Physiol Genomics 42A: 39–51, 2010.
First published May 4, 2010; doi:10.1152/physiolgenomics.00054.2010.

Genomic analysis of miRNAs in an extreme mammalian hibernator, the Arctic ground squirrel

Yuting Liu,[1]* Wenchao Hu,[1]* Haifang Wang,[1] Minghua Lu,[2] Chunxuan Shao,[1] Corinna Menzel,[3] Zheng Yan,[1] Ying Li,[1] Sen Zhao,[1] Philipp Khaitovich,[1] Mofang Liu,[2] Wei Chen,[3] Brian M. Barnes,[4] and Jun Yan[1]

[1]*CAS-MPG Partner Institute for Computational Biology,* [2]*Core Facility for Noncoding RNA, Institute of Biochemistry and Cell Biology, Shanghai Institutes of Biological Sciences, Shanghai, China;* [3]*Max Planck Institute for Molecular Genetics, Berlin, Germany; and* [4]*Institute of Arctic Biology, University of Alaska Fairbanks, Fairbanks, Alaska*

Submitted 22 March 2010; accepted in final form 3 May 2010

Liu Y, Hu W, Wang H, Lu M, Shao C, Menzel C, Yan Z, Li Y, Zhao S, Khaitovich P, Liu M, Chen W, Barnes BM, Yan J. Genomic analysis of miRNAs in an extreme mammalian hibernator, the Arctic ground squirrel. *Physiol Genomics* 42A: 39–51, 2010. First published May 4, 2010; doi:10.1152/physiolgenomics.00054.2010.—MicroRNAs (miRNAs) are 19- to 25-nucleotide-long small and noncoding RNAs now well-known for their regulatory roles in gene expression through posttranscriptional and translational controls. Mammalian hibernation is a physiological process involving profound changes in set-points for food consumption, body mass and growth, body temperature, and metabolic rate in which miRNAs may play important regulatory roles. In an initial study, we analyzed miRNAs in the liver of an extreme hibernating species, the Arctic ground squirrel (*Spermophilus parryii*), using massively parallel Illumina sequencing technology. We identified >200 ground squirrel miRNAs, including 18 novel miRNAs specific to ground squirrel and *mir-506* that is fast evolving in the ground squirrel lineage. Comparing animals sampled after at least 8 days of continuous torpor (late torpid), within 5 h of a spontaneous arousal episode (early aroused), and 1–2 mo after hibernation had ended (nonhibernating), we identified differentially expressed miRNAs during hibernation, which are also compared with the results from two other miRNA profiling methods: Agilent miRNA microarray and real-time PCR. Among the most significant miRNAs, miR-320 and miR-378 were significantly underexpressed during both stages of hibernation compared with nonhibernating animals, whereas miR-486 and miR-451 were overexpressed in late torpor but returned in early arousal to the levels similar to those in nonhibernating animals. Analyses of their putative target genes suggest that these miRNAs could play an important role in suppressing tumor progression and cell growth during hibernation. High-throughput sequencing data and microarray data have been submitted to GEO database with accession: GSE19808.

microRNA; deep sequencing; hibernation

MicroRNAs (miRNAs) are noncoding regulatory RNAs, 19–25 nucleotides (nt) in length that control gene expression in animals and plants at the posttranscriptional and translational levels (4). Primary miRNA transcripts (pri-miRNAs), transcribed by Pol II polymerase similar to the protein coding mRNAs, are cleaved in the nucleus by Drosha, a RNase III member, producing 60–80 nt precursor miRNAs (pre-miRNA). The pre-miRNAs form hairpin structures and are exported to the cytoplasm, where they are further processed to 19–25 nt double-stranded RNA duplexes through a second cleavage by another RNase III protein, Dicer. One strand is incorporated into the RNA-induced silencing complex (RISC), producing an active mature miRNA and an inactive miRNA*, in most cases. Once assembled within the RISC complex, miRNAs can induce the downregulation of target genes by causing mRNA degradation and/or repression of translation. miRNAs associate with their mRNA targets by base-pair complementarities from 5′ positions 2–8 of mature miRNAs, that are termed "seed regions." In animals, imperfect base-pairings are common in miRNA-target interactions, complicating the miRNA target prediction when using bioinformatics approaches. miRNAs have been shown to play important regulatory roles in important biological processes such as development in a wide range of species. Differential expression of miRNAs can lead to differential regulation of their target genes and, therefore, can be an essential part of molecular adaptation to environmental changes.

Hibernation is an energy-saving strategy adopted seasonally by a wide range of mammalian species to survive extreme environments through regulated metabolic suppression (10). Mammalian hibernators have evolved the ability to sustain profoundly low body temperatures and rates of oxygen consumption through complex cellular and molecular reorganization. Understanding the molecular mechanisms of hibernation should have important medical implications for the development of novel therapies to treat or prevent cardiovascular diseases, injury from ischemia/reperfusion, as well as for prolonging organ preservation. The Arctic ground squirrel (AGS, *Spermophilus parryii*) living in Alaska is an extreme hibernator that sustains core body temperature as low as −2.9°C while hibernating in the wild (3). Hibernation in Arctic ground squirrels, as with all small hibernators, is interrupted every 10–21 days by arousal episodes as they spontaneously rewarm to euthermic body temperatures (36–37°C) for 15–24 h before slowly cooling and reentering torpor. The molecular mechanisms of hibernation and cellular and molecular functions of spontaneous arousal are central issues in hibernation research.

Hibernation is a highly regulated physiological process. On the molecular level, regulatory control and response to hibernation can take place at multiple levels including transcriptional, posttranscriptional, translational, and posttranslational levels. At the mRNA level, we and other laboratories have utilized large-scale genomic approaches to investigate differential gene expression during hibernation in several species of ground squirrel (7, 31–33). At the protein level, we recently

* Y. Liu and W. Hu contributed equally to this work.

Address for reprint requests and other correspondence: J. Yan, CAS-MPG Partner Inst. for Computational Biology, Shanghai Institutes of Biological Sciences, 320 Yue Yang Rd., Shanghai, 200031, China (e-mail: junyan@picb.ac.cn).

conducted a global high-throughput proteomic study in liver tissue of AGS sampled during different hibernating states using an unbiased LC-MS/MS approach (28). Comparing proteomic and gene-expression data in the same tissues, we observed inconsistent changes for a large number of genes between mRNA and protein levels indicating substantial posttranscriptional regulation during hibernation. Potentially, miRNA represents an important posttranscriptional regulatory mechanism in hibernation that could explain differences between transcriptional and translational levels. In the first report of miRNA changes in hibernators, Morin et al. (25) examined selected miRNAs expression in four organs of hibernating thirteen-lined ground squirrels (*Spermophilus tridecemlineatus*) by RT-PCR. Although there were no changes found in liver, they reported underexpression of miR-24 in heart and skeletal muscle, underexpression of miR-122a in skeletal muscle, and overexpression of miR-1 and miR-21 in kidney of torpid animals, compared with euthermic, nonhibernating animals. So far, however, systematic study on miRNA in any hibernating species has been lacking.

The advent of massively parallel sequencing technologies such as Illumina/Solexa sequencing has greatly facilitated discovery and analysis of small RNAs. Ultrahigh-throughput sequence in synthesis technology can generate several million small RNA sequences between 17 and 35 nt in each small RNA library in one run. It can discover novel miRNAs by mapping the sequence reads onto the genome, and the numerical frequency of specific small RNAs can represent their expression levels in the library. So far such miRNA sequencing technology has been mainly applied in model organisms (15, 26). In this paper, we report results of the first study using Illumina sequencing technology to describe the profile of small RNAs in a hibernating species, the AGS. We focus on liver in this study because it is one of the most important metabolic organs responsible for the significant metabolic shift during hibernation. It is chosen also because of the existence of microarray and proteomic data for the same liver samples through our previous studies (28, 32). Ground squirrel species are only distantly related to other organisms whose miRNAs have been studied (Fig. 1). Recently, a low-coverage genome of thirteen-lined ground squirrel (*S. tridecemlineatus*) has become available through the mammalian genome project. AGS is a close-related species of thirteen-lined ground squirrel as they split only ~4.6 million years ago (Mya) (6). The mapping of Illumina sequence reads from the small RNAs in AGS onto the thirteen-lined ground squirrel genome can help to identify the novel and fast-evolving miRNAs in ground squirrel species. Information on ground squirrel miRNAs will also contribute to the study of miRNA evolution in mammals.

MATERIALS AND METHODS

Animals. As described in our previous studies (28, 32), *S. parryii* were trapped during July near Toolik Lake (68°N 149°W, elevation 809 m) and transported to the University of Alaska Fairbanks. Animals were housed at 20 ± 2°C with a 16:8-h light-dark photoperiod and provided with Mazuri rodent chow and water ad libitum. Hibernating animals were implanted with temperature-sensitive radio transmitters and maintained at a 4:20-h photoperiod. Body temperature (T_b) was monitored for precise stage of torpor and arousal by an automated telemetry system that measures and records T_b every 10 min (8). Animals late in a torpor bout (LT; $n = 4$) were collected after 80–90% of the duration of the previous torpor bout (8–12 days). Animals sampled early after spontaneously arousing from torpor (EA; $n = 4$) were collected 1–2 h after core T_b had increased above 30°C during rewarming. Postreproductive summer euthermic animals sampled in May and June 1–2 mo after ending hibernation (PR; $n = 6$) were used as nonhibernating controls. These animals had completed

