

RESEARCH ARTICLE

Open Access

# MicroRNA profiling of rhesus macaque embryonic stem cells

Zhenghua Sun<sup>1,3,4†</sup>, Qiang Wei<sup>2,3,4†</sup>, Yanfeng Zhang<sup>1,3,4</sup>, Xiechao He<sup>2,3</sup>, Weizhi Ji<sup>2,3\*</sup> and Bing Su<sup>1,3\*</sup>

## Abstract

**Background:** MicroRNAs (miRNAs) play important roles in embryonic stem cell (ESC) self-renewal and pluripotency. Numerous studies have revealed human and mouse ESC miRNA profiles. As a model for human-related study, the rhesus macaque is ideal for delineating the regulatory mechanisms of miRNAs in ESCs. However, studies on rhesus macaque (r)ESCs are lacking due to limited rESC availability and a need for systematic analyses of fundamental rESC characteristics.

**Results:** We established three rESC lines and profiled microRNA using Solexa sequencing resulting in 304 known and 66 novel miRNAs. MiRNA profiles were highly conserved between rESC lines and predicted target genes were significantly enriched in differentiation pathways. Further analysis of the miRNA-target network indicated that gene expression regulated by miRNAs was negatively correlated to their evolutionary rate in rESCs. Moreover, a cross-species comparison revealed an overall conservation of miRNA expression patterns between human, mouse and rhesus macaque ESCs. However, we identified three miRNA clusters (miR-467, the miRNA cluster in the imprinted Dlk1-Dio3 region and C19MC) that showed clear interspecies differences.

**Conclusions:** rESCs share a unique miRNA set that may play critical roles in self-renewal and pluripotency. MiRNA expression patterns are generally conserved between species. However, species and/or lineage specific miRNA regulation changed during evolution.

## Background

ESCs are derived from the inner cell mass (ICM) of blastocyst-stage embryos [1,2]. Self-renewal and pluripotency enable ESCs to be a renewable and versatile model to study developmental biology. Moreover, ESCs have potential applications in regenerative medicine. Rhesus macaques (*Macaca mulatta*) are a well-studied primates, and with genetic and physiological similarities to humans, rhesus macaques have become an ideal model for ESC-based therapies [3]. However, the study and application of rESCs are lacking compared with those of mouse and human ESCs due to limited rESC line availability and a need for systematic analyses of fundamental rESC characteristics.

MiRNAs are small endogenous non-coding transcripts (~19-25 nt) with diverse roles in development, differentiation and oncogenesis. MiRNAs bind to complementary sites within cognate mRNA 3' UTRs, resulting in degradation, deadenylation or translational repression, which provide a crucial level of post-transcriptional regulation [4]. Moreover, tissue- and cell type-specific miRNA expression patterns have been described [5-9], which elucidate various miRNA functions in specific conditions. MiRNAs also play important roles in ESCs as demonstrated by deletion of Dicer or DGCR8 in mouse ESCs resulting in proliferation and differentiation defects [10-12]. Previous studies of miRNA expression patterns in mouse and human ESCs have revealed a unique miRNA set that is distinct from other cell types and tissues [6,13,14]. Several miRNAs preferentially expressed in human and mouse ESCs, and down-regulated in differentiated cells are key regulators of 'stemness' [15-19]. However, the miRNA expression profile of rESCs is unknown.

\* Correspondence: wji@mail.kiz.ac.cn; sub@mail.kiz.ac.cn

† Contributed equally

<sup>1</sup>State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, China

<sup>2</sup>Department of Reproduction and Development, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, China

Full list of author information is available at the end of the article

ESC lines derived from the same species may contain distinct miRNA profiles and share only a small number of miRNAs [20]. This observation is likely caused by various ESC culture conditions rather than inherent genetic variation within embryos used for ESC derivation [21,22]. Another factor is the use of numerous analyses for detecting miRNA patterns due to the limited resolution of techniques such as microarray analysis [23]. However, recent advancements in next-generation sequencing technology provide an ideal tool for analyzing the miRNA transcriptome with high resolution to identify novel miRNAs [24].

In this study, we isolated and characterized three rESC lines, and performed miRNA profiling using Solexa sequencing. Our miRNA study of rESCs and cross-species comparison may assist future studies for understanding and modulating ESC regulatory networks.

## Results

### Isolation and characterization of rESC lines

Thirteen expanded rhesus macaque blastocysts with a prominent ICM were selected by immunosurgery and 11 ICMs were isolated and plated onto feeder cells. ICMs attached to feeder cells within 48 h and three ESC-like ICM outgrowths appeared after 7-8 days. ICM outgrowths were manually dissociated into 4-6 smaller clumps using a microscalpel, excised from feeder cells and replated onto fresh mouse embryonic fibroblasts (mEFs). Clones with distinct boundaries and high nuclear to cytoplasm ratios were selected for further propagation. Three rESC lines were established and designated as IVF1.2, IVF3.2 and IVF3.3. IVF3.2 and IVF3.3 were derived using the same sperm and oocyte donors. rESCs shared common morphologies with other primate ESCs such as being flat with a distinct boundary against feeder cells. Cells showed high nuclear to cytoplasm ratios and prominent nucleoli (Figure 1A). IVF1.2 and IVF3.3 were cultured for >60 passages and IVF3.2 for >80. Pluripotency markers were highly expressed in all rESC lines including Oct-4, Nanog, SSEA-4, TRA-1-60 and TRA-1-81 (Figure 1A). Other crucial transcription factors, such as Sox-2 and Rex-1, were also detected by RT-PCR (data not shown).

