRNA isolation kit (Ambion) following the manufacturer's total RNA isolation procedure. Ten micrograms of total RNA was size fractionated on 15% tris-borate-EDTA (TBE) urea polyacrylamide gel and small RNA fraction in the range 18–30 nucleotides was extracted. 5′ and 3′ RNA adapters were ligated to the gel extracted small RNA fraction, and cDNA was generated according to the manufacturer (Illumina). The resulting cDNA was amplified and sequenced using the Illumina's Genome Analyzer.

Small RNA genome mapping

Small RNA genome mapping and quantification was performed as described previously (Morin et al., 2008) as well as by the Pash alignment algorithm (Kalafus et al., 2004; Coarfa and Milosavljevic, 2008). The Pash algorithm was used to map the reads onto the mouse genome assembly (NCBI Build 37, mm9) and mapping parameters were guided by both the number of reads mapped as a measure of sensitivity and the number of mappings to known RNAs as a measure of specificity. Read mappings were clustered based on overlapping mapping coordinates and intersected with known miRNAs (http://genome.ucsc.edu/cgi-bin/hgTrackUi?g=miRNA), snoRNAs (http://gene.fudan.sh.cn/snoRNAbase.nsf), piRNA Bank (Sai Lakshmi and Agrawal, 2008) and repeats (http://genome.ucsc.edu/cgi-bin/hgTrackUi?g=rmsk).

Novel miRNA prediction

Read clusters not intersecting with known RNAs and consisting of at least 100 reads were identified as putative novel miRNAs. The cutoff of 100 reads was chosen arbitrary to enrich for more abundant small RNAs. The mouse putative novel RNAs were mapped to human and rat using Blast (Altschul et al., 1997), UCSC Alignment Nets and UCSC Alignment Chains (Kent et al., 2003) for isolating conserved putative small RNAs across three species. miR-abela (Sewer et al., 2005) was used to identify putative small RNAs with miRNA hairpin structures.

Novel miRNA prediction was approached as described previously (Morin et al., 2008). In essence, this method uses RNALfold (Hofacker, 2003) to determine whether genomic regions flanking sites yielding small RNAs are able to fold into miRNA-like hairpins. MiPred, which relies on an RF algorithm, then determines whether each hairpin has structural properties similar to known miRNAs (Jiang et al., 2007). A second approach, termed miRDeep, was specifically designed to use aligned small RNA reads to identify signatures of Drosha and Dicer cleavage in addition to a pre-miRNA-like structure (Friedlander et al., 2008). Predictions from both methods were refined using EvoFold, which identifies evolutionarily conserved hairpins based on the conservation of structural properties in the genomic sequence (Pedersen et al., 2006).

miRNA amplification

Semi-quantitative PCR was performed as described previously (Ro et al., 2007) with modifications. Small RNAs from 11 different mouse tissues (brain, lung, heart, stomach, liver, muscle, kidney, uterus, adult ovary, testis and NB ovary) were isolated using the mirVana™ miRNA isolation kit (Ambion). The small RNA fraction was poly-adenylated at 37°C for 1 h with poly(A) polymerase (Ambion). The poly(A)-tailed small RNA fraction was treated with DNase I (Invitrogen, Carlsbad, CA, USA) at room temperature for 15 min. RTQ primer (CGAATTCTAGAGCTCGAGGCAGGCGACATGGCTGGCTAGT TAAGCTTGGTACCGAGCTCGGATCCACTAGTCC(T)$_{25}$-) was annealed to the cDNA and reverse transcription was carried out with 200 U of SuperScript II reverse-transcriptase (Invitrogen). Following treatment with RNase H (Invitrogen), the cDNA was used to perform RT−PCR with a miRNA-specific primer, and a universal reverse primer, RTQ-UNIV (CGAATTCTAGAGCTCGAGGCAGG), was used for

PCR amplification of each miRNA. miRNA-specific primer sequences and annealing temperatures used to verify NB ovary preferential miRNAs are listed in Supplementary Table S1. HotStart *Taq* (Qiagen, Valencia, CA, USA) was used to perform PCR. The annealing temperature was adjusted according to the Tm of each miRNA. PCR products were analyzed by electrophoresis on a 1.5% agarose gel.

Results

Sequencing and annotation of small RNAs from the NB mouse ovary

The Genome Analyzer sequencing of NB ovary isolated small RNA fraction produced 4 655 992 Genome Analyzer reads of 33 bp each representing 153 647 736 bp of sequence data. After uniqueness filtering, the Pash method was used to map Solexa reads from small, non-coding RNAs onto the mouse reference genome. Several Pash runs were performed using different parameters in order to optimize the mapping of Solexa reads derived from small RNAs. Optimization of Pash mapping parameters was guided by both the number of reads mapped as a measure of sensitivity and the number of mappings to known RNAs as a measure of specificity. Sequences were mapped based on their overlaps with publicly available genome annotations including miRNAs, rRNAs, tRNAs, other small RNAs and genomic repeats. Read mappings were clustered based on overlapping mapping coordinates and intersected with known miRNAs, snoRNAs, piRNA clusters and genomic repeats (Table I). The most optimal Pash run mapped 50.13% (2 333 996 reads) of the reads. A total of 398 distinct miRNA sequences were identified in the NB mouse ovary transcriptome, out of 492 known miRNAs (Supplementary Table SII). One hundred and thirty-six out of 211 known snoRNAs (Supplementary Table SIII) and 564 390 out of 1 396 863 known piRNAs (Supplementary Table SIV) were found in the NB ovary small ovary transcriptome. Among reads that were mapped to the genome, 25.24% mapped to miRNAs, 25.5% to genomic repeats, 3.5% to piRNAs, 1.88% to tRNAs and 1% to rRNAs, snRNAs, scRNAs (small cytoplasmic RNAs), srpRNAs (signal recognition particle RNAs) and snoRNAs (Fig. 1A and Table II). Among 398 miRNAs expressed from mouse NB ovaries, 53.02% derived from intergenic regions, 40.95% from intronic regions, 3.27% from larger non-coding RNAs, 1.76% from exons and 0.75% from 3′UTR (Fig. 1B, Table III). These results indicate that mouse NB ovaries, composed of both the somatic and the germ cell components, express a large repertoire of small RNAs.