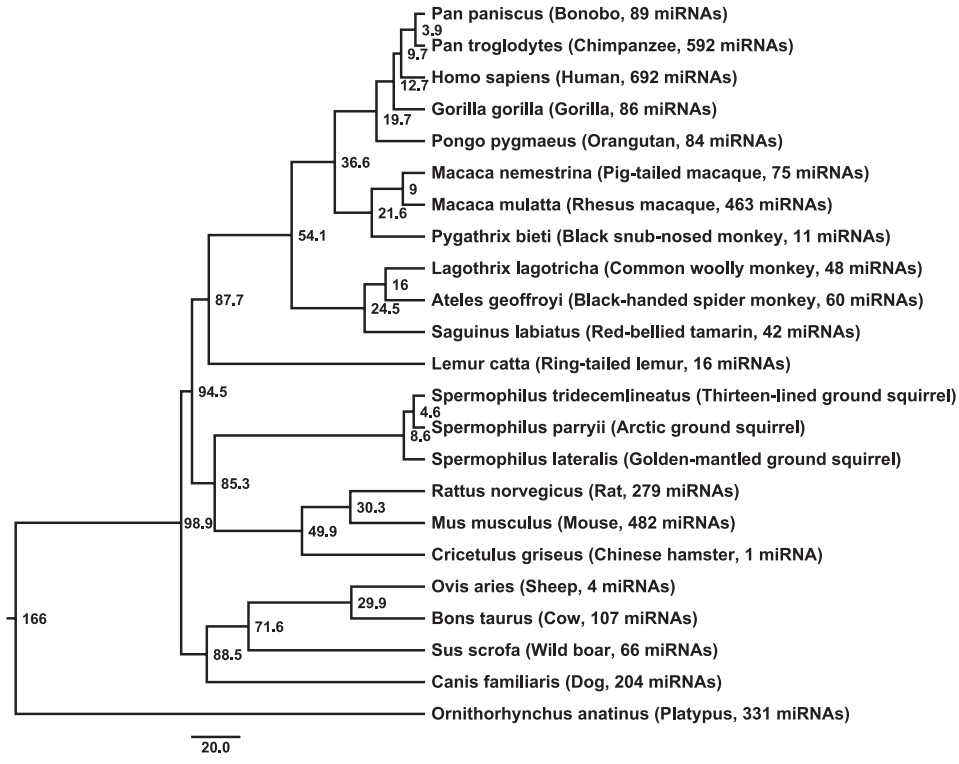

Fig 1. The phylogenetic relationship of ground squirrel species with other mammal species whose microRNAs (miRNAs) are present in miRBase 12.0. The tree is constructed based on data from Bininda-Emonds et al. (6). The lengths of tree branches are in millions of years ago (Mya).

reproductive regression as assessed by external inspection of gonads and genitalia and had entered molt. Torpid animals were euthanized by decapitation without anesthesia. Aroused and postreproductive animals were deeply anesthetized with isoflurane and decapitated. Liver tissue was rapidly dissected, frozen in liquid N_2 and stored at −80°C. Animal protocols were approved by the University of Alaska Fairbanks Institutional Animal Care and Use Committee.

Small RNA library preparation and sequencing. Liver tissue samples from 14 AGS were used in this study including four early aroused animals (EA), four late torpid animals (LT), and six postreproductive animals (PR) as nonhibernating animals. Total RNA from each tissue sample was extracted using TRIzol (Invitrogen) according to the manufacturer's instructions. Total RNA quality was assessed by Bioanalyzer 2100 (Agilent). Total RNA of AGS from the same stage were mixed in the equal amount into three pooled samples: EA, LT, and PR. We loaded ~10 μg total RNA from each pooled samples onto a denaturing 15% polyacrylamide gel. The nucleotides from 18 to 30 nt in length were excised, and RNA was eluted for 4 h with 0.3 M NaCl at room temperature. The small RNAs were recovered by ethanol precipitation. The small RNAs were then ligated sequentially to 5′ and 3′ RNA/DNA chimeric oligonucleotide adapters (TCGTATGCCGTCTTCTGCTTGT). Reverse transcription was preformed after ligation with adapters, followed by 15 cycles PCR amplification. The resulting PCR products were sequenced following the Solexa Genome Analysis User Guide (Illumina) with no modifications.

Illumina sequence data analysis. All known miRNAs in human, mouse, rat, and dog were also aligned to ground squirrel genome by National Center for Biotechnology Information (NCBI) basic local alignment tool (BLAST, parameters: -p blastn -W 11, with all default parameters). RepeatMasker was run with parameters: (-species rodentia -e wublast). 327 putative miRNA precursors on ground squirrel genome were labeled with prefix "mir_clst_" (Supplemental Table S1).[1] The unique unambiguous Illumina sequences were aligned to thirteen-lined ground squirrel genome by NCBI MegaBLAST (parameters: -W 18 -p 100, with all default parameters). The regions containing "N"s on three miRNA precursors (mir_clst_4, mir_clst_90, mir_clst_297) were extended by Illumina sequence reads. To identify 5′ or 3′ mature miRNAs, we divided the regions on the miRNA precursors into 5′ and 3′ arms relative to the center of precursors. We defined the 5′ end position on the miRNA precursor mapped with the most sequence reads on the 5′ or 3′ arm as the representative 5′ end of the 5′ or 3′ mature miRNAs, respectively. The sequence stretch mapped on the miRNA precursor starting from the representative 5′ end of mature miRNAs with the most sequence reads was defined as representative 5′ or 3′ mature miRNA sequences, respectively. To remove the accidental hits, only the mature miRNAs with at least 10 Illumina sequence reads were kept. The secondary structures of putative ground squirrel precursors were identified using RNA folding program (RNAfold parameters: -noPS -d2 -noLP). miRNA precursors with invalid secondary structure were discarded after manual examination. The miRDeep program was run with parameters: (-c 10 and default parameters). While quantifying miRNA expression, all sequence reads mapped on 5′ or 3′ arms of the precursors were summed up as the sequence reads on corresponding 5′ or 3′ mature miRNA, respectively, to represent their expression level. As indicated by the total number of reads mapped onto genome, the variation of total miRNA expression between states is very small, for example LT is only 3.5% greater than EA and PR (Table 1). Only mature miRNAs with at least 40 reads were used for miRNA quantification. To evaluate the statistical significance of differential miRNA expression in Illumina data, we used Fisher's exact test to compute the *P* values between any two stages of EA, LT, and PR stages. The differences in total miRNA amounts between states have been taken into account automatically in Fisher's exact test. The minimum of *P* values in three pair-wise comparisons was used as a measure of overall significance as we aimed to select the miRNAs where at least one pair-wise comparison was significant.

miRNA microarray. miRNA microarrays were manufactured by Agilent Technologies, and each contains probes for 567 mouse and 10 viral microRNAs from miRBase v10.1. Labeling and hybridization of total RNA were performed according to the manufacturer's protocol with no modification. We used 100 ng total RNA from three pooled samples as input in the labeling reaction. The entire reaction was hybridized to the array for 20 h at 55°C. Microarray results were extracted using Agilent Microarray Scanner (Agilent) and Agilent Feature Extraction software (v9.5.3.1), subsequently analyzed by Gene-Spring GX 7.3.1 software (Agilent).

miRNA microarray data analysis. The miRNA signals and errors on miRNA microarray were obtained from gTotalGeneSignal and gTotalGeneError respectively in the output files of Agilent Feature Extraction software. Comparing the miRNA expression between two stages i and j, the statistical significance of differential expression was measured by the statistic

$$u = \frac{x_i - x_j}{\sqrt{\varepsilon_i^2 + \varepsilon_j^2}}$$

where x_i and x_j are total signals, ε_i and ε_j are total errors, and i and j denote any two stages in EA, LT, and PR stages. *P* values were computed by $P(|x| > u)$, where $P(x)$ is the probability distribution of standard normal distribution.

Quantitative miRNA real-time PCR assay. Taqman miRNA real-time PCR assays (Applied Biosystems) were used to quantify the mature miRNA expression following the protocols by Applied Biosystems. SnoRNA Sno135 (Applied Biosystems) was used as the endogenous control for miRNA expression assay. Each reverse transcription (RT) reaction consisted of 500 ng of total RNA, 1× RT buffer, dNTPs (each at 1 mM), 50 U/μl MultiScribe Reverse transcriptase, 50 nM stem-loop RT primer and 0.38 U/μl RNase inhibitor (Applied Biosystems). RT reactions were incubated at 16°C for 30 min, 42°C for 30 min and 85°C for 5 min. Real-time PCR reactions were performed in double replicates in 20 μl volumes. The real-time reaction mix consisted of 1.33 μl RT product, 1 μl of 20× TaqMan miRNA assay mix, 10 μl TaqMan 2× universal PCR Master Mix. Quantitative miRNA real-time PCR assays were conducted on LightCycler 480 real-time PCR System (Roche). 35-C_T was used as the $\log_2$(expression value), where C_T is the cycle number crossing a fixed

[1] The online version of this article contains supplemental material.

Table 1. *Mapping result of Illumina reads on ground squirrel genome*

	Total Reads			Unique Sequences		
Stage	Read Number	Mapped	Percentage	Total Number	Mapped	Percentage
EA	8,788,698	5,611,673	63.85%	1,724,246	255,868	14.84%
LT	8,084,896	5,860,476	72.49%	1,235,120	190,722	15.44%
PR	8,086,791	5,614,441	69.43%	1,336,005	249,950	18.71%
Average	8,320,128	5,695,530	68.45%	1,431,790	232,180	16.22%

EA, early aroused; LT, late torpid; PR, postreproductive.

threshold. ANOVA was used to compute the *P* values of differential miRNA expression in EA, LT, and PR stages.