All rESC lines showed strong alkaline phosphatase activity (Figure 1A) and could spontaneously differentiate into cell lineages from the three embryonic germ layers. After 2 days in suspension culture, rESC clones formed embryoid bodies (EBs). After 5 days, cavities were observed in EBs, which are termed cystic EBs. EBs attached to gelatin-coated dishes and were cultured for 5-7 days until various differentiated cells appeared including neuron-like cells, contractive cardiomyocytes and endoderm-like cells (Figure 1B). Teratomas were formed 6-7 weeks after rESC injection into the hind leg

muscles of SCID mice. Histological characterization revealed that ectoderm (neural tube and squamous epithelium), mesoderm (cartilage, muscle and adipose) and endoderm (intestinal epithelia and glands) like structures were present in teratomas (Figure 1B). These results demonstrated that rESC lines could differentiate into all three embryonic germ layers *in vitro* and *in vivo*.

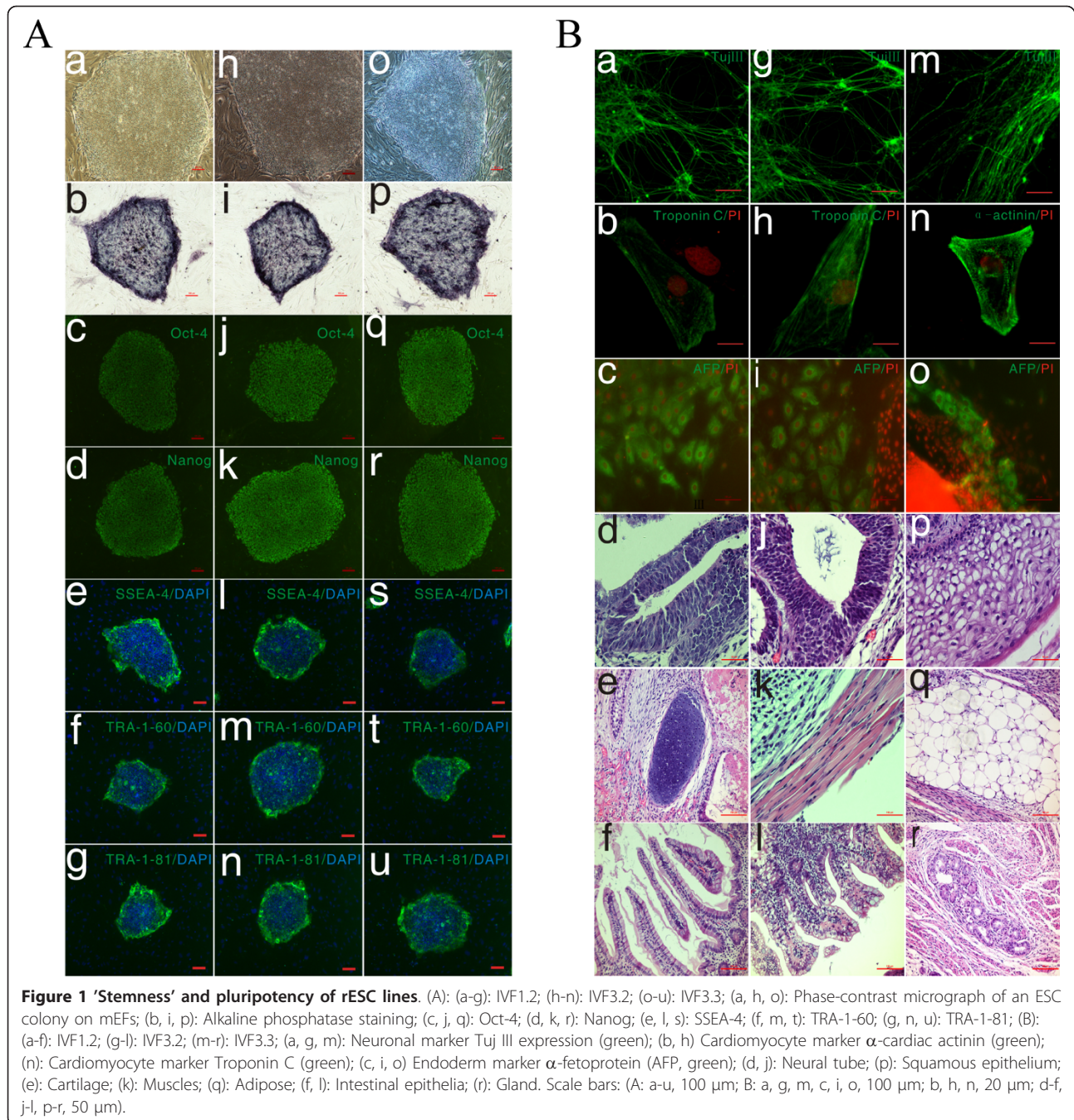
Detailed G-banding analysis revealed that all rESC lines were karyotypically normal with a diploid set of 42 chromosomes (Fig. S1 in Additional file 1), even after long-term culture *in vitro* and repeated freeze-thaw procedures. IVF1.2 and IVF3.3 were derived from male blastocysts (40+XY) and IVF3.2 was derived from a female blastocyst (40+XX).

### Small RNA Sequencing

Newly established rESCs were tested for purity using Oct-4 staining, which resulted in >95% Oct-4-positive cells (Fig. S2 in Additional file 1). A small RNA library from each cell line was then constructed for sequencing. Small RNA library sequencing yielded  $12.66 \times 10^6$ ,  $13.12 \times 10^6$  and  $11.57 \times 10^6$  raw reads from IVF1.2, IVF3.2 and IVF3.3, respectively. After filtering,  $10.89 \times 10^6$  (IVF1.2),  $10.60 \times 10^6$  (IVF3.2) and  $9.26 \times 10^6$  (IVF3.3) clean reads (18-30 nt) were obtained. The length distribution of clean reads centered at 22-23 nt, coinciding with the miRNA length range. Sequence alignment with the rhesus macaque genome (rheMac2) using the Short oligonucleotide alignment program (SOAP) [25] demonstrated that  $8.36 \times 10^6$  (IVF1.2),  $7.68 \times 10^6$  (IVF3.2) and  $7.15 \times 10^6$  (IVF3.3) sequences were mapped to the genome (Figure 2A). MiRNAs were the major component of small RNA libraries from rESCs (Figure 2B) being 60% of sequences, which were matched to reference miRNAs in each sample. The proportions of other annotated RNAs was <15%. Notably, ~25% of sequences could not be annotated because of low coverage ( $6 \times$ ) and relatively poor annotations of the available rhesus macaque genome. Unknown sequences were further analyzed to identify novel miRNAs.