## Chromosomal location and characteristics of the different classes of small RNAs

miRNA was the most abundant small RNA population mapped to the genome. mmu-mir-320 was the most abundant miRNA sequence in the NB ovary, followed by mmu-let-7i and mmu-let-7d (Table IV). miRNA expression, as determined by the total number of reads mapped per chromosome, originated mostly from the X chromosome (73 586 reads mapped), followed by chromosome 2, 1 and 10 (Fig. 2A). The most abundant miRNAs from the X chromosome included a cluster that encodes mmu-mir-322, mmu-mir-503, mmu-mir-351, mmu-mir-542, mmu-mir-450a-1, mmu-mir-450a-2 and mmu-mir-450b. It is interesting to note that the number of reads in the mmu-mir-322 cluster, decreased from the 5′ to the 3′ direction with mmu-mir-322 (located at the 5′ end of the cluster) represented by 16 458 reads and mmu-mir-450b (located at the 3′ end of the cluster) represented by 315 reads. Other notable miRNAs were mmu-mir-672 and the mmu-mir-465 family, which includes mmu-mir-465a-5p, mmu-mir-465b-5p, mmu-mir-465c-5p and mmu-mir-465-3p. Semi-quantitative RT−PCR on multiple tissues showed that X chromosome expressed mmu-mir-503, mmu-mir-672 (Fig. 3A) and all members of the mmu-mir-465 family (mir-465a-5p, mir-465b-5p, mir-465c-5p and mir-465c-3p) are preferentially expressed in the gonads (Fig. 3B). In addition to the X chromosome expressed miRNAs, mmu-mir-202 and mmu-mir-298, expressed from chromosomes 7 and 2, respectively, also showed preferential expression in the gonads (Fig. 3A). Most known miRNA genes map to chromosome 2, followed by chrX and chr12 (Fig. 2B). As expected, most miRNA transcript reads from the NB ovary mapped to chrX and chr2 (Fig. 2A and B); however, expression did not simply follow gene density, as more reads mapped to chr10 than chr12, despite six times as many miRNA genes on chr12 rather than chr10.

Massive parallel sequencing also detected snoRNAs. One hundred and thirty-six out of 211 known snoRNAs were encountered in the NB ovary (Supplementary Table SIII). U3B.2 and Z12 were the most abundant snoRNAs detected in the NB ovary and map to chromosome 11. In addition to snoRNAs, we detected piRNAs. piRNAs are small RNA clusters initially discovered in mammalian testes and associate with PIWI proteins. Although thought to be exclusively expressed in the testes, piRNA clusters were detected in the NB ovary. piR_028252 from chromosome 7 was the most abundantly expressed piRNA in the NB ovary with 19 374 reads, followed by piR_033077 with 6153 reads and piR_010700 with 2996 reads, both from the X chromosome. Among 14 piRNA sequences with over 1000 reads, there were more individually expressed piRNAs

Table I Comparisons of Pash mapped Solexa reads to known small RNAs.

Small RNA type	Known RNAs	RNAs overlapped by mappings	% RNAs overlapped by mappings	Mappings to RNA	[d]Average mappings/RNA
[a]miRNA	492	398	80.89%	589 033	1479.98
[b]snoRNA	211	136	64.45%	4156	30.55882353
[c]piRNA clusters	1 396 863	564 390	40.40%	81 594	0.144570244

[a]Known RNAs were obtained from UCSC Genome Browser miRNA track on mm9, [b]snoRNAbase (http://gene.fudan.sh.cn/snoRNAbase.nsf), and [c]piRNA Bank (Sai Lakshmi and Agrawal, 2008). [d]The average mappings/RNA for piRNA is lower than 1 because most reads that map to piRNA map to multiple locations of the piRNA in the genome.
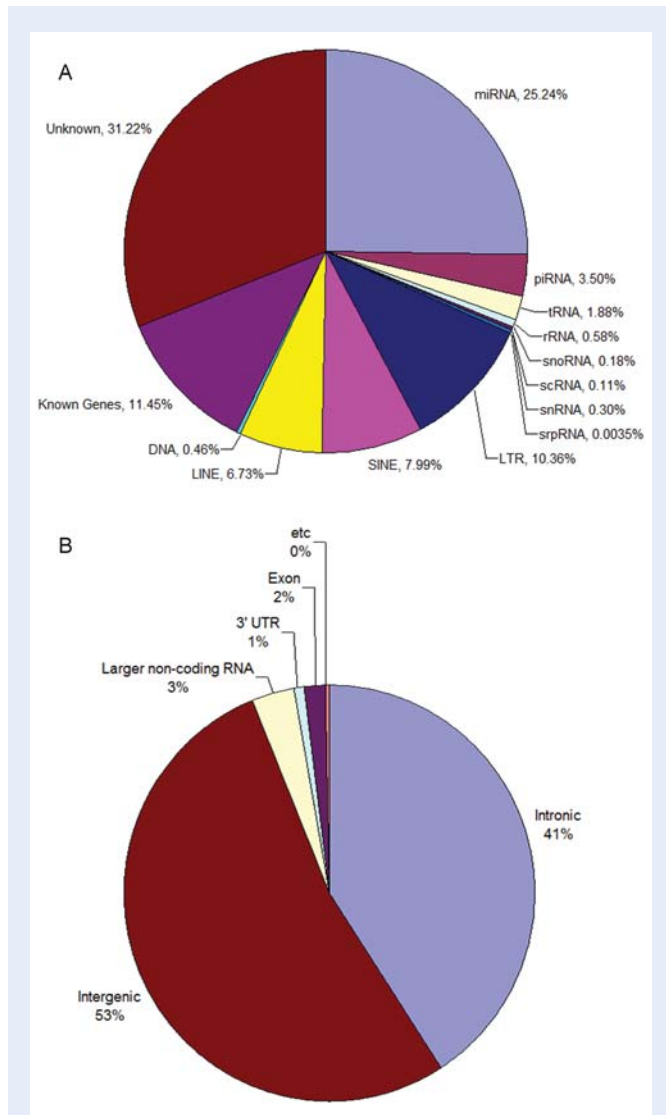
Figure 1 Pie chart of (**A**) sequences that mapped to known RNAs, repeats and genes and (**B**) genomic context of sequences mapped to known miRNAs. Sequences were clustered based on overlapping mapping coordinates and intersected with known miRNAs, snoRNAs, piRNA clusters and repeats. scRNA, small cytoplasmic RNA; srpRNA, signal recognition particle RNA; SINE, short interspersed repetitive elements; LINE, long interspersed repetitive elements.