Real-time PCR with poly(A) tailing for small RNA. Small RNA from each tissue sample was extracted with the mirVana miRNA isolation kit (Ambion) according to the manufacturer's instructions. U6 RNA was used as the endogenous control. Poly(A) tail was added to small RNA with Poly(A) Polymerase (Ambion). Reverse transcription (RT) was performed with the PrimeScript RT reagent Kit (Takara). A primer, containing oligo(dT) and a universal PCR primer sequence, anneals to the poly-adenylated RNA. RT reaction synthesizes cDNA containing the small RNA complement sequence, a poly(dT) tract, and the universal PCR primer sequence. Poly(A) tailing and RT were performed in the same reaction consisted of 500 ng of total RNA, polyA polymerase (Ambion), ATP (10 mM), 5× PrimeScript buffer (contains dNTPs), RT primer (10 μM), PrimeScript RT enzyme (Takara), and 20 U/μl RNasin in 20 μl volumes. Reactions were incubated at 37°C for 60 min, 95°C for 2 min, and 10°C for 30 min. Real-time PCR was performed with Power SYBR Green PCR Master Mix (ABI) in three replicates in 20 μl volumes. The real-time reaction mix consisted of 1 μl RT product, 10 μl of 2× SYBR Green PCR Master Mix (ABI), universal primer (10 μM), and small RNA specific primer (10 μM). Real-time PCR assays were conducted on Applied Biosystems 7500 fast Real-time PCR System (ABI).

miRNA target prediction. Due to the lack of full-length mRNAs available for ground squirrel, we tried to extract ground squirrel 3′ untranslated regions (UTRs) through homology search from other species where 3′ UTRs were known. Human 3′ UTRs were extracted from Ensembl core database (release 51) using Ensembl BioMart 1.2. The genomic repeats on human genome were masked by RepeatMasker (version open-3.2.2) using repeat sequences database for human (RepBase release 20071204). The longest 3′ UTR of each gene in human were selected if multiple 3′ UTRs exist. The repeat masked human 3′ UTRs were aligned to thirteen-lined ground squirrel genome by NCBI BLAST. If the human 3′ UTR was aligned to the downstream of stop codon of its homologous gene in ground squirrel, the sequence from stop codon to the end of the furthest alignment on ground squirrel genome was defined as the 3′ UTR of homologous ground squirrel gene. Due to the large evolutionary distance between human and ground squirrel, the 3′ UTR alignments between human and ground squirrel were generally poor and did not stretch for very long, i.e., ~20% of 3′ UTR in human. So we required the alignment length of 3′ UTR to be at least 20% of 3′ UTR in human to guarantee that we can obtain sufficiently long 3′ UTRs. The 3′ UTRs that aligned to the opposite direction as the mRNA on ground squirrel genome were removed. In the end, we obtained 7,881 ground squirrel 3′ UTRs. To compensate for the lack of 3′ UTR information in ground squirrel, for the miRNAs known to exist in other species, we extracted the targets predicted by TargetScan program (release: 4.1) directly, which considered conservation information spanning a wide range of mammalian species including human, rat, mouse, chicken, and dog. For ground squirrel-specific miRNA (mirdeep-192–5p) and fast-evolving miRNA (spa-miR-506–3p), we used miRanda algorithm (version 4.1) (22, 23) to predict the target sites on ground squirrel 3′ UTRs obtained from human homology search.

Identification of coexpressed miRNA clusters. Human tissue miRNA expression data were obtained from Agilent human miRNA microarray experiment (1), ABI real-time PCR experiment (24), and small RNA library cloning experiment (19) to identify coexpressed miRNAs in human. The coexpressed miRNA pairs were defined by the criteria that the Pearson correlation coefficient of miRNA pair across tissues is larger than a certain cutoff (0.95 in Agilent microarray data, 0.9 in ABI real-time PCR data, and 0.8 in small RNA library cloning data) in at least one dataset. Many miRNAs have been found to be transcribed from one transcript (9, 12, 21). If two human miRNAs fall into the same intergenic or intronic regions on the human genome, they are considered to be colocalized on the human genome. If two ground squirrel miRNAs fall into the same scaffold on the ground squirrel genome, they are considered to be colocalized on the ground squirrel genome. If two miRNAs are identified as coexpressed in human miRNA expression data and colocalized on both human and ground squirrel genomes, they were identified as cotranscribed miRNAs in ground squirrel. We used a relaxed *P*-value cutoff ($P < 0.05$) of statistical significance when we searched for miRNA clusters in the differentially expressed miRNAs in our Illumina and Agilent data.

RESULTS

Small RNA library construction. We constructed small RNA libraries from the AGS livers sampled in three different physiological stages: EA 1–2 h after T_b had increased >30°C during spontaneous arousal from torpor in hibernation, LT after 8–10 days of continuous torpor, and PR after AGS had ended hibernation and passed through the subsequent reproduction phase. These three stages were chosen because EA and LT are the most contrasting stages during the hibernation torpor-arousal cycle according to our previous gene expression study (32), and PR is used as the control group to represent the nonhibernating stage. The same groups of animals were used in

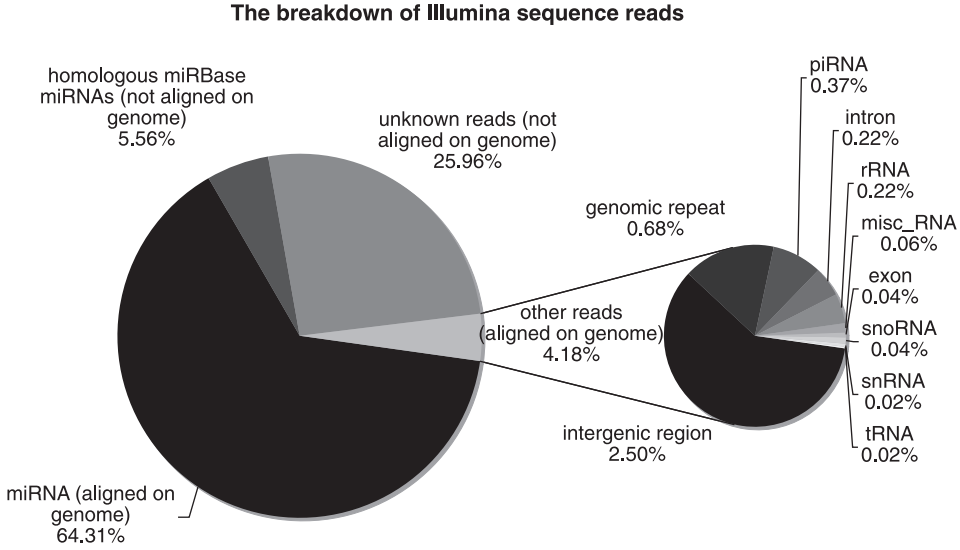

Fig. 2. The distribution of Illumina sequence reads according to their annotations.

our genomics and proteomics studies (28, 32). Massive sequencing of the small RNA libraries by Illumina Genome Analyzer yielded 8,788,698 (EA), 8,084,896 (LT), and 8,086,791 (PR) unfiltered sequence reads.

Ground squirrel genome annotation. We annotated putative miRNA precursors from the ground squirrel genome by aligning miRNA precursors of human, mouse, rat, and dog in miRBase 12.0 (17) onto the thirteen-lined ground squirrel genome. The putative miRNA precursors on the ground squirrel genome were identified to be the mutual best hits with homologous known miRNA precursors on human, mouse, rat, and dog genomes, respectively. In total, 327 putative ground squirrel miRNA precursors were identified corresponding to 319 unique miRNA precursors (Supplemental Table S1). The genomic regions of piRNAs, tRNAs, and rRNAs were identified by aligning the homologous sequences from human onto ground squirrel genome using NCBI BLAST (2). Other annotations for snRNAs, snoRNAs, miscellaneous RNAs such as 7SK and Vault RNAs, integenic regions, exons, and introns on the ground squirrel genome were obtained from Ensembl directly. The genomic repeats on ground squirrel genome were identified by RepeatMasker program (29).

Illumina sequence mapping. We extracted 30 nt sequences from the 5′ ends of Illumina sequence reads. After removing the sequences with low complexity, we aligned them to the thirteen-lined ground squirrel genome with the megablast program, allowing only perfect matches for at least 18 nt extending from the 5′ end; 5,614,441 (PR), 5,611,673 (EA), and 5,860,476 (LT) sequence reads can be mapped onto at least one locus on the ground squirrel genome (Table 1). The mapped sequences were annotated as one of the above-mentioned RNA categories if their 5′ ends fell into the corresponding annotated regions. The distribution of sequence counts in different RNA categories is shown in Fig. 2. A total of 16,050,943 reads from three small RNA libraries can be mapped to at least one putative ground squirrel miRNA precursor, corresponding to 93.9% of total sequence reads mapped on the ground squirrel genome. We mapped 234 putative ground squirrel miRNA precursors with at least one Illumina sequence read (Supplemental Table S2).

The 5′ or 3′ mature miRNAs on miRNA precursors were identified by Illumina reads mapped on the 5′ or 3′ arm of precursors (see MATERIALS AND METHODS). We found 168 putative ground squirrel miRNA precursors contained at least one 5′ or 3′ mature miRNA sequence corresponding to 221 unique mature miRNAs. We obtained the secondary structures of these putative ground squirrel precursors using RNA folding program (RNAfold) (18). After we manually examined the secondary structures of these precursors, we identified 161 precursors with the stem-loop hairpin structures characteristic to miRNA precursors, corresponding to 219 unique mature miRNAs (Supplemental Table S3).

A total of 7,865,844 Illumina sequence reads failed to map onto the ground squirrel genome. To estimate the number of miRNAs that could have been missed in our analysis due to the low coverage of the ground squirrel genome, we aligned these sequence reads onto the miRBase miRNA precursors in human, mouse, rat, and dog, whose homologous miRNA precursors cannot be found in the ground squirrel genome assembly. We can map 1,381,944 sequence reads onto at least one homologous miRBase miRNA precursors. After the redundancy of miRNA precursors between species is removed, 84 unique miRNA precursors including 103 mature miRNAs such as miR-148a and miR-323 contain at least 10 perfect matches of Illumina sequence reads that cannot be aligned on the ground squirrel genome (Supplemental Table S4). They most likely fell into the missing regions on the ground squirrel genome. We sequenced the flanking regions of miR-148a precursor (245 nts, NCBI accession: GU233959) and miR-323 precursor (234 nts, NCBI accession: GU233960), respectively, on the AGS genome using conserved PCR primers between human, mouse, and rat. We verified that these genomic sequences cannot be mapped to ENSEMBL thirteen-lined ground squirrel genome, indicating they were indeed missed in the current genome assembly.