### MiRNA expression profiles of rESC lines

After matching to 466 reference rhesus macaque miRNAs, 326, 329 and 326 known miRNAs were detected in IVF1.2, IVF3.2 and IVF3.3, respectively, with absolute counts ranging from 1-1,702,656 (Table S1 in Additional file 1). Among them, 304 miRNAs were shared by the three pools and remaining miRNAs were expressed at low levels with <10 counts. From a genome-wide perspective, very low miRNAs were likely sequencing artifacts or miRNA machinery byproducts [26]. Therefore, an absolute count of 30 was set for each pool as the minimum threshold for further analyses.

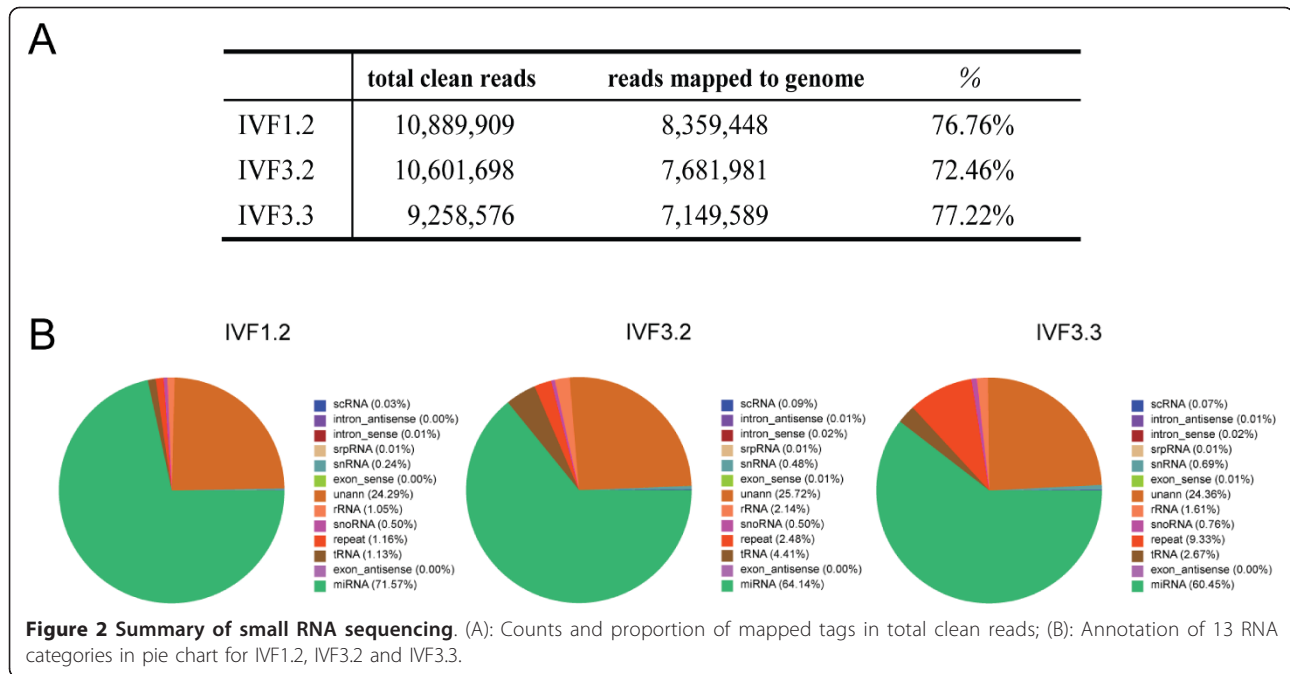


To compare miRNA expression patterns between rESC lines, we evaluated the expression variability (Var) and coefficient of variation (C.V.) of miRNA clusters (Figure 3A, 3B). The results revealed similar expression patterns of miRNA clusters in rESC lines. We also compared the expression of total miRNAs using a differential index (D.I.) and Kappa Statistical analyses. For the D.I., most miRNAs were consistently expressed (Figure 3C), whereas a slight deviation existed in IVF1.2 due to low-level expressed miRNAs. For Kappa Statistical

analysis, there was high repeatability between any two cell lines (Figure 3D). Taken together, miRNA expression patterns in rESC lines were highly consistent.

#### Identification of novel miRNAs

We predicted novel miRNA genes in unannotated sequences using Mireap (<http://sourceforge.net/projects/mireap/>). Briefly, according to the characteristic hairpin structure of miRNA precursors, unannotated small RNA sequences mapped to genome were downloaded with



flanking sequences and analyzed for secondary structure, Dicer cleavage sites and minimum free energy (mfe). After analysis, 141,635 reads that contained 135 novel miRNA precursors with >30 reads were identified. There were 62 novel miRNA precursors (66 mature miRNAs) shared between the three rESC lines (Table S2 in Additional file 1) and 19 (23 mature miRNAs) were homologous to reference miRNAs (miRBase V16.0) in other species (Table 1). Of the 19 miRNA precursors with homologs, 16 precursors were only conserved in humans and other primates, but not rodents.