than piRNAs expressed in clusters (Supplementary Table SIV). piRNAs mapped predominantly to chromosomes 1, X and 2.

Novel small RNAs

Read clusters not intersecting with known RNA species and consisting of at least 100 reads were further considered as putative novel miRNAs. We identified a total of 30 'novel' miRNAs that met our previously established criteria for miRNA (Morin *et al.*, 2008). We examined a subset of the predicted miRNAs for multi-tissue expression pattern. Most of the examined putative miRNAs were ubiquitously expressed. However, four novel miRNAs, novel 11, novel 12, novel 18 and novel 20, derived from four distinct regions in the genome

Table II Proportion of reads mapping to known RNAs, repeats and genes.

[e]Genomic element	Percent mapped reads
RNA	
miRNA	25.24%
piRNA	3.50%
tRNA	1.88%
rRNA	0.58%
snoRNA	0.18%
scRNA	0.11%
snRNA	0.30%
srpRNA	0.0035%
Repeats	
LTR	10.36%
SINE	7.99%
LINE	6.73%
DNA	0.46%
Genes	
Known Genes	11.45%

[e]scRNA, small cytoplasmic RNA; srpRNA, signal recognition particle; SINE, short interspersed repetitive elements; LINE, Long interspersed repetitive elements.

(Table V) showed preferential expression in the reproductive tract (Fig. 4A). However, the expression was not exclusive to the ovaries. One miRNA, mmu-mir-1981, was initially classified as a novel miRNA but was recently identified in mouse ES cells (Babiarz *et al.*, 2008). mmu-mir-1981 expression was confined to the gonads (Fig. 3A).

A total of 522 distinct sequences not overlapping known RNA species were identified in addition to the novel putative miRNAs and did not appear to derive from the pre-miRNA transcripts (Supplementary Table SV). Twenty-five out of the 522 such transcripts shared 100% conservation with rat and human sequences, whereas 40 transcripts shared 90% or higher conservation with rat and human. We examined the expression of one of the conserved small RNAs that was not predicted to derive from pre-miRNA, NovelRNA 361, and found that this transcript was preferentially expressed in the gonads (Fig. 4A). Interestingly this transcript was derived from an intron within *Astn2* gene, which is preferentially expressed in the brain, turbinates, eye and dorsal root ganglion (Wheeler *et al.*, 2008). These findings suggest that some of these unidentified transcripts may play an important function during gonadal development.

Identification of isomiR sequences in the NB ovary miRNAs

It has been previously shown that miRNA variants are detectable on massive parallel sequencing (Morin *et al.*, 2008). These multiple mature variants have been referred to as isomiRs (Morin *et al.*, 2008). The origin of isomiRs is most likely due to the Dicer or Drosha variant cleavage within the pre-miRNA hairpin loop, although the reason for this variant cleavage is unknown. We identified eight isomiRs with reads more than 100 (Supplementary Table SVI).

Table III Summary of miRNA genome context.

	No. of miRNAs in the database	No. of miRNAs found by mapping	% miRNA with mapping	Number of mapped reads	Reads/miRNA
Intron	215	163	0.758139535	213 894	1312.233129
Intergenic	246	211	0.857723577	421 296	1996.663507
Larger non-coding RNA	13	13	1	1771	136.2307692
5' UTR	1	0	0	0	N/A
3' UTR	5	3	0.6	8775	2925
Exon	11	7	0.636363636	17 918	2559.714286
ᶠetc	1	1	1	3	3
Total	492	398			

ᶠNo longer included in UCSC track.