Ground squirrel miRNA annotation. Under the current miRBase naming convention, we named AGS miRNA precursors as *spa-mir-X*, where X is the homologous precursor ID in human, mouse, rat, or dog in miRBase. The mature miRNAs on the arm with predominant expression were annotated as spa-miR-X, while the mature miRNAs on the other arm as spa-

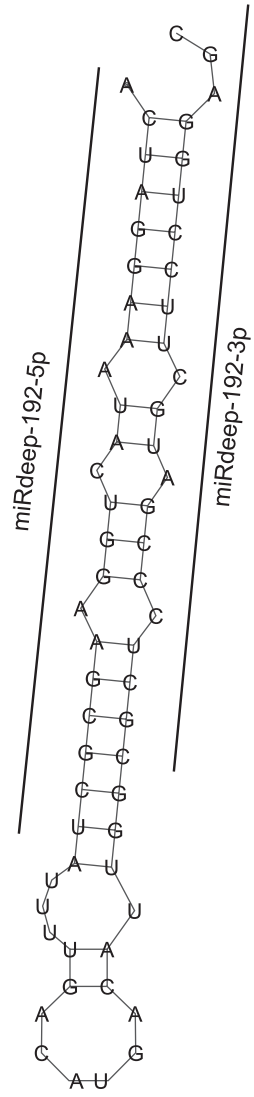

Fig. 3. Secondary structure of the novel ground squirrel specific miRNA (*mirdeep-192*) predicted by miRDeep program.

miR-X*. We defined the predominance of expression if the sequence matches on one arm are at least two times higher than those on the opposite arm. Among the 161 AGS miRNA precursors, 74 of them are 5′ predominant and 70 of them are 3′ predominant; 17 of them did not show 5′ or 3′ predominance and were annotated as spa-miR-X-5p and spa-miR-X-3p accordingly (Supplemental Table S3). Comparing our AGS miRNA annotations with miRNA-miRNA* annotations of their homologous miRNAs in human, mouse, rat, and dog in miRBase, we observed general consistency ($P < 10^{-26}$, Fisher's exact test) with all four species. However, there are also 16 precursors showing opposite miRNA-miRNA* annotation with the miRNAs in at least one of the four species. For example, hsa-miR-223 was annotated in the 3′ and hsa-miR-223* was in the 5′ of precursor in human. However, spa-miR-223 was annotated in the 5′ arm and was five times more abundant than spa-miR-223* in the 3′ arm in AGS liver. In fact, such naming inconsistencies have already existed in miRBase annotations for human, mouse, rat, and dog. This might be due to the species-specific and/or tissue-specific effects. It calls for a more sophisticated miRNA naming system.

Iso-miRs in ground squirrel. The variations of 5′ and 3′ end positions of mature miRNAs on the miRNA precursors can generate so-called iso-miRs, i.e., mature miRNAs with base-shifts (26). As the variations in the 5′ end might lead to different seed regions, we will be mainly concerned with the iso-miRs due to 5′ variation. Typically, the number of sequence reads mapped on the 5′ end of mature miRNA are much more numerous than the neighboring positions. However, there are four miRNA precursors, *spa-mir-101*, *spa-mir-506*, *spa-mir-23b*, and *spa-mir-24,* where equally abundant Illumina reads have been mapped to more than one position close to each other. For example, 68,227 and 66,076 sequence reads were mapped onto two adjacent loci on the 3′ arm of *spa-mir-101* precursor, respectively. This leads to two mature miRNAs with only a 1 nt shift. In miRBase 12.0, the latter position was annotated as the mature sequence of *mir-101* in most species except in chicken where the former position was annotated (*gga-mir-101*). This might lead to multiple iso-miRs with different seeds targeting different genes.

Evolution of miRNAs in ground squirrel. As miRNA mature regions fulfill their ultimate regulatory functions, we expect that changes in these regions will confer the most significant functional changes in miRNA evolution. We compared the mature miRNA sequences that we identified in ground squirrel with known mature miRNA sequences in human. Ten ground squirrel miRNAs contained small numbers of nucleotide shifts of their 5′ ends relative to human homologous miRNAs, including seven with a 1 nt shift, two with a 2 nt shift, and one with a 3 nt shift. Such shifts of 5′ end of miRNAs can lead to miRNAs with new seed regions targeting different sets of genes in ground squirrel. Furthermore, we identified a total of 16 substitutions within 14 mature ground squirrel miRNAs compared with their homologous miRNAs in human. The distribution of these substitutions relative to 5′ end of mature miRNA is shown in Supplemental Fig. S1. Only two such substitutions occur inside of seed regions that might significantly affect the target genes. We further divided the substitutions into those that fall into 5′ and 3′ mature regions and nonmature regions on the miRNA precursor, respectively. The substitutions showed significant preferences in nonmature regions compared with mature regions ($P < 2.2 \times 10^{-16}$, Fisher's exact test), indicating much stronger selection pressure in the mature regions compared with nonmature regions.

The miRNA mature sequences contain very limited substitutions across species and therefore are not very useful in inferring their evolutionary history. On the other hand, miRNA

Table 2. *Top 20 most significant ground squirrel miRNAs identified in Illumina sequencing as differentially expressed during hibernation*

spa_id[a]	hsa_id[b]	mmu_id[c]	rno_id[d]	cfa_id[e]	EA[f]	LT[g]	PR[h]	pval_min[i]	max_logFC[j]
spa-miR-378	hsa-miR-378	mmu-miR-378	rno-miR-378	cfa-miR-378	17896	18227	25815	0	0.53
spa-miR-1	hsa-miR-1	mmu-miR-1	rno-miR-1	cfa-miR-1	17001	12509	8772	0	0.95
spa-miR-184	hsa-miR-184	mmu-miR-184	rno-miR-184	NA	1467	2489	4206	2.45E-301	1.52
spa-miR-143	hsa-miR-143	mmu-miR-143	rno-miR-143	cfa-miR-143	12959	8301	10317	2.95E-271	0.64
spa-miR-152	hsa-miR-152	mmu-miR-152	rno-miR-152	cfa-miR-152	2739	2988	5194	1.56E-169	0.92
spa-miR-320	hsa-miR-320a	mmu-miR-320	rno-miR-320	cfa-miR-320	4963	4016	6639	3.18E-169	0.73
spa-miR-181a	hsa-miR-181a	mmu-miR-181a	rno-miR-181a	cfa-miR-181a	6200	3749	3816	1.17E-158	0.73
spa-miR-181b	hsa-miR-181b	mmu-miR-181b	rno-miR-181b	cfa-miR-181b	2742	1648	1699	1.40E-72	0.73
spa-miR-9-5p	hsa-miR-9	mmu-miR-9	rno-miR-9	cfa-miR-9	2067	1384	1399	3.23E-38	0.58
spa-miR-486	hsa-miR-486-5p	mmu-miR-486	NA	NA	269	663	507	1.85E-35	1.30
spa-miR-9-3p	hsa-miR-9*	mmu-miR-9*	rno-miR-9*	NA	1287	783	745	1.09E-33	0.79
spa-miR-451	hsa-miR-451	mmu-miR-451	rno-miR-451	NA	449	766	356	2.44E-31	1.11
spa-miR-10a	hsa-miR-10a	mmu-miR-10a	rno-miR-10a-5p	cfa-miR-10	420	344	681	1.34E-29	0.99
spa-miR-142	hsa-miR-142-5p	mmu-miR-142-5p	rno-miR-142-5p	cfa-miR-142	1773	2479	1666	1.67E-29	0.57
spa-miR-223	hsa-miR-223*	NA	NA	NA	161	27	34	6.36E-26	2.58
spa-miR-151	hsa-miR-151-3p	mmu-miR-151-3p	rno-miR-151*	NA	390	385	686	2.61E-23	0.83
spa-miR-508	hsa-miR-508-3p	NA	NA	NA	221	178	65	3.00E-21	1.77
spa-miR-138	hsa-miR-138	mmu-miR-138	rno-miR-138	cfa-miR-138b	717	432	429	4.63E-20	0.74
spa-miR-506-5p	NA	NA	NA	NA	223	188	76	4.21E-18	1.55
spa-miR-129	hsa-miR-129*	NA	NA	NA	427	229	214	2.11E-17	1.00

Homologous miRNAs in human, mouse, rat, and dog are also shown. [a]spa_id, Arctic ground squirrel (AGS) mature microRNAs (miRNAs) ID annotated as described in the text; [b]hsa_id, homologous mature miRNA ID in human; [c]mmu_id, homologous mature miRNA ID in mouse; [d]rno_id, homologous mature miRNA ID in rat; [e]cfa_id, homologous mature miRNA ID in dog; [f]EA, number of Illumina reads corresponding to the miRNAs in EA stage; [g]LT, number of Illumina reads corresponding to the miRNAs in LT stage. [h]PR, number of Illumina reads corresponding to the miRNAs in PR stage. [i]pval_min, minimal *P* values in pair-wise comparisons; [j]max_logFC, maximum $\log_2$-transformed fold changes in pair-wise comparisons. NA, not available.