#### MiRNA target gene prediction and functional classification

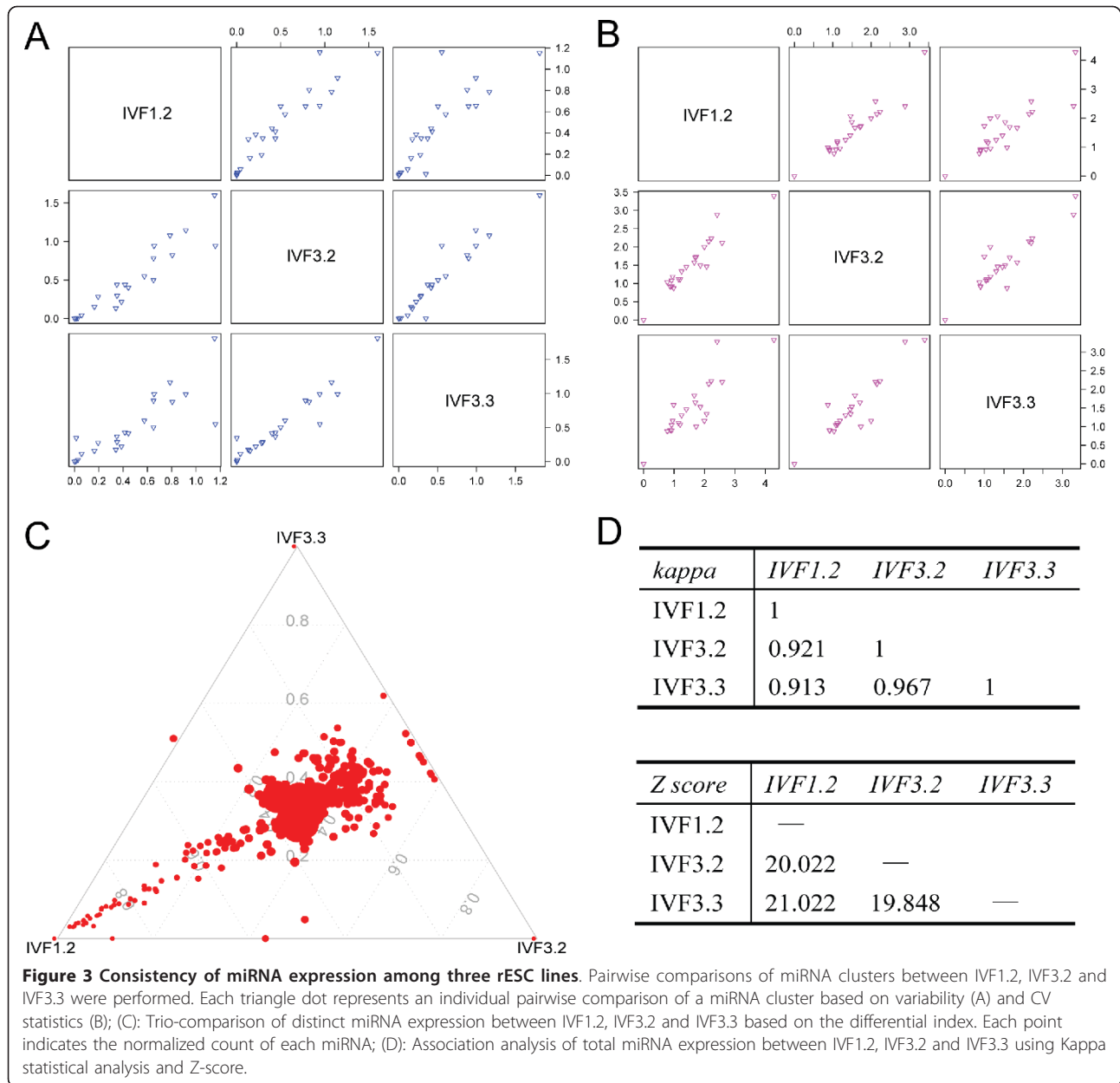
Three programs were used to predict miRNA target genes. Using the TargetScan program, a custom set of Perl codes were downloaded from the TargetScan database (Release 5.1, <http://www.targetscan.org/>). Conserved and non-conserved predicted patterns of miRNA targets were considered separately. For PITA and miRanda, we used 238 expressed miRNAs (number of reads  $\geq 30$ ) and rhesus macaque 3'UTRs orthologous to annotated human 3'UTRs as inputs to predict miRNA targets. Over-representation of predicted miRNA targets for both conserved and non-conserved patterns based on the TargetScan program in the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway is shown in Table S3 in Additional file 1 ( $P \leq 0.05$ , after Benjamini Hochberg correction). Gene enrichment patterns obtained by various methods were similar to each other, indicating that miRNAs correlated to functionally

critical roles in rESCs. For example, predicted targets involved in the MAPK (hsa04010), Wnt (hsa04310) and TGF-beta signaling pathways (hsa04350), and pathways involved in cancer (hsa05200) were significantly enriched (Table S3 in Additional file 1).

#### MiRNA-target interaction network analysis

A systematic survey of miRNAs that interacted with targets may provide a better understanding of miRNA regulation in stem cells. Thus, we analyzed potential interactions between miRNAs and their targets, which is termed as the miRNA-target network (MT network). All miRNAs and their targets were used to generate a bipartite graph of miRNA-target interactions. Nodes represent miRNAs or target genes and edges correspond to interactions between miRNAs with target genes. Node degree in the MT network is the number of connections or edges with other nodes. For each target gene node, degrees represent the number of miRNAs targeting the gene. Globally, the MT network was comprised of 8,934 nodes with 48,546 edges. Using MT network degree analysis (see methods), the degree of genes regulated by miRNAs was negatively correlated ( $r = -0.564$  with  $p = 0.018$  based on Pearson's correlation) with their evolutionary rate (dN/dS) (Figure 4A). Similarly, a previous study reported a negative correlation between the degree and evolutionary rate of proteins in a protein-protein interaction network [27].

To exclude factors with potential effects on the observed negative correlation, we tested the effect of 3'UTR length. After length normalization, the negative



correlation remained (Figure 4B,4C). Another factor was a potential ascertainment bias caused by miRNA target prediction and identification of orthologous macaque 3'UTRs. To evaluate this possibly, we randomized the scale-free MT network based on Barabasi-Albert model. The degree of miRNA and target genes in the random network was significantly lower ( $p < 2.2e^{-16}$  with a Wilcoxon rank sum test) compared with that observed in the MT network (Figure 4D). A significant correlation was not observed in the random network, which further supported the negative correlation in the MT network.