Table IV Top 20 expressed miRNAs based on read counts.

miRNA Name	Genomic location	Mature sequence (5p/3p)	Count	Genomic context
mmu-mir-320	chr14:70843317-70843398	AAAAGCTGGGTTGAGAGGGCGA	51 264	Intergenic
mmu-let-7i	chr10:122422696-122422780	TGAGGTAGTAGTTTGTGCTGTT	43 029	Intergenic
mmu-let-7d	chr13:48631381-48631483	AGAGGTAGTAGGTTGCATAGTT	34 173	Intergenic
mmu-mir-298	chr2:174093005-174093086	GGCAGAGGAGGGCTGTTCTTCCC	30 259	Intergenic
mmu-mir-199a-1	chr9:21300939-21301008	CCCAGTGTTCAGACTACCTGTTC/ ACAGTAGTCTGCACATTGGTTA	23 252	Intron
mmu-mir-30a	chr1:23279108-23279178	TGTAAACATCCTCGACTGGAAG	22 864	Intergenic
mmu-mir-140	chr8:110075144-110075213	CAGTGGTTTTACCCTATGGTAG	22 796	Intron
mmu-mir-322	chrX:50407432-50407526	CAGCAGCAATTCATGTTTTGGA	16 458	Intergenic
mmu-mir-152	chr11:96711707-96711779	TCAGTGCATGACAGAACTTGG	14 976	Intron
mmu-mir-26a-2	chr10:126432586-126432669	TTCAAGTAATCCAGGATAGGCT	14 858	Intron
mmu-mir-423	chr11:76891566-76891674	TGAGGGGCAGAGAGCGAGACTTT/ AGCTCGGTCTGAGGCCCCTCAGT	14 666	Intron
mmu-mir-99b	chr17:17967152-17967221	CACCCGTAGAACCGACCTTGCG	14 362	Intergenic
mmu-mir-199a-2	chr1:164147945-164148054	CCCAGTGTTCAGACTACCTGTTC/ ACAGTAGTCTGCACATTGGTTA	13 953	Intron
mmu-mir-199b	chr2:32173980-32174089	ACAGTAGTCTGCACATTGGTTA	13 393	Intron
mmu-mir-127	chr12:110831056-110831125	TCGGATCCGTCTGAGCTTGGCT	13 166	Exon
mmu-mir-25	chr5:138606549-138606632	CATTGCACTTGTCTCGGTCTGA	12 696	Intron
mmu-mir-143	chr18:61808850-61808912	TGAGATGAAGCACTGTAGCTC	12 634	Intergenic
mmu-mir-181a-2	chr2:38708255-38708330	AACATTCAACGCTGTCGGTGAGT	11 437	Intron
mmu-mir-181a-1	chr1:139863032-139863118	AACATTCAACGCTGTCGGTGAGT	11 199	Intergenic
mmu-mir-26a-1	chr9:118940914-118941003	TTCAAGTAATCCAGGATAGGCT	10 772	Intron

mmu-mir-135 and mmu-mir-871 isomiRs were most abundant, with 1249 and 472 reads detected on sequencing. mmu-mir-135 is expressed in multiple tissues; however, the isomiR detected in the NB ovary, mmu-mir-135a-2-3p, was expressed preferentially in the gonads (Fig. 4B). mmu-mir-871 is expressed from the X chromosome and shows preferential expression in the gonads, whereas its isomiR, mmu-mir-871-3p, is also expressed preferentially in the gonads (Fig. 4B). Interestingly, mmu-mir-871 originates from the same cluster as mmu-mir-465, which we previously showed to be expressed

preferentially in the gonads (Fig. 3B). Both mmu-mir-135a-2-3p and mmu-mir-871-3p isomiR sequences were absent from the mouse miRBase database, corroborating RT−PCR results that these sequences are preferentially expressed in the gonads.

Discussion

NB ovary is characterized by heterogenous group of structures including germ cell clusters and primordial follicles. It is during this stage of
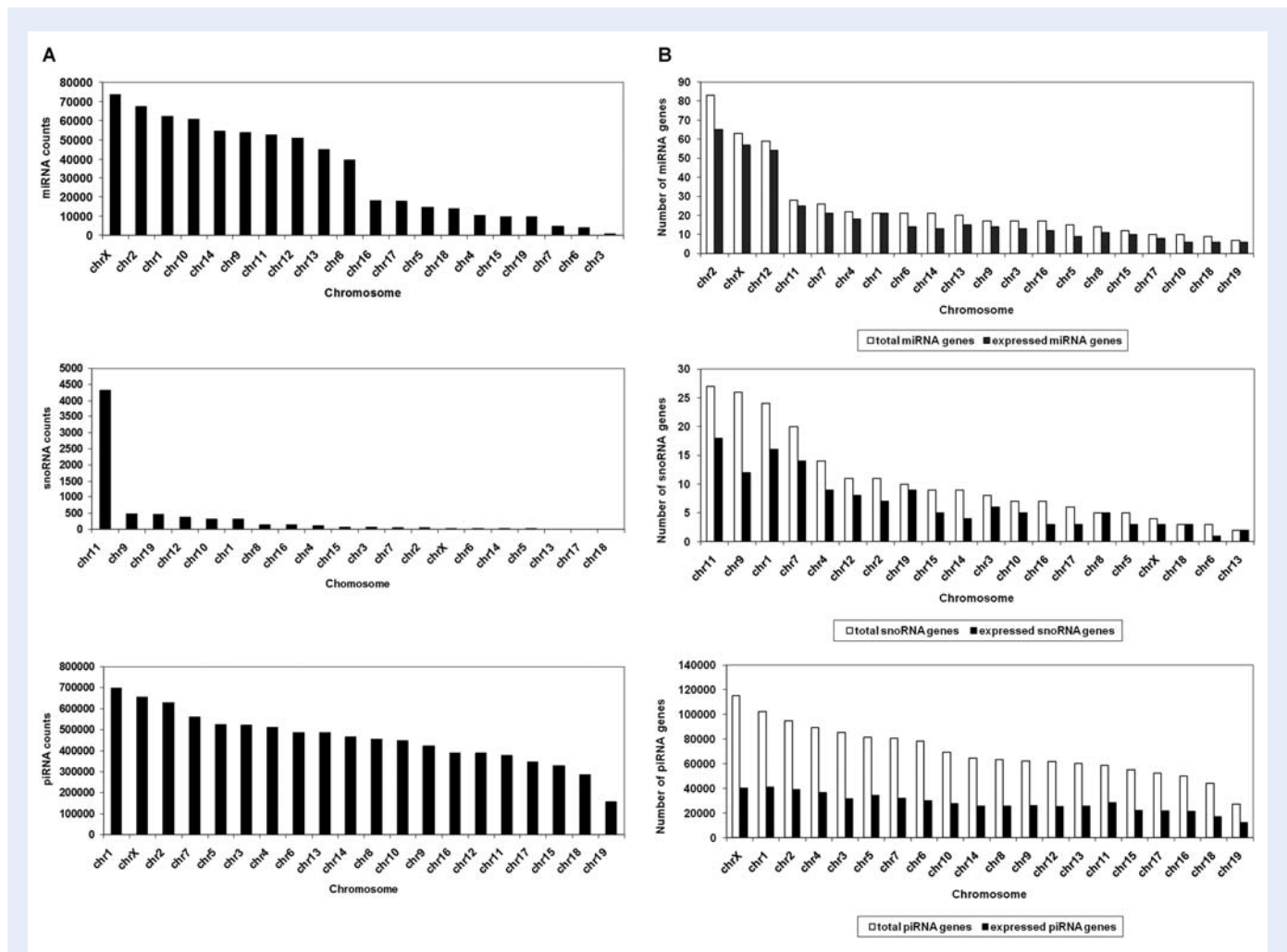
Figure 2 Chromosomal location of miRNAs, piRNAs and snoRNAs based on the number of sequence reads (A) and number of genes (B).