Table 3. *Top 20 most significant miRNAs identified in Agilent mouse miRNA microarray as differentially expressed during hibernation*

miRNA Name[a]	*P* Value[b]	max_logFC[c]	EA[d]	LT[e]	PR[f]
mmu-miR-1224	0	1.71	685.2	316.0	209.3
mmu-miR-190b	0	1.75	9.9	33.4	28.0
mmu-miR-211	0	2.18	6.6	9.2	29.9
mmu-miR-709	0	1.95	101.8	41.2	26.4
mmu-miR-144	4.00E-15	1.41	46.5	109.8	41.3
mmu-miR-486	8.79E-13	1.25	100.3	163.4	68.8
mmu-miR-451	5.07E-12	1.17	2753.1	4341.6	1924.5
mmu-miR-494	2.01E-11	1.25	211.5	148.5	89.0
mmu-miR-378	2.39E-10	1.09	66.9	79.9	142.6
mmu-miR-452	1.43E-08	1.08	58.9	46.7	27.8
mmu-miR-500	7.17E-08	1.17	18.0	40.5	35.0
mmu-miR-705	9.36E-08	1.45	18.2	17.9	6.7
mmu-miR-34a	1.83E-07	0.96	55.0	106.9	106.8
mmu-miR-15b	6.79E-07	0.80	335.2	416.5	239.7
mmu-miR-184	2.60E-06	0.72	167.0	213.4	275.1
mmu-miR-200b	3.85E-06	0.97	35.4	30.3	59.3
mmu-miR-31*	4.80E-06	0.79	74.4	76.8	44.3
mmu-miR-449a	5.80E-06	0.78	1441.6	913.9	842.4
mmu-miR-222	3.39E-05	0.72	42.3	49.4	69.7
mmu-miR-574-5p	7.60E-05	0.68	85.3	65.3	53.3

[a]miRNA_name: mouse mature miRNA ID in miRBase 10.1; [b]*P* value, minimal *P* value in pair-wise comparisons; [c]max_logFC, maximum $\log_2$-transformed fold change in pair-wise comparisons; [d]EA, normalized miRNA probe signal in EA stage; [e]LT, normalized miRNA probe signal in LT stage; [f]PR, normalized miRNA probe signal in PR stage.

precursor sequences determine the secondary structures during miRNA maturation and are also under the selection pressure. They contain much more informative substitutions than in mature regions. Thus it is more appropriate to use the substitutions in precursor sequences to characterize how fast the miRNAs are evolving. We generated four-species alignments of homologous miRNA precursor sequences in human, ground squirrel, mouse, and rat. Using {human, {ground squirrel, {mouse, rat}}} as the underlying phylogenetic tree, we partitioned the substitutions among four species into ground squirrel, mouse, rat, mouse-rat common ancestor, and human specific substitutions in the precursor sequences respectively (Supplemental Table S5). In total, we identified 164 substitutions in ground squirrel, 128 in mouse, 172 in rat, 329 in mouse-rat common ancestor, and 163 in human. Among all miRNA precursors, *spa-mir-506* showed significantly higher substitutions in ground squirrel lineage than expected ($P = 0.005$, Fisher's exact test; Supplemental Fig. S2). It appears to be undergoing rapid evolution in the ground squirrel lineage.

miRNA prediction. We applied miRDeep program (13) on the Illumina sequences aligned on thirteen-lined ground squirrel genome to predict putative miRNAs. miRDeep incorporates the position and frequency of sequenced small RNAs with the secondary structure of miRNA precursor and can reliably discover novel miRNAs. The miRDeep program returned 196 unique miRNA precursor sequences, which we named *mirdeep-X* where Xs are numeric numbers (Supplemental Table S6); 162 of them matched identically to our previously annotated miRNA precursors in AGS, leaving 34 RNA sequences to be determined. Seven of them have high sequence similarities to existing noncoding RNAs in other species in Rfam databases including one U2 RNA, two tRNA, and four snoRNA. One of them (*mirdeep-24*) contains sequence with low complexity and is identical to *hsa-mir-1277* in human. We mapped the rest of the 26 putative novel miRNAs to other available mammalian genomes; 18 of them can only be found in ground squirrel genome and thus seemed to be ground squirrel specific. The secondary structure of novel ground squirrel specific miRNA

Table 4. *miRNAs differentially expressed in both Agilent microarray and Illumina sequencing. Pearson correlations were computed between Illumina results and Agilent results across EA, LT, and PR stages*

		Agilent					Illumina				
miRNA Name[a]	Cor[b]	*P* Value[c]	max_logFC[c]	EA[c]	LT[c]	PR[c]	*P* Value[d]	max_logFC[d]	EA[d]	LT[d]	PR[d]
spa-miR-320	1.00	7.37E-04	0.70	31.1	25.3	41.2	3.18E-169	0.73	4963	4016	6639
spa-miR-184	1.00	2.60E-06	0.72	167.0	213.4	275.1	2.45E-301	1.52	1467	2489	4206
spa-miR-378	0.99	2.39E-10	1.09	66.9	79.9	142.6	0.00E+00	0.50	17896	18227	25815
spa-miR-451	0.99	5.07E-12	1.17	2753.1	4341.6	1924.5	2.44E-31	1.11	449	766	356
spa-miR-152	0.99	2.84E-04	0.68	30.6	29.2	46.6	1.56E-169	0.92	2739	2988	5194
spa-miR-142-5p	0.96	4.89E-04	0.77	18.6	31.8	21.0	1.67E-29	0.57	1773	2479	1666
spa-miR-144	0.93	4.00E-15	1.41	46.51	109.82	41.27	8.99E-15	1.52	71	203	113
spa-miR-486	0.56	8.79E-13	1.25	100.26	163.36	68.82	1.85E-35	1.30	269	663	507
spa-miR-1	0.41	5.69E-03	1.44	8.57	12.99	4.80	0.00E+00	0.95	17001	12509	8772

[a]miRNA_name: annotated AGS miRNAs ID; [b]Cor, Pearson correlation coefficient of expression value between Illumina and Agilent result; [c]*P* value, max_logFC, EA, LT, PR in Agilent array: same as described in Table 2; [d]*P* value, max_logFC, EA, LT, PR in Illumina sequencing: same as described in Table 3.

with the highest Illumina sequence reads (*mirdeep-192*) is shown in Fig. 3.

Differential expression of miRNAs from Illumina data. To reliably quantify miRNA expression, we restricted our analysis to miRNAs with at least 40 Illumina sequence reads and used the number of sequence reads mapped onto the miRNAs as the measure of their expression. The most abundant miRNA in Illumina sequence data is spa-miR-122 with 4,067,950 (EA), 4,474,717 (LT), and 4,209,809 (PR) Illumina sequence reads, which account for 74.6% of total Illumina sequence reads mapped to all AGS miRNAs. We tested the differential expression of ground squirrel miRNAs between any two stages among EA, LT, and PR stages (Supplemental Table S7A) and found 32 miRNAs to be differentially expressed under the criteria $P < 10^{-10}$ (Fisher's exact test) and $|\log_2$ fold change$|>0.5$. Both spa-miR-506–5p and spa-miR-506–3p from the fast-evolving *spa-mir-506* showed significant overexpression in both EA and LT during hibernation compared with PR. The top 20 most significant miRNAs are shown in Table 2. The significant differentially expressed novel ground squirrel specific miRNAs predicted by miRDeep program were also shown in (Supplemental Table S7B). Among them, mirdeep-192–5p from ground squirrel specific *mirdeep-192* was significantly overexpressed in EA and LT compared with PR.

Comparison with Agilent microarray and real-time PCR. We used Agilent mouse miRNA microarray covering 567 mouse miRNAs and 10 viral mouse miRNAs in Sanger 10.1 miRBase (30) as an independent miRNA profiling method to search for additional differentially expressed miRNAs. As in Illumina small RNA library construction, three total RNA pools from the AGS in EA, LT, and PR stages were hybridized to three Agilent mouse miRNA microarrays. Agilent microarray detected 204 miRNAs in at least one of three stages, with 181 in EA, 187 in LT, and 167 in PR, and 156 miRNAs were detected in all three stages. Using $P < 0.01$ and $|\log_2$ fold change$|>0.5$ as the criterion (see MATERIALS AND METHODS), we identified 42 differentially expressed miRNAs among three stages (Supplemental Table S8). The top 20 most significant miRNAs are shown in Table 3.

There were 84 miRNAs detected in both Illumina sequencing method (sequence read number ≥ 40) and Agilent mouse miRNA microarray. The log-transformed miRNA sequence counts from Illumina sequencing correlate positively ($r = 0.44$,

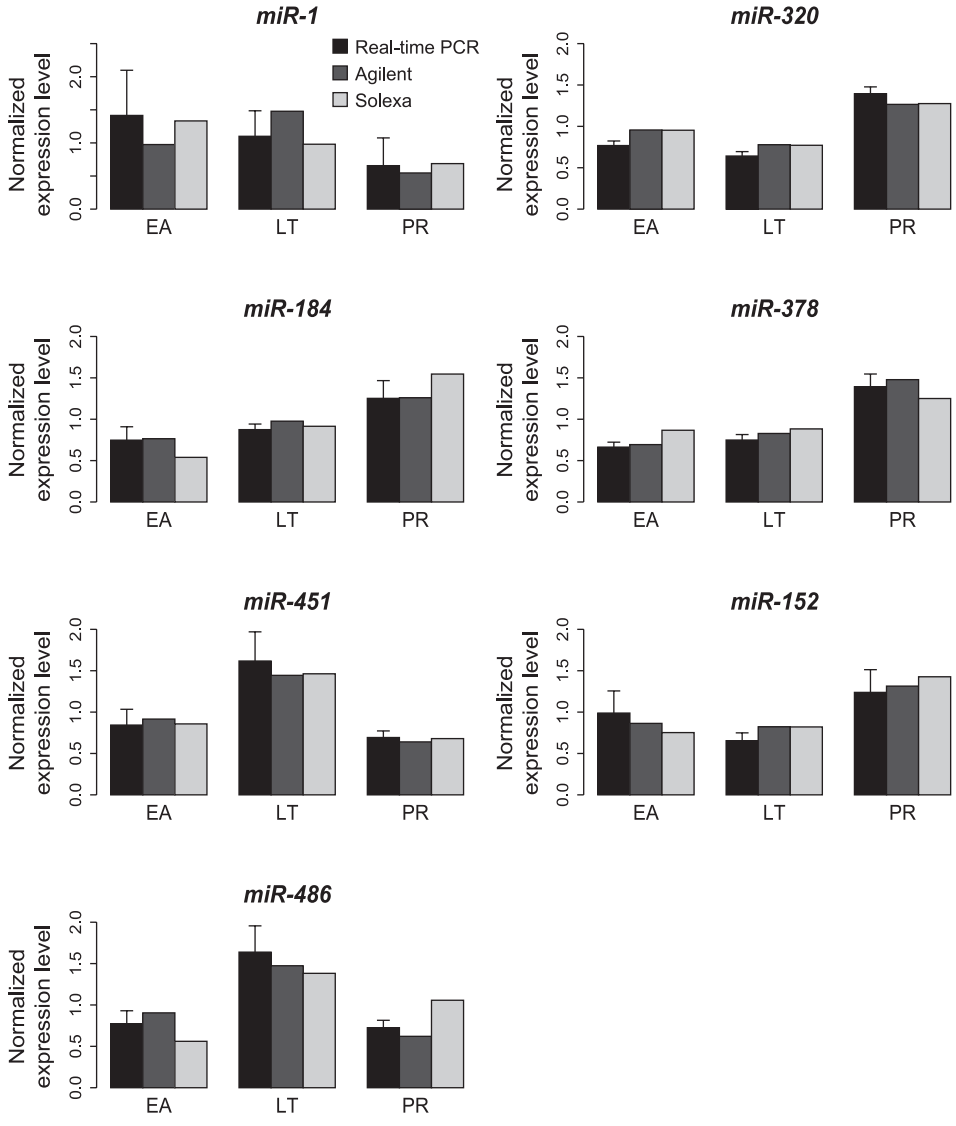

Fig. 4. Real-time PCR validation of differentially expressed miRNAs supported by both Illumina sequencing and Agilent microarray. EA, early aroused; LT, late torpid; PR, postreproductive.