#### Comparison of miRNA profiles from mouse, human and rhesus macaque

To compare ESC miRNA profiles between species, we retrieved previously reported mouse and human ESC miRNA profiles. To avoid a cross-talking bias due to various sequencing techniques and data sizes, we selected Solexa sequencing datasets from mouse and human ESCs [14,28]. MiRNA expression frequencies were normalized to each other by determining the expected frequency in mapped reads per million. MiRNA expression levels were calculated as the mean of three samples. The three species expressed 138 common miRNAs. Clustering analysis

**Table 1 Novel miRNAs identified in rESCs**

Mature miRNA sequence	MiRNA precursor location	Name	Reads		
			IVF1.2	IVF3.2	IVF3.3
TATGGAGGTCTCTGTCTGGCT	chr1:205,580,537-205,580,616:-	miR-1843-5p	124	159	152
TCTGATCGTCCCTCCATACA	chr1:205,580,537-205,580,616:-	miR-1843-3p	127	174	193
TGATGGGTGAATTTGTAGAAGG	chr1:70,976,129-70,976,208:-	miR-1262-3p	2,557	2,267	2,150
GTTGGGACAAGAGAACGGTCTT	chr1:158,332,926-158,333,000:-	miR-3122-5p	252	371	214
AATCCCTTATGGATAATCTGG	chr2:80519804-80519882:+	miR-3938-3p	138	227	104
CATGCTAGAACAGAAAGAATGGG	chr3:106,443,245-106,443,327:+	miR-3146-3p	37	120	61
TATTTTGTAGTGTGGGAATTGA	chr4:125,407,861-125,407,939:+	miR-3145-3p	35	36	31
GAAAATGATGAGTAGTGACTGATG	chr4:128869861-128869951:+	miR-3622-3p	72	243	73
AAGAGCTTTTGGGAATTCAGGTAG	chr5:144,708,317-144,708,405:-	miR-3140-3p	151	358	200
ATATACAGGGGGAGACTCTTAT	chr7:164,327,476-164,327,555:+ chr7:164,328,696-164,328,775:+	miR-1185-3p	52	176	105
CGGGAACGTCGAGACTGGAGC	chr7:164,844,819-164,844,896:-	miR-1247-3p	127	131	32
TTAGGGCCCTGGCTCCATCTCC	chr9:73,887,081-73,887,164:+	miR-1296-5p	33	123	53
TCGACCGGACTCGACCGGCTCG	chr9:103,084,702-103,084,783:-	miR-1307-5p	291	872	944
ACTCGGCGTGGCGTCGGTCGTGG	chr9:103,084,702-103,084,783:-	miR-1307-3p	1,936	3,975	3,030
GACTCTAGCTGCCAAAGGCGCT	chr11:98,713,141-98,713,223:+	miR-1251-5p	40	97	59
TGTGGACCTCTGGCCTTGGC	chr11:105706113-105706195:+	miR-3922-3p	192	250	248
TGCGGGCTAGGGCTAACAGCA	chr16:11,829,804-11,829,901:+	miR-744-5p	11,447	24,797	14,171
CTGTTGCCACTAACCTCAACC	chr16:11,829,804-11,829,901:+	miR-744-3p	75	81	104
TTCCGGCTCGCGTGGGTGTGT	chr16:18,891,345-18,891,423:-	miR-1180-3p	30	133	77
CCGTCCTAAGTTGTTGAGTT	chrX:68,990,098-68,990,174:+	miR-676-3p	46	206	206
TTCATTGCGCTGTCCAGATGTA	chrX:113,236,410-113,236,505:+	miR-1298-5p	246	893	852
CATCTGGCAACTGACTGAACT	chrX:113,236,410-113,236,505:+	miR-1298-3p	30	98	58
TGAGTACCGCCATGTCTGTTGGG	chrX:113,271,154-113,271,232:+	miR-1911-5p	88	843	473

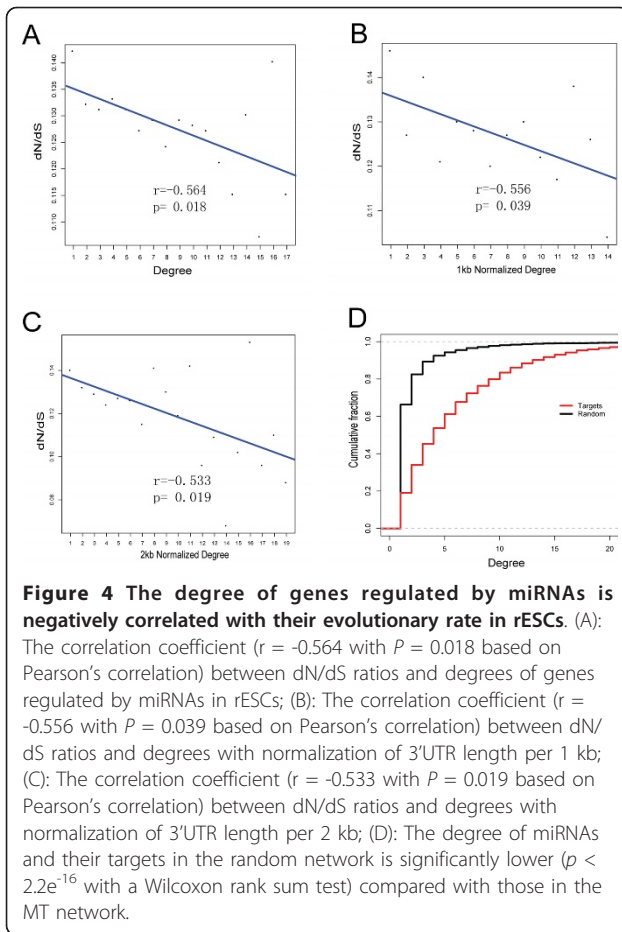
based on the 138 shared miRNAs revealed that miRNA expression profiles were more similar between human and rhesus macaque (Fig. S3 in Additional file 1), which was consistent with their evolutionary relationship. A number of miRNAs were stably and highly expressed between species, such as the miR-17 family, miR-21 and miR-103, which suggested conserved functions in ESCs during evolution.