ovarian development that many oocytes within the cluster are eliminated to yield a final pool of primordial follicles (Pangas and Rajkovic, 2006). A number of germ cell specific RNA binding proteins have been identified such as MSY2, DAZLA, and NANOS3, and their respective knockouts show the importance of such proteins and RNA metabolism in the female gonadal development (Ruggiu *et al.*, 1997; McNeilly *et al.*, 2000; Tsuda *et al.*, 2003; Yang *et al.*, 2006). Conditional knockouts of Dicer1 and Argonaute2 in growing mouse oocytes (Murchison *et al.*, 2007; Kaneda *et al.*, 2009), two important enzymes in miRNA biogenesis, show normal ovarian development, but post-ovulation oocyte maturation is defective likely due to perturbed completion of meiosis I. The role of miRNAs during NB ovary development is unclear. We have previously shown that a number of germ cell specific transcriptional regulators are highly expressed in the NB ovary (Choi and Rajkovic, 2006). We therefore examined the population of small RNAs in the NB ovary. We used massive parallel sequencing to determine the small RNA transcriptome in the 18–30 nucleotide range in the NB mouse ovary, and identified 398 known miRNAs and 30 novel small RNAs predicted to be miRNAs.

We also identified a significant expression of piRNAs, likely originating from the oocytes within the NB ovaries, and snoRNAs. piRNAs interact with PIWI proteins and knockouts of individual PIWI family members result in male sterility; however, female fertility is not affected by such knockouts (Deng and Lin, 2002; Kuramochi-Miyagawa *et al.*, 2004; Carmell *et al.*, 2007). The significance of piRNAs in the NB ovaries is unclear, although recent evidence suggest that piRNAs are expressed in the Drosophila and mouse female germlines (Chambeyron *et al.*, 2008; Tam *et al.*, 2008) and important in germline silencing of transposable elements. Whether redundant, piRNA-independent pathways exist in the female germline and suppress transposable elements remains to be seen. It is perhaps of interest that almost 522 transcripts matched un-annotated sites in the genome, one of which showed preferential expression in the gonads. The role of such transcripts in the general nucleic acid metabolism of the germ cells remains to be established.

We showed that the total number of miRNA sequences mapped to the X chromosome more than any other chromosome, followed by chr2, chr1 and chr10. The number of miRNA genes is highest on chromosomes 2, chrX and chr12, and thus perhaps it is not
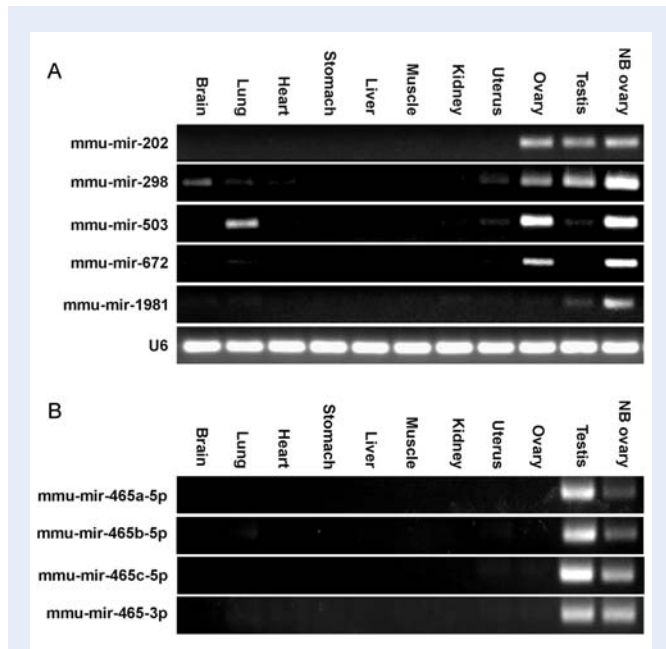
Figure 3 Multi-tissue RT–PCR of selected known miRNAs. Small RNA fraction was isolated and cDNA synthesized from 11 different mouse tissues followed by PCR. Expression profiles of (**A**) mmu-mir-202, mmu-mir-503 and mmu-mir-672 and (**B**) mmu-mir-465 cluster (mmu-mir-465a, mmu-mir-465b, mmu-mir-465c and mmu-mir-465-3p) are shown. U6 RNA amplification was used as a positive control.

unexpected that chrX and chr2 express more miRNA transcripts than other chromosomes. However, correlation between gene density on a particular chromosome, and expression is not uniform. For example, more reads mapped to chr10 than chr12, despite six times as many miRNA genes on chr12 rather than chr10. The significance of the preferential expression of miRNAs from certain chromosomes in the NB ovary is unclear.