Table 5. *Expression patterns of cotranscribed miRNA clusters in Illumina sequencing data and Agilent microarray data, respectively*

miRNA Clusters	Expression Pattern
Illumina data	
miR-200a, miR-200b	PR>LT~EA
miR-144, miR-451	LT>EA~PR
miR-181a, miR-181b	EA>PR~LT
miR-9, miR-9*	EA>LT~PR
miR-143, miR-145	EA>LT
Agilent data	
miR-27a, miR-23a	PR>EA~LT
miR-144, miR-451	LT>EA~PR
miR-200a, miR-200b	PR>LT~EA
miR-20a, miR-19a, miR-19b	EA~LT>PR

A>B, the miRNA expression in *stage A* is significantly higher ($P < 0.05$) than in *stage B*; A~B, the miRNA expression is not significantly different between *stages A* and *B*, where A and B can be the EA, LT, or PR stage.

$P = 3.2 \times 10^{-5}$) with the log-transformed normalized intensities on the Agilent mouse miRNA microarray. Among the differentially expressed miRNAs identified in pair-wise comparison by at least one technology, the log fold-changes between the two technologies are also positively correlated with $r = 0.78$ ($P = 0.005$) in EA vs. LT, $r = 0.60$ ($P = 0.002$) in EA vs. PR, and $r = 0.88$ ($P = 4 \times 10^{-7}$) in LT vs. PR (Supplemental Fig. S3). There are nine miRNAs identified as the significantly differentially expressed miRNAs in both Illumina sequencing and Agilent miRNA microarray (Table 4). All of them showed consistent changes among the three stages between both technologies.

We tested the expression of 15 miRNAs identified as significantly differentially expressed either in Illumina sequencing or Agilent microarray technologies using ABI Taqman miRNA real-time PCR assay. The results are shown in Supplemental Table S9. The differentially expressed miRNAs supported by both technologies have a higher success rate to be validated compared with those only supported by one technology. As illustrated in Fig. 4, we validated that spa-miR-320 and spa-miR-378 were significantly ($P < 0.05$) underexpressed in EA and LT vs. PR, whereas spa-miR-486 and spa-miR-451 were significantly overexpressed in LT vs. PR but returned in EA to a level similar to that in PR. Although spa-miR-1, spa-miR-184, and spa-miR-152 were not considered as significant in real-time PCR, their changes of expression were still highly consistent with the results in Illumina sequencing and Agilent microarray. The differential expression of spa-miR-181a supported by Illumina sequencing alone and spa-miR-211, spa-miR-452, and spa-miR-190b supported by Agilent microarray alone were also validated in real-time PCR.

We validated the existence of ground squirrel-specific miRNA (mirdeep-192-5p) and its differential expression during hibernation using poly(A) tailing reverse transcription and SYBR Green Real-time PCR. mirdeep-192-5p can be detected in AGS liver but not in mouse liver. It is also mildly ($P = 0.07$) overexpressed in EA and LT vs. PR consistent with Illumina sequencing result. We also tested spa-miR-506-3p, with no available primers from ABI, using poly(A) tailing real-time PCR. It was significantly ($P = 0.01$) overexpressed in EA vs. PR but showed no significant difference between LT vs. PR.

DISCUSSION

AGS liver miRNA profiles. In our study, spa-miR-122 accounted for 74.6% of miRNA counts mapped on the ground squirrel genome. This result is consistent with miRNA expression in human liver where, through sequencing of small RNA cloning libraries, hsa-miR-122 accounts for ~72% of all miRNA counts and is the most abundant miRNA among all human tissues (16, 17). Unlike mRNA expression, miRNA expression in a certain tissue can be dominated by one or a few miRNAs. We also cross-compared our AGS liver results on Agilent mouse miRNA microarray with human liver results on Agilent human miRNA microarray (1). There are 81 miRNAs with identical mature sequences that can be detected in both arrays, and the global expression profiles in AGS and human livers are highly correlated ($r = 0.77$). However, 57 of 81 miRNAs showed significant ($P < 0.01$) differential expression between human and AGS. For example, miR-30c and miR-221 are significantly ($P < 10^{-8}$) overexpressed in AGS liver compared with human liver, whereas miR-145 and miR-143 are significantly underexpressed, reflecting substantial miRNA expression differences between two species.

Cotranscribed miRNA clusters. Several miRNAs often can be transcribed from the same primary transcript. To identify these cotranscribed miRNA clusters, we first combined multiple human tissue miRNA expression data including that from Agilent human miRNA microarray experiments (1), ABI real-time PCR experiments (24), and small library cloning experi-

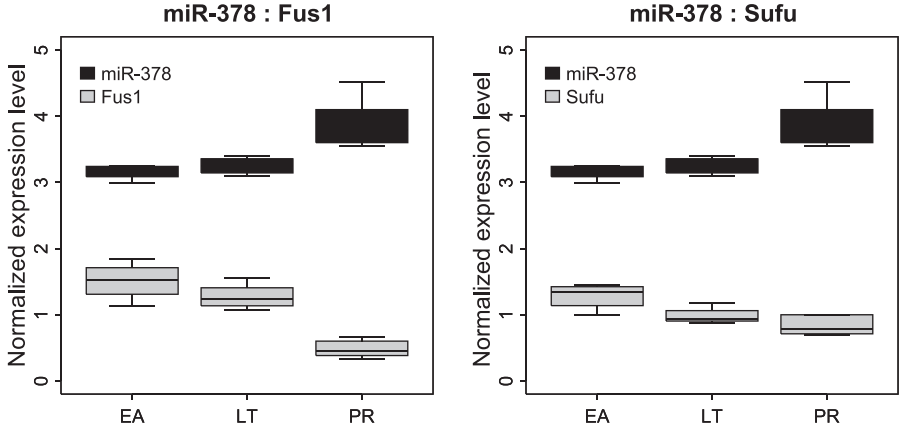

Fig. 5. The expression patterns of miR-378 and its targets: Fus1 and Sufu.

Table 6. *Enriched GO categories of the predicted targets of differentially expressed miRNAs during hibernation with FDR <0.05*

GO Category	FDR
Spa-miR-1	
GO:0019222_regulation_of_metabolic_process	0
GO:0016481_negative_regulation_of_transcription	0
GO:0010467_gene_expression	0
GO:0045934_negative_regulation_of_nucleobase__nucleoside__nucleotide_and_nucleic_acid_metabolic_process	0
GO:0043283_biopolymer_metabolic_process	4.35E-04
GO:0048523_negative_regulation_of_cellular_process	8.33E-04
GO:0016043_cellular_component_organization_and_biogenesis	1.92E-03
GO:0030154_cell_differentiation	3.67E-03
GO:0032502_developmental_process	4.38E-03
GO:0045892_negative_regulation_of_transcription__DNA-dependent	4.24E-03
GO:0016192_vesicle-mediated_transport	2.11E-02
GO:0048634_regulation_of_muscle_development	2.46E-02
GO:0048514_blood_vessel_morphogenesis	2.40E-02
GO:0048646_anatomical_structure_formation	3.19E-02
GO:0045787_positive_regulation_of_cell_cycle	4.44E-02
spa-miR-486	
GO:0043283_biopolymer_metabolic_process	0
GO:0050789_regulation_of_biological_process	3.00E-02
spa-miR-211	
GO:0045449_regulation_of_transcription	0
GO:0016070_RNA_metabolic_process	0
GO:0043283_biopolymer_metabolic_process	0
GO:0019222_regulation_of_metabolic_process	0
GO:0016044_membrane_organization_and_biogenesis	4.50E-03
GO:0016192_vesicle-mediated_transport	4.29E-03
GO:0032502_developmental_process	1.04E-02
GO:0006897_endocytosis	1.68E-02
GO:0031365_N-terminal_protein_amino_acid_modification	4.41E-02
spa-miR-378	
GO:0050789_regulation_of_biological_process	0
GO:0043170_macromolecule_metabolic_process	0
GO:0030518_steroid_hormone_receptor_signaling_pathway	1.18E-02
GO:0016331_morphogenesis_of_embryonic_epithelium	2.35E-02
GO:0018193_peptidyl-amino_acid_modification	3.21E-02
GO:0007243_protein_kinase_cascade	3.43E-02
GO:0007167_enzyme_linked_receptor_protein_signaling_pathway	3.32E-02
GO:0048598_embryonic_morphogenesis	4.33E-02
GO:0010001_glial_cell_differentiation	4.38E-02
GO:0016540_protein_autoprocessing	4.31E-02
GO:0007275_multicellular_organismal_development	4.16E-02
GO:0051726_regulation_of_cell_cycle	4.13E-02
GO:0007242_intracellular_signaling_cascade	4.03E-02
spa-miR-320	
GO:0065007_biological_regulation	0
GO:0043283_biopolymer_metabolic_process	0
GO:0006350_transcription	0
GO:0007399_nervous_system_development	0
GO:0044237_cellular_metabolic_process	7.50E-03
GO:0007155_cell_adhesion	1.21E-02
GO:0000080_G1_phase_of_mitotic_cell_cycle	2.20E-02
GO:0006464_protein_modification_process	3.98E-02
GO:0032502_developmental_process	4.18E-02
GO:0006793_phosphorus_metabolic_process	4.28E-02
mirdeep-192	
GO:0006464_protein_modification_process	0
spa-miR-506-3p	
GO:0007242_intracellular_signaling_cascade	0
GO:0006464_protein_modification_process	0
GO:0006810_transport	0
GO:0051179_localization	1.43E-03
GO:0045941_positive_regulation_of_transcription	2.00E-03
GO:0031047_gene_silencing_by_RNA	1.44E-02
GO:0007275_multicellular_organismal_development	1.56E-02
GO:0043283_biopolymer_metabolic_process	2.65E-02
GO:0032502_developmental_process	3.90E-02
GO:0031325_positive_regulation_of_cellular_metabolic_process	3.91E-02

GO, Gene Ontology; FDR, false discovery rate.

ments (19) in human. The coexpression of miRNAs was integrated with miRNA genomic locations on the human genome to identify 66 cotranscribed miRNA clusters in human. Of these, 34 were also colocalized on the ground squirrel genome. Seven of these miRNA clusters contain at least one miRNA showing differential expression in our Illumina sequencing data and/or Agilent microarray data (Table 5). The significant miRNAs within each miRNA cluster all showed consistent differential expression patterns during hibernation. This suggests that the cotranscriptions of these miRNAs have been conserved in ground squirrel species.