Of the 138 common miRNAs the 'stemness' miR-290-295 cluster [13,29] was the highest expressed in mouse ESCs. However, primate homologs the miR-290-295 and miR-371-373 clusters were not highly expressed in human and rESCs. Instead, the miR-302 cluster was the highest expressed in human and rhesus macaque [14,30,31], indicating the functional divergence of stemness miRNA clusters in primate lineages. The miR-290-295 cluster contains multiple mature miRNAs with seed sequences similar or identical to those in the miR-302 cluster [6]. MiR-290-295 and miR-302 clusters are transcriptionally regulated by Oct-4/Sox2 [29]. Taken together, these observations suggest that these two clusters are similarly expressed, regulate common target genes and are functionally conserved in ESCs derived from various species.

There are three miRNA clusters differentially expressed in mouse, human and rESCs. The chromosome 19 miRNA cluster (C19MC), a primate-specific miRNA cluster [32], was enriched in human ESCs, but almost absent in rESCs (Figure 5C). A conserved miRNA cluster within the imprinted *Dlk1-Dio3* region was highly expressed in mouse and rhesus macaque ESCs, but rarely expressed in human ESCs (Figure 5B), which was consistent with results from four other human ESC lines [31]. The miR-467 cluster in the *sfmbt2* gene intron region was only expressed in mouse ESCs (Figure 5A), embryos and newborn ovaries [33,34], but not in human and rhesus macaque ESCs, consistent with previous reports [29].

## Discussion

We established three rESC lines (IVF1.2, IVF3.2 and IVF3.3) and profiled miRNA expression using Solexa sequencing. MiRNA expression profiles were generally conserved between the three rESC lines, sharing 92.4% of expressed miRNAs. IVF3.2 and IVF3.3 were more similar because of derivation from identical sperm and ovum donors. Similarly, Navara *et al.* [35] reported that pedigreed rESCs express homogeneous gene profiles.



Compared with known miRNAs in our data set, most novel miRNAs were moderately expressed. Novel miRNA biological functions and association with stemness are yet to be elucidated.

We predicted miRNA targets in rESCs using three independent algorithms. Targets involved in TGF- $\beta$ , MAPK and Wnt signaling pathways were significantly enriched. These pathways play important roles in mouse and human ESC differentiation [36-38]. This observation suggests that in addition to transcriptional level regulation by stemness-associated factors such as Oct4 and Nanog, miRNAs may also play an important role in maintaining ESC self-renewal and pluripotency at the post-transcriptional level. Moreover, over-represented miRNA targets involved in cancer pathways support the observation that cancer and ESCs share pathways for self-renewal and proliferation [38].

Our systematic investigation of the MT network suggested that the degree of genes regulated by miRNAs was negatively correlated with their evolutionary rate in rESCs. Genes with more miRNA regulators evolve more slowly. This was anticipated in accordance to a rule of "the more conserved, the deeper regulated" has been

proposed in protein-protein interaction networks, in which hubs or central proteins are functionally essential and usually lethal after gene knockout [27]. Therefore in the MT network, hub targets should be more regulated due to selective pressure, which indicates the fine regulation of miRNAs in ESCs.

The cross-species comparison indicated that the majority of shared miRNAs were stably expressed between species and consistent with miRNA sequence-level conservation. This observation suggested that miRNA expression is under stabilizing selection due to functional constraints on miRNAs in ESCs. Despite this close relationship, miRNA expression differences existed between human and rhesus macaque ESCs. The chromosome 19 miRNA cluster (C19MC) is a primate-specific miRNA cluster and contains >30 mature miRNAs. These clustered miRNAs are expressed in human ESCs, placenta and fetal brain [14,31,32,39], suggesting that they are involved in embryogenesis. However, the C19MC cluster is almost absent in rESCs. This expression difference suggests that even between closely related species, there may be a significant miRNA expression divergence and regulatory differences during embryogenesis between rhesus macaques and humans. In addition, the miRNA cluster in the imprinted Dlk1-Dio3 region was conserved at the sequence level between mice, rhesus macaques and humans. Liu *et al* reported that the expression of this cluster is positively correlated to mouse stem cell pluripotencies (ESCs vs iPS cells) [40]. We found that this cluster was enriched in mouse and rhesus macaque ESCs, but rare in human ESCs. Recently, it was reported that diverse temporal origins may be responsible for observed differences between mouse and human ESCs [41]. Our data suggests that miRNA expression patterns may provide a new strategy for investigating interspecies differences in ESC pluripotency.

## Conclusions

Our results indicate that miRNA profiles are highly conserved and consistent between the three rESC lines, and these miRNAs may play critical roles in the maintenance of self-renewal and pluripotency in rESCs, particularly in differentiation pathways. Further analysis of the miRNA-target network indicates that the degree of genes regulated by miRNAs is negatively correlated with their evolutionary rate in rESCs. This observation suggests important roles of miRNAs in post-transcriptional regulation. Moreover, cross-species comparison of ESCs revealed an overall conservation of miRNA expression patterns, indicating stabilizing selection on miRNA expression in ESCs during evolution. However, we identified three miRNA clusters (miR-467, the miRNA cluster in the imprinted Dlk1-Dio3 region and C19MC) that show distinct differences between mouse, human and

A	miRNA	reads(mmu)
	miR-297a	317.35
	miR-297b	163.33
	miR-466	642.55
	miR-467a	758.47
	miR-467b	639.36
	miR-669a	2,074.47
	miR-669b	552.19
	miR-669c	3,238.71