We discovered a total of 30 potential novel miRNAs using methodologies described previously (Morin et al., 2008). The number of novel small RNA species mapped in our study is much higher than previously described in the adult ovaries (Ro et al., 2007; Mishima et al., 2008) and is largely due to the use of massive parallel sequencing technology. For example, Mishima et al. (2008) only sequenced 10 000 clones from the adult ovary and mapped them to only 154 known miRNA genes, whereas we obtained more than 4 000 000 reads that mapped to 398 known miRNA genes. There are also significant differences between NB and adult ovaries. RNA contribution in adult ovaries is mostly derived from the somatic cell component, whereas NB ovaries are highly enriched in the germ cell component. The novel miRNAs we discovered were not very abundant based on the sequence count, with total number of reads varying between 12 and 1100. Most of these miRNAs were expressed ubiquitously, although few did show preferential expression in the gonads. The isomiRs expressed in the NB ovary were not abundant with read counts numbering less than 100 for most of them. It is interesting to note that the top two expressed isomiRs, mmu-mir-135a-2-3p and mmu-mir-871-3p, were both preferentially expressed in gonads, suggesting that gonadal milieu plays a role in processing of these transcripts.

Table V List of predicted novel miRNAs.

Name	Genomic location	Most abundant mature sequence	Count	[a]Expression profile
novel 11	chr6:83056012-83056041	TCACTTTGTAGACCAGGCTGG	1082	Specific
novel 12	chr16:43933215-43933245	TCACTCTGTAGACCAGGCTGG	674	Specific
novel 13	chr16:4736237-4736259	ATCCCGGACGAGCCCCC	851	Ubiquitous
novel 14	chr19:7537680-7537697	TGAGGTAGGAGATTG	413	Ubiquitous
novel 15	chr3:47577930-47577953	ATCCCGGACGAGCCCCC	825	Ubiquitous
novel 16	chr19:24686349-24686366	AGCAGCATTGGACAG	284	Ubiquitous
novel 17	chr7:115354363-115354391	GGGGATGTAGCTCAGTGGTAG	185	N/A
novel 18	chr4:10772729-10772759	TGAGATCCAACTGTAAGGCATT	866	Specific
novel 19	chr12:9018369-9018389	TGAGGTAGTAGGTTG	717	N/A
novel 20	chr7:6407724-6407747	TAGGCTAGAGAGAGGTTGGGGA	200	Specific
novel 21	chr10:126159550-126159568	TGAGGTAGTGGGTTG	423	Ubiquitous
novel 22	chr12:78299884-78299905	ATCCCGGACGAGCCCCC	851	Ubiquitous
novel 23	chr4:3464245-3464275	TGAGATCCAACTGTAAGGCATT	727	Specific
novel 24	chr17:19595919-19595942	ATCCCGGACGAGCCCCC	825	Ubiquitous
novel 25	chr4:131788084-131788114	GGGGATGTAGCTCAGTGGTAG	185	N/A
novel 26	chr2:155262385-155262401	TGAGGTAGGAGATTG	413	Ubiquitous
novel 27	chr1:53641015-53641038	ATCCCGGACGAGCCCCC	825	Ubiquitous
novel 28	chr2:118090634-118090658	GGGGATGTAGCTCAGTGGTAG	185	N/A
novel 53	chr1:190308113-190308137	GGTAGGTGTATGTTCATATGCCGC	104	N/A
novel 441	chr7:6756349-6756376	AGGCTAGAGAGAGGTTGGGGATGGGGA	607	Ubiquitous

Continued

Table V *Continued*

Name	Genomic location	Most abundant mature sequence	Count	[a]Expression profile
novel 42	chrX:141473887-141473957	TAATAGCCAGAAGCTGGAAAGAACC	12	N/A
novel 165	chr8:98487444 -98487507	TGGGATTAAAGGCATGCACCAC	17	N/A
novel 166	chrX:159312853-159312922	TGGAGAGATGGCTCAGCC	21	N/A
novel 215	chr12:60168183-60168229	CGTTTCCCGGGCGGCG	22	N/A
novel 220	chr5:34529822-34529868	CGCGCGGGCGGGGCCGGG	18	N/A
novel 221	chr1:52289212-52289295	CGCGCGGGCGGGGCCGGG	18	N/A
novel 567	chr11:97298088-97298125	TGAGGTAGGAGGTTGGA	223	N/A
novel 572	chr12:110986103-110986159	TGCCCCCTCCAGGAAGCCTTCT	29	Ubiquitous
novel 635	chr2:158464319-158464376	CCCTGGGAGGAGACGTGGATTC	20	N/A
novel 766	chrX:136845487-136845551	TCTGGAGGCACATGGTTTGAA	49	N/A