Rapid evolution of mir-506 family. In human, *mir-506* is a members of an X chromosome-linked miRNA cluster including *mir-506*, *mir-507*, *mir-508*, *mir-509*, *mir-510*, *mir-513*, and *mir-514*. This miRNA cluster was first identified by Bentwich et al. (5) in the testis of primates. Later, Zhang et al. (35) reported rapid evolution of this miRNA cluster in primates. In fact, *mir-509*, *mir-510*, *mir-513*, and *mir-514* have multiple copies that are highly variable among primates, indicating their recent duplications and expansions in the primate lineage. miRNAs including *mir-506*, *mir-507*, and *mir-508* have only one copy in primate genomes and their homologous sequences can be found in both mouse and dog indicating a more ancient history. In addition to *spa-mir-506*, we also identified *spa-mir-507*, *spa-mir-508*, and *spa-mir-509* in ground squirrel. The relative genomic positions of *spa-mir-506, spa-mir-507*, and *spa-mir-508* are conserved among primates, mouse, dog, and ground squirrel. Both spa-miR-508 and spa-miR-509 showed overexpression in both EA and LT during hibernation compared with PR, similar to spa-miR-506–5p and spa-miR-506–3p. Therefore, *spa-mir-506*, *spa-mir-507*, *spa-mir-508*, and *spa-mir-509* are likely cotranscribed from the same transcript. Like *spa-mir-506*, the mature and precursor sequences of *spa-mir-507* and *spa-mir-508* have also undergone significant changes. Comparative analysis of their miRNA precursor sequences suggests that the rapid evolution of this X chromosome-linked miRNA cluster may be also occurring in ground squirrel species. Expression profiles of these miRNAs in human tissues showed that they were only expressed in male testis and female ovary (1). Zhang et al. (35) showed they were all underexpressed in the testis of adult compared with infant primates and suggested that they could play an important role in testis development. In our study, they were detected in liver for the first time, indicating their functions in development may go beyond testis. Their significant overexpression could inhibit cell growth and differentiation during hibernation, which is consistent with our previous gene and protein expression studies.

Comparison between miRNA profiling platforms. In this study, we applied three independent methods: the next-generation Illumina/Solexa deep-sequencing, Agilent miRNA microarray, and real-time PCR technologies to investigate the miRNA profiles in an extreme hibernator, the AGS. The next-generation sequencing technology has reached unprecedented depth compared with conventional sequencing methods. It helped us to discover hundreds of miRNAs including

Table 7. *The miRNA target genes that showed negative correlations between miRNA and gene expression at protein level and mRNA level*

miRNA ID[a]	Target Gene Symbol[b]	LT_EA_p[c]	PR_EA_p[d]	PR_LT_p[e]	mean_EA[f]	mean_LT[g]	mean_PR[h]	*P* Value[i]	Cor[j]
				Proteomics result					
spa-miR-1	GCLC	0.11	0.01	0.07	2.69E-05	4.75E-05	1.24E-04	3.44E-02	−0.61
spa-miR-1	HSPD1	0.26	0.05	0.20	5.93E-03	6.68E-03	7.83E-03	3.68E-02	−0.61
spa-miR-378	HDLBP	0.51	0.18	0.50	3.10E-04	2.76E-04	2.32E-04	1.68E-02	−0.67
spa-miR-506-3p	SHMT2	0.44	0.03	0.10	8.37E-04	9.37E-04	1.35E-03	2.04E-02	−0.66
spa-miR-506-3p	AOX1	0.06	7.94E-04	1.21E-03	1.20E-04	6.25E-05	5.84E-04	2.28E-02	−0.65
spa-miR-506-3p	RET	0.08	0.04	0.64	1.55E-05	3.32E-05	4.23E-05	2.40E-02	−0.64
spa-miR-506-3p	DAK	0.50	0.01	0.01	9.56E-04	8.70E-04	1.89E-03	4.35E-02	−0.59
spa-miR-506-3p	ACADSB	0.08	0.02	0.01	1.30E-03	9.30E-04	2.07E-03	4.53E-02	−0.59
mirdeep-192-5p	GSTM3	0.75	0.01	0.03	7.87E-05	7.37E-05	2.25E-04	3.13E-03	−0.77
mirdeep-192-5p	MAT1A	0.55	0.14	0.13	1.25E-03	1.16E-03	1.59E-03	5.29E-03	−0.75
mirdeep-192-5p	FYCO1	0.04	0.05	0.03	1.91E-05	2.67E-06	4.44E-05	5.67E-03	−0.74
mirdeep-192-5p	BHMT2	0.02	0.08	0.01	5.73E-04	2.79E-04	8.24E-04	2.60E-02	−0.64
				Real-time PCR result					
spa-miR-506-3p	SORD	0.36	0.10	0.15	6.29E +06	7.26E +06	9.04E +06	2.53E-02	−0.64
spa-miR-506-3p	PAH	0.13	0.06	0.01	3.86E +06	2.80E +06	6.25E +06	3.72E-02	−0.60
spa-miR-506-3p	AOX1	0.63	1.81E-03	6.80E-04	1.14E +03	1.06E +03	1.31E +04	4.63E-02	−0.58
spa-miR-506-3p	ACADSB	0.68	0.01	0.01	1.38E +04	1.30E +04	3.92E +04	4.83E-02	−0.58
mirdeep-192-5p	ARNTL	0.16	0.09	0.03	7.92E +04	5.92E +04	1.20E +05	3.37E-02	−0.61
				Illumina result					
spa-miR-506-3p	SORD	0.02	1.34E-03	0.03	4.24E-01	7.33E-01	1.17E +00	9.70E-04	−0.82
spa-miR-506-3p	AGPAT3	0.07	7.79E-04	3.41E-03	1.37E-01	2.16E-01	5.42E-01	9.38E-03	−0.71
spa-miR-506-3p	GPD1	0.01	0.01	0.95	4.41E-02	1.29E-01	1.28E-01	3.02E-02	−0.62
spa-miR-506-3p	GCK	0.85	4.30E-04	2.68E-04	−4.44E-03	1.36E-02	1.50E-01	3.60E-02	−0.61

Pearson correlations were computed between miRNA expression (real-time PCR result) and gene expression at the protein (proteomics result) and mRNA (real-time PCR and Illumina BeadArray results) levels. [a]miRNA ID, AGS miRNA ID; [b]Target Gene Symbol, gene symbol of predicted miRNA target; [c]EA_LT_p, significant *P* value in EA vs. LT comparison; [d]EA_PR_p, significant *P* value in EA vs. PR comparison. [e]LT_PR_p, significant *P* value in LT vs. PR comaprison; [f]mean_EA, the average expression level for EA; [g]mean_LT, the average expression level for LT; [h]mean_PR, the average expression level for PR; [i]*P* value, *P* value of association test between miRNA expression and protein/mRNA expression. [j]Cor, Pearson correlation coefficient.

ground squirrel-specific and fast-evolving miRNAs. However, its capability for miRNA quantification has not been fully compared with more traditional hybridization-based high-throughput miRNA microarrays and real-time PCR. Here we compared the results of Illumina sequencing and Agilent miRNA microarray technologies on the common samples. Our results showed that there is only limited overlap in terms of the number of significant miRNAs identified by both Illumina sequencing and Agilent microarray technologies, but the consistency is very high for the overlapping miRNAs. The differentially expressed miRNAs supported only by Illumina sequencing have a low success rate in real-time PCR validation. This could be due to the fact that the rapid amplification step during Illumina sequencing can potentially introduce biases that may skew the real biological differences between samples. Another explanation is that only pooled samples were used in Illumina sequencing, whereas each individual animal was tested in real-time PCR. Because of the expense and limited number of samples per run in next-generation sequencing technology, pooling samples from the same physiological state has been a routine practice in such experiments. Our study suggests that there is a tradeoff between cost and precision in next-generation sequencing experiments and points out the importance of biological replicates during miRNA quantification.

Potential functions of differentially expressed miRNAs. The miRNAs identified to be differentially expressed during hibernation in this study can be largely divided into two groups according to their expression patterns: spa-miR-211, spa-miR-378, spa-miR-184, spa-miR-200a, and spa-miR-320 are significantly underexpressed in EA and LT vs. PR; spa-miR-144, spa-miR-486, spa-miR-451, and spa-miR-142 are significantly overexpressed in LT vs. PR but return in EA to the levels similar to those in PR. Previous studies showed that the inhibition of hsa-miR-211 can cause a decrease in cell growth in the HeLa cell (11). Lee et al. (20) also showed that miR-378 can promotes cell survival, tumor growth, and angiogenesis by targeting SuFu and Fus-1 expression. We used real-time PCR to measure the mRNA expression changes of SuFu and Fus-1 among the same animals in three stages of hibernation. SuFu and Fus-1 showed significant overexpression in both EA and LT vs. PR, exactly opposite to the expression change in spa-miR-378 during hibernation (Fig. 5). These indicate that the underexpression of spa-miR-211 and spa-miR-378 may suppress tumor genesis during hibernation. There are also studies showing that miR-486, miR-451, miR-144, and miR-142 inhibit tumor growth (14, 27). Together, the synergistic effects of two groups of miRNAs may help to suppress tumor growth during torpor and resume cell cycle during arousal.