B	miRNA	reads(mmu)	reads(mml)	reads(hsa)
	miR-127	5,228.81	3,531.18	195.13
	miR-134	712.32	14,791.50	24.09
	miR-136	6,193.40	59.33	0.00
	miR-154	3,244.52	5.30	0.00
	miR-299	833.75	344.08	0.00
	miR-300	1,578.57	NA	0.00
	miR-323	956.06	142.73	0.00
	miR-329	442.27	3.37	0.00
	miR-337	1,366.36	27.53	0.00
	miR-341	129.50	NA	0.00
	miR-369	2,025.21	566.13	0.00
	miR-370	391.09	3,670.14	0.00
	miR-376a	2,174.19	36.74	0.00
	miR-376b	2,627.53	5.58	0.00
	miR-376c	935.23	23.02	0.00
	miR-377	869.67	14.81	0.00
	miR-379	1,128.50	25,147.95	16.86
	miR-380	813.86	294.81	0.00
	miR-381	806.35	2,501.28	0.00
	miR-382	1,515.73	1,825.11	0.00
	miR-409	1,801.92	1,247.36	0.00
	miR-410	886.07	154.22	0.00
	miR-411	2,668.86	1,036.93	0.00
	miR-412	40.31	3.96	0.00
	miR-431	807.18	21.39	0.00
	miR-432	0.00	1,421.26	0.00
	miR-433	778.32	54.75	26.50
	miR-434	2,390.01	NA	0.00
	miR-485	739.39	165.13	0.00
	miR-487b	517.05	27.84	0.00
	miR-493	0.00	1,053.07	0.00
	miR-494	1,359.55	212.05	0.00
	miR-495	1,637.62	7.71	0.00
	miR-496	232.13	2.69	0.00
	miR-539	233.50	21.04	0.00
	miR-541	5,738.05	0.00	0.00
	miR-543	2,240.25	0.00	0.00
	miR-654-3p	0.00	302.87	0.00
	miR-654-5p	0.00	205.55	0.00
	miR-665	1,200.14	0.00	0.00
	miR-758	300.19	33.49	0.00
	miR-889	0.00	576.35	0.00

C	miRNA	reads(mml)	reads(hsa)
	miR-498	0.00	156.59
	miR-512	3.20	2,204.24
	miR-516	0.37	626.34
	miR-517a	0.00	785.34
	miR-517b	0.00	19.27
	miR-517c	NA	9.64
	miR-518a	0.16	180.68
	miR-518b	0.14	517.94
	miR-518c	0.00	103.59
	miR-518d	0.00	79.50
	miR-518e	0.00	103.59
	miR-518f	0.21	202.36
	miR-519a	0.00	156.59
	miR-519b	0.00	93.95
	miR-519c	0.14	106.00
	miR-519d	0.00	122.86
	miR-519e	NA	12.05
	miR-520a	0.00	874.47
	miR-520b	0.00	79.50
	miR-520c	0.00	9.64
	miR-520d	0.22	74.68
	miR-520e	0.00	16.86
	miR-520f	0.00	81.91
	miR-520g	0.00	178.27
	miR-520h	0.00	50.59
	miR-521	0.00	14.45
	miR-522	0.00	33.73
	miR-523	0.00	84.32
	miR-524	0.00	93.95
	miR-525	0.00	57.82
	miR-526a	0.00	48.18
	mir-526b	0.00	91.54
	mir-527	0.00	36.14

**Figure 5 Expression differences of three miRNA clusters among human, rhesus macaque and mouse ESCs.** (A): MiR-467cluster expression in mouse (mmu) ESCs; (B): Expression of the miRNA cluster in the imprinted Dlk1-Dio3 region in mouse, rhesus macaque (mml) and human (hsa) ESCs; (C): C19MC expression in rhesus macaque (mml) and human (hsa) ESCs. Reads were normalized to the mapped reads per million.



rhesus macaque ESCs, which suggests species and/or lineage specific miRNA regulatory changes during evolution.

## Methods

### rESC isolation and characterization

rESCs were derived from the ICM of rhesus macaque blastocyst-stage embryos and characterized using standard protocols. rESC isolation and characterization details are provided in the supplemental Methods in Additional file 1.

### RNA extraction and small RNA sequencing

Total RNA was extracted using a mirVana miRNA isolation kit (Ambion). According to the manufacturer's instructions (Illumina), the small RNA fraction between 18-30 nt was isolated from total RNA by PAGE (polyacrylamide gel electrophoresis) purification and ligated to a pair of adaptors at the 5' and 3' ends. Small RNA molecules were converted to cDNA and amplified by RT-PCR using adaptor primers. Unique 6 bp indexed sequences were added to 5' adaptors for multiplex sequencing. Purified DNA was directly used for cluster generation and sequencing analysis using an Illumina Genome Analyzer. These data have been deposited in the GEO database (GSE27886).

### Small RNA annotation and novel miRNA prediction

For clean reads, the original 35 nt reads from sequencing were filtered by trimming the 5' and 3' adaptors, eliminating contaminants, and inadequate (<18 nt) and low quality reads. For detecting known miRNAs, clean reads were matched to reference miRNA precursor sequences in the rhesus macaque (miRBase version 16.0) using Blastall (-p blastn -F F -e 0.01). MiRNA absolute expression was the sum of clean reads mapped to the precursor. For annotating degradation tags of rRNA, tRNA, small cytoplasmic RNA (scRNA), small nuclear RNA (snRNA) and small nucleolar RNA (snoRNA), clean reads were aligned with Genbank and Rfam 9.1 databases using Blastall. For identifying repeat associated RNA and degradation tags of mRNA, clean reads were overlapped with repeat sequences, exons and introns of mRNAs. In the initial alignment and annotation, some small RNA tags were mapped to more than one category. For uniquely mapped reads, the following priority rules were applied: rRNA/tRNA/scRNA/snRNA/snoRNA (Genbank > Rfam) > known miRNA > repeat > exon > intron. Unannotated small RNA tags were termed as 'unann'. Mireap (<http://sourceforge.net/projects/mireap/>) was used to predict novel miRNAs.