[a]Ubiquitous, expression in all examined tissue types; Specific, preferential expression in gonads; N/A, PCR not amplified or data not available.
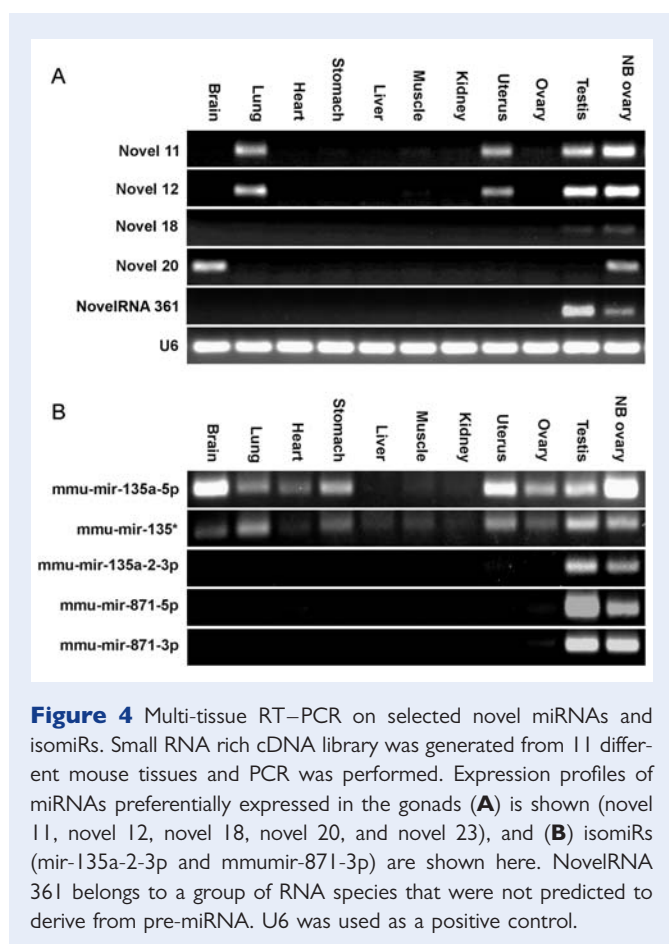
Figure 4 Multi-tissue RT−PCR on selected novel miRNAs and isomiRs. Small RNA rich cDNA library was generated from 11 different mouse tissues and PCR was performed. Expression profiles of miRNAs preferentially expressed in the gonads (**A**) is shown (novel 11, novel 12, novel 18, novel 20, and novel 23), and (**B**) isomiRs (mir-135a-2-3p and mmumir-871-3p) are shown here. NovelRNA 361 belongs to a group of RNA species that were not predicted to derive from pre-miRNA. U6 was used as a positive control.

The role of any of these miRNAs in ovarian follicular development, either novel or isomiRs, is unknown. Moreover, it is unclear whether these miRNAs are redundant in their function, as has been suggested for other miRNAs (Miska *et al.*, 2007). Future studies ablating specific miRNAs using transgenic technologies will help us better understand the role of these clusters as well as individual miRNAs in gonadal development and ovarian folliculogenesis.

Supplementary data

Supplementary data are available at http://molehr.oxfordjournals.org/.

Acknowledgements

Funding

References

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 1997; **25**:3389−3402.

Babiarz JE, Ruby JG, Wang Y, Bartel DP, Blelloch R. Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. *Genes Dev* 2008;**22**:2773−2785.

Ballow DJ, Xin Y, Choi Y, Pangas SA, Rajkovic A. Sohlh2 is a germ cell-specific bHLH transcription factor. *Gene Expr Patterns* 2006; **6**:1014−1018.

Carmell MA, Girard A, van de Kant HJ, Bourc'his D, Bestor TH, de Rooij DG, Hannon GJ. MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. *Dev Cell* 2007; **12**:503−514.

Carthew RW, Sontheimer EJ. Origins and Mechanisms of miRNAs and siRNAs. *Cell* 2009;**136**:642−655.

Chambeyron S, Popkova A, Payen-Groschene G, Brun C, Laouini D, Pelisson A, Bucheton A. piRNA-mediated nuclear accumulation of retrotransposon transcripts in the *Drosophila* female germline. *Proc Natl Acad Sci USA* 2008;**105**:14964−14969.

Choi Y, Rajkovic A. Genetics of early mammalian folliculogenesis. *Cell Mol Life Sci* 2006;**63**:579−590.

Choi Y, Qin Y, Berger MF, Ballow DJ, Bulyk ML, Rajkovic A. Microarray analyses of newborn mouse ovaries lacking Nobox. *Biol Reprod* 2007; **77**:312–319.

Choi Y, Yuan D, Rajkovic A. Germ cell-specific transcriptional regulator sohlh2 is essential for early mouse folliculogenesis and oocyte-specific gene expression. *Biol Reprod* 2008;**79**:1176–1182.

Coarfa C, Milosavljevic A. Pash 2.0: scaleable sequence anchoring for next-generation sequencing technologies. *Pac Symp Biocomput* 2008; 102–113.

Deng W, Lin H. miwi, a murine homolog of piwi, encodes a cytoplasmic protein essential for spermatogenesis. *Dev Cell* 2002;**2**:819–830.

Doench JG, Sharp PA. Specificity of microRNA target selection in translational repression. *Genes Dev* 2004;**18**:504–511.

Fiedler SD, Carletti MZ, Hong X, Christenson LK. Hormonal regulation of microRNA expression in periovulatory mouse mural granulosa cells. *Biol Reprod* 2008;**79**:1030–1037.

Filipowicz W, Bhattacharyya SN, Sonenberg N. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat Rev Genet* 2008;**9**:102–114.

Friedlander MR, Chen W, Adamidi C, Maaskola J, Einspanier R, Knespel S, Rajewsky N. Discovering microRNAs from deep sequencing data using miRDeep. *Nat Biotechnol* 2008;**26**:407–415.

Griffiths-Jones S, Grocock RJ, van Dongen S, Bateman A, Enright AJ. miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res* 2006;**34**:D140–D144.

Hofacker IL. Vienna RNA secondary structure server. *Nucleic Acids Res* 2003;**31**:3429–3431.