To search for the potential broader functions of these miRNAs, we predicted the putative targets of the most significantly differentially expressed miRNAs using Targetscan. Targets of ground squirrel-specific miRNA (mirdeep-192–5p) and fast-evolving miRNA (spa-miR-506–3p) were predicted using MIRANDA program (MATERIALS AND METHODS). Gene Ontology (GO) analysis using GoMiner (34) identified enriched functional groups (false discovery rate <0.05) among miRNA putative targets (Table 6). As miRNA targets may be affected at the mRNA and protein levels, we searched for the genes showing significant differential expression at the protein level from our previous proteomics result (28) and at the mRNA level from our real-time PCR result (28) and Illumina BeadArray result (32) among miRNA putative targets. As these experiments have been conducted on the same set of animals, we computed the Pearson correlations between miRNA expression and their putative target gene expression at the protein and mRNA levels. The overlapped genes that also showed negative correlations between miRNA and gene expression are more likely to be the true targets of miRNAs. They are listed in Table 7. However, future functional studies will be needed to reveal the specific functions of differential expressed miRNAs in animal hibernation.

ACKNOWLEDGMENTS

Present address for W. Chen: Berlin Inst. for Medical Systems Biology, Max-Delbrück-Center for Molecular Medicine, Berlin 13125, Germany.

GRANTS

We acknowledge support from National Basic Research Program of China Grant 2006CB910700, Shanghai Science and Technology Committee Grant 08QA1407500 (J. Yan), SA-SIBS Scholarship Program (J. Yan), National Science Foundation Grants 0076039 and 0732755, and US Army Medical Research and Materiel Command Grant #05178001 (B. M. Barnes).

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

REFERENCES

1. **Ach RA, Wang H, Curry B.** Measuring microRNAs: comparisons of microarray and quantitative PCR measurements, and of different total RNA prep methods. *BMC Biotechnol* 8: 69, 2008.
2. **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. *Nucl Acids Res* 25: 3389, 1997.
3. **Barnes BM.** Freeze avoidance in a mammal: body temperatures below 0 degree C in an Arctic hibernator. *Science* 244: 1593–1595, 1989.
4. **Bartel DP.** MicroRNAs: genomics, biogenesis, mechanism, function. *Cell* 116: 281–297, 2004.
5. **Bentwich I, Avniel A, Karov Y, Aharonov R, Gilad S, Barad O, Barzilai A, Einat P, Einav U, Meiri E, Sharon E, Spector Y, Bentwich Z.** Identification of hundreds of conserved and nonconserved human microRNAs. *Nat Genet* 37: 766–770, 2005.
6. **Bininda-Emonds OR, Cardillo M, Jones KE, MacPhee RD, Beck RM, Grenyer R, Price SA, Vos RA, Gittleman JL, Purvis A.** The delayed rise of present-day mammals. *Nature* 446: 507–512, 2007.
7. **Brauch KM, Dhruv ND, Hanse EA, Andrews MT.** Digital transcriptome analysis indicates adaptive mechanisms in the heart of a hibernating mammal. *Physiol Genomics* 23: 227–234, 2005.
8. **Buck CL, Barnes BM.** Effects of ambient temperature on metabolic rate, respiratory quotient, and torpor in an arctic hibernator. *Am J Physiol Regul Integr Comp Physiol* 279: R255–R262, 2000.
9. **Cai X, Hagedorn CH, Cullen BR.** Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA* 10: 1957–1966, 2004.
10. **Carey HV, Andrews MT, Martin SL.** Mammalian hibernation: cellular and molecular responses to depressed metabolism and low temperature. *Physiol Rev* 83: 1153–1181, 2003.
11. **Cheng AM, Byrom MW, Shelton J, Ford LP.** Antisense inhibition of human miRNAs and indications for an involvement of miRNA in cell growth and apoptosis. *Nucl Acids Res* 33: 1290–1297, 2005.
12. **Cullen BR.** Transcription and processing of human microRNA precursors. *Mol Cell* 16: 861–865, 2004.
13. **Friedlander MR, Chen W, Adamidi C, Maaskola J, Einspanier R, Knespel S, Rajewsky N.** Discovering microRNAs from deep sequencing data using miRDeep. *Nat Biotech* 26: 407–415, 2008.
14. **Gal H, Pandi G, Kanner AA, Ram Z, Lithwick-Yanai G, Amariglio N, Rechavi G, Givol D.** MIR-451 and Imatinib mesylate inhibit tumor growth of glioblastoma stem cells. *Biochem Biophys Res Commun* 376: 86–90, 2008.
15. **Glazov EA, Cottee PA, Barris WC, Moore RJ, Dalrymple BP, Tizard ML.** A microRNA catalog of the developing chicken embryo identified by a deep sequencing approach. *Genome Res* 18: 957–964, 2008.

16. **Griffiths-Jones S, Grocock RJ, van Dongen S, Bateman A, Enright AJ.** miRBase: microRNA sequences, targets and gene nomenclature. *Nucl Acids Res* 34: D140–D144, 2006.
17. **Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ.** miRBase: tools for microRNA genomics. *Nucl Acids Res* 36: D154–D158, 2008.
18. **Gruber AR, Lorenz R, Bernhart SH, Neubock R, Hofacker IL.** The Vienna RNA websuite. *Nucl Acids Res* 36: W70–W74, 2008.
19. **Landgraf P, Rusu M, Sheridan R, Sewer A, Iovino N, Aravin A, Pfeffer S, Rice A, Kamphorst AO, Landthaler M, Lin C, Socci ND, Hermida L, Fulci V, Chiaretti S, Foa R, Schliwka J, Fuchs U, Novosel A, Muller RU, Schermer B, Bissels U, Inman J, Phan Q, Chien M, Weir DB, Choksi R, De Vita G, Frezzetti D, Trompeter HI, Hornung V, Teng G, Hartmann G, Palkovits M, Di Lauro R, Wernet P, Macino G, Rogler CE, Nagle JW, Ju J, Papavasiliou FN, Benzing T, Lichter P, Tam W, Brownstein MJ, Bosio A, Borkhardt A, Russo JJ, Sander C, Zavolan M, Tuschl T.** A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell* 129: 1401–1414, 2007.
20. **Lee DY, Deng Z, Wang CH, Yang BB.** MicroRNA-378 promotes cell survival, tumor growth, and angiogenesis by targeting SuFu and Fus-1 expression. *Proc Natl Acad Sci USA* 104: 20350–20355, 2007.
21. **Lee Y, Kim M, Han J, Yeom KH, Lee S, Baek SH, Kim VN.** MicroRNA genes are transcribed by RNA polymerase II. *EMBO J* 23: 4051–4060, 2004.
22. **Lewis BP, Burge CB, Bartel DP.** Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* 120: 15–20, 2005.
23. **Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB.** Prediction of mammalian microRNA targets. *Cell* 115: 787–798, 2003.
24. **Liang Y, Ridzon D, Wong L, Chen C.** Characterization of microRNA expression profiles in normal human tissues. *BMC Genomics* 8: 166, 2007.
25. **Morin P Jr, Dubuc A, Storey KB.** Differential expression of microRNA species in organs of hibernating ground squirrels: a role in translational suppression during torpor. *Biochim Biophys Acta* 1779: 628–633, 2008.
26. **Morin RD, O'Connor MD, Griffith M, Kuchenbauer F, Delaney A, Prabhu AL, Zhao Y, McDonald H, Zeng T, Hirst M, Eaves CJ, Marra MA.** Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. *Genome Res* 18: 610–621, 2008.
27. **Qian S, Ding JY, Xie R, An JH, Ao XJ, Zhao ZG, Sun JG, Duan YZ, Chen ZT, Zhu B.** MicroRNA expression profile of bronchioalveolar stem cells from mouse lung. *Biochem Biophys Res Commun* 377: 668–673, 2008.
28. **Shao C, Liu Y, Ruan H, Li Y, Wang H, Kohl F, Goropashnaya AV, Fedorov VB, Zeng R, Barnes BM, Yan J.** Shotgun proteomics analysis of hibernating Arctic ground squirrels. *Mol Cell Proteomics* 9: 313–326, 2010.
29. **Smit A, Hubley R, Green P.** *RepeatMasker* Open-3.0. *Institute for Systems Biology*, http://www.repeatmasker.org 1996–2004.
30. **Wang H, Ach RA, Curry B.** Direct and sensitive miRNA profiling from low-input total RNA. *RNA* 13: 151–159, 2007.
31. **Williams DR, Epperson LE, Li W, Hughes MA, Taylor R, Rogers J, Martin SL, Cossins AR, Gracey AY.** Seasonally hibernating phenotype assessed through transcript screening. *Physiol Genomics* 24: 13–22, 2005.
32. **Yan J, Barnes BM, Kohl F, Marr TG.** Modulation of gene expression in hibernating arctic ground squirrels. *Physiol Genomics* 32: 170–181, 2008.
33. **Yan J, Burman A, Nichols C, Alila L, Showe LC, Showe MK, Boyer BB, Barnes BM, Marr TG.** Detection of differential gene expression in brown adipose tissue of hibernating arctic ground squirrels with mouse microarrays. *Physiol Genomics* 25: 346–353, 2006.
34. **Zeeberg BR, Feng W, Wang G, Wang MD, Fojo AT, Sunshine M, Narasimhan S, Kane DW, Reinhold WC, Lababidi S, Bussey KJ, Riss J, Barrett JC, Weinstein JN.** GoMiner: a resource for biological interpretation of genomic and proteomic data. *Genome Biol* 4: R28, 2003.
35. **Zhang R, Peng Y, Wang W, Su B.** Rapid evolution of an X-linked microRNA cluster in primates. *Genome Res* 17: 612–617, 2007.