Jiang P, Wu H, Wang W, Ma W, Sun X, Lu Z. MiPred: classification of real and pseudo microRNA precursors using random forest prediction model with combined features. *Nucleic Acids Res* 2007;**35**: W339–W344.

Kalafus KJ, Jackson AR, Milosavljevic A. Pash: efficient genome-scale sequence anchoring by positional hashing. *Genome Res* 2004; **14**:672–678.

Kaneda M, Tang F, O'Carroll D, Lao K, Surani MA. Essential role for Argonaute2 protein in mouse oogenesis. *Epigenetics Chromatin* 2009;**2**:9.

Kent WJ, Baertsch R, Hinrichs A, Miller W, Haussler D. Evolution's cauldron: duplication, deletion, and rearrangement in the mouse and human genomes. *Proc Natl Acad Sci USA* 2003;**100**:11484–11489.

Kuramochi-Miyagawa S, Kimura T, Ijiri TW, Isobe T, Asada N, Fujita Y, Ikawa M, Iwai N, Okabe M, Deng W et al. Mili, a mammalian member of piwi family gene, is essential for spermatogenesis. *Development* 2004;**131**:839–849.

McNeilly JR, Saunders PT, Taggart M, Cranfield M, Cooke HJ, McNeilly AS. Loss of oocytes in Dazl knockout mice results in maintained ovarian steroidogenic function but altered gonadotropin secretion in adult animals. *Endocrinology* 2000;**141**:4284–4294.

Meister G, Tuschl T. Mechanisms of gene silencing by double-stranded RNA. *Nature* 2004;**431**:343–349.

Mishima T, Takizawa T, Luo SS, Ishibashi O, Kawahigashi Y, Mizuguchi Y, Ishikawa T, Mori M, Kanda T, Goto T et al. MicroRNA (miRNA) cloning analysis reveals sex differences in miRNA expression profiles between adult mouse testis and ovary. *Reproduction* 2008;**136**:811–822.

Miska EA, Alvarez-Saavedra E, Abbott AL, Lau NC, Hellman AB, McGonagle SM, Bartel DP, Ambros VR, Horvitz HR. Most *Caenorhabditis elegans* microRNAs are individually not essential for development or viability. *PLoS Genet* 2007;**3**:e215.

Morin RD, O'Connor MD, Griffith M, Kuchenbauer F, Delaney A, Prabhu AL, Zhao Y, McDonald H, Zeng T, Hirst M et al. Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. *Genome Res* 2008;**18**:610–621.

Morozova O, Hirst M, Marra MA. Applications of new sequencing technologies for transcriptome analysis. *Annu Rev Genomics Hum Genet* 2009;**10**:135–151.

Murchison EP, Stein P, Xuan Z, Pan H, Zhang MQ, Schultz RM, Hannon GJ. Critical roles for Dicer in the female germline. *Genes Dev* 2007;**21**:682–693.

Olsen PH, Ambros V. The lin-4 regulatory RNA controls developmental timing in *Caenorhabditis elegans* by blocking LIN-14 protein synthesis after the initiation of translation. *Dev Biol* 1999;**216**:671–680.

Pangas SA, Rajkovic A. Transcriptional regulation of early oogenesis: in search of masters. *Hum Reprod Update* 2006;**12**:65–76.

Pedersen JS, Bejerano G, Siepel A, Rosenbloom K, Lindblad-Toh K, Lander ES, Kent J, Miller W, Haussler D. Identification and classification of conserved RNA secondary structures in the human genome. *PLoS Comput Biol* 2006;**2**:e33.

Pepling ME, Spradling AC. Mouse ovarian germ cell cysts undergo programmed breakdown to form primordial follicles. *Dev Biol* 2001; **234**:339–351.

Ro S, Song R, Park C, Zheng H, Sanders KM, Yan W. Cloning and expression profiling of small RNAs expressed in the mouse ovary. *RNA* 2007;**13**:2366–2380.

Ruggiu M, Speed R, Taggart M, McKay SJ, Kilanowski F, Saunders P, Dorin J, Cooke HJ. The mouse Dazla gene encodes a cytoplasmic protein essential for gametogenesis. *Nature* 1997;**389**:73–77.

Sai Lakshmi S, Agrawal S. piRNABank: a web resource on classified and clustered Piwi-interacting RNAs. *Nucleic Acids Res* 2008; **36**:D173–D177.

Sewer A, Paul N, Landgraf P, Aravin A, Pfeffer S, Brownstein MJ, Tuschl T, van Nimwegen E, Zavolan M. Identification of clustered microRNAs using an ab initio prediction method. *BMC Bioinformatics* 2005;**6**:267.

Soyal SM, Amleh A, Dean J. FIGalpha, a germ cell-specific transcription factor required for ovarian follicle formation. *Development* 2000; **127**:4645–4654.

Tam OH, Aravin AA, Stein P, Girard A, Murchison EP, Cheloufi S, Hodges E, Anger M, Sachidanandam R, Schultz RM et al. Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. *Nature* 2008;**453**:534–538.

Tsuda M, Sasaoka Y, Kiso M, Abe K, Haraguchi S, Kobayashi S, Saga Y. Conserved role of nanos proteins in germ cell development. *Science* 2003;**301**:1239–1241.

Wheeler DL, Barrett T, Benson DA, Bryant SH, Canese K, Chetvernin V, Church DM, Dicuccio M, Edgar R, Federhen S et al. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* 2008;**36**:D13–D21.

Wilhelm D, Palmer S, Koopman P. Sex determination and gonadal development in mammals. *Physiol Rev* 2007;**87**:1–28.

Yang J, Medvedev S, Yu J, Schultz RM, Hecht NB. Deletion of the DNA/RNA-binding protein MSY2 leads to post-meiotic arrest. *Mol Cell Endocrinol* 2006;**250**:20–24.