### MiRNA expression profiling analysis

Based on the randomly selected 10,000 bins for each chromosome with each bin spanning a 50 bp window (210,000 bins in 21 chromosomes), miRNAs with sporadic reads were filtered out and miRNAs with >30 reads ( $p < 0.0001$  based on Poisson distribution simulation) were treated as true reads for further analyses. miRNA expression profiles were compared and evaluated in rESCs at cluster and individual miRNA levels. A miRNA cluster was defined as consecutive miRNAs with a  $\leq 10$  kb inter-miRNA distance (MID). For each level, two independent methods were used for confirmation. Var and C.V. were used to survey miRNA member divergences in a cluster. Var was used to represent a divergence measure based on Shannon entropy [42]. D.I. and Kappa Statistical analysis ( $\kappa$ ) were used to compare total expressed miRNAs in rESCs. D.I. was used to evaluate relative deviation of individual miRNA expression levels in rESCs. Kappa statistical analysis was used to measure the repeatability of total miRNAs expression between any two samples (supplemental methods in Additional file 1).

### MiRNA target prediction

To predict miRNA target genes, TargetScanS [43], miRanda (<http://www.microrna.org/microrna/home.do>) and PITA ([http://genie.weizmann.ac.il/pubs/mir07/mir07\\_prediction.html](http://genie.weizmann.ac.il/pubs/mir07/mir07_prediction.html)) were used. Common miRNAs ( $\geq 30$  reads) were used to predict target genes. For each miRNA, the reference mature miRNA sequence in miRBase was used. For 3'UTR identification in the rhesus macaque, human annotated 3'UTRs from Refseq and orthologous 3'UTRs from six species (human, chimpanzee, macaque, mouse, rat and dog) were extracted from multi17way aligned files downloaded from the UCSC database [44]. Orthologous 3'UTRs were aligned using the MUSCLE [45] program to obtain multi6way 3'UTR files required for analysis.

Gene ontology (GO) analyses of predicted miRNA targets were performed using t DAVID Bioinformatics Resources [46] (<http://david.abcc.ncifcrf.gov/home.jsp>). Ascertainment biases such as a relatively low number of annotated miRNAs, 3'UTRs in the rhesus macaque,  $\geq 30$  cutoffs for miRNA reads and false predicted miRNA targets were assessed by simulating the bias with 240 randomly synthesized *in silico* 22 nt miRNAs (equal number with true expression).

Potential correlations between miRNA (or target) degrees connected in the MT network, miRNA expression and target evolution were evaluated (supplemental methods in Additional file 1).

## Additional material

**Additional data file 1:** is a DOC file with Supplemental Methods, Figures (S1, S2, S3, S4) and Tables (S1, S2, S3, S4).

### Abbreviations

ESCs: embryonic stem cells; ICM: inner cell mass; EB: embryoid body; miRNA: microRNA; nt: nucleotide; scRNA: small cytoplasmic RNA; snRNA: small nuclear RNA; snoRNA: small nucleolar RNA; Var: variability; C.V.: coefficient of variation; D.I.: differential index;  $\kappa$ : Kappa Statistic; MT network: miRNA-target regulatory network; iPS cell: induced pluripotent stem cell; C19MC: chromosome 19 miRNA cluster; mfe: minimum free energy; SOAP: Short oligonucleotide alignment program; SCID: severe combined immune deficiency; mEFs: mouse embryonic fibroblasts; dN/dS: nonsynonymous vs. synonymous substitution rate.

### Acknowledgements

We thank Hui Zhang, Yuyu Niu, Bin Lu, Tao Tan and Xiangyu Guo for technical assistance. This study was supported by grants from the 973 program of China (2011CBA01101), the National 863 project of China (2006AA02A116 and 2006AA02A101), the Major Project of Ministry of Science and Technology (2009ZX09501-028), the Chinese Academy of Sciences (KSCX1-YW-R-34, Westlight Doctoral Program) and the National Natural Science Foundation of China (30871343).

### Author details

<sup>1</sup>State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, China.

<sup>2</sup>Department of Reproduction and Development, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, China. <sup>3</sup>Kunming Biomed International and National Engineering Research Center of Biomedicine and Animal Science, Kunming 650051, China. <sup>4</sup>Graduate School of Chinese Academy of Sciences, Beijing 100049, China.

### Authors' contributions

ZS and QW: data collection, assembly, analysis and interpretation, and manuscript writing; YZ: data analysis and interpretation, and manuscript writing; XH: provision of study material; WJ and BS: conception and design, data interpretation, financial support, manuscript writing and final approval of manuscript. All authors read and approved the final manuscript.

Received: 30 December 2010 Accepted: 31 May 2011

Published: 31 May 2011

### References

1. Thomson J, Kalishman J, Golos T, Durning M, Harris C, Becker R, Hearn J: **Isolation of a primate embryonic stem cell line.** *Proc Natl Acad Sci USA* 1995, **92**:7844.
2. Thomson J, Itskovitz-Eldor J, Shapiro S, Waknitz M, Swiergiel J, Marshall V, Jones J: **Embryonic stem cell lines derived from human blastocysts.** *Science* 1998, **282**:1145.
3. Wolf D: **Nonhuman primate embryonic stem cells: an underutilized resource.** *Regenerative Medicine* 2008, **3**:129-131.
4. He L, Hannon G: **MicroRNAs: small RNAs with a big role in gene regulation.** *Nature Reviews Genetics* 2004, **5**:522-531.
5. Sempere LF, Freemantle S, Pitha-Rowe I, Moss E, Dmitrovsky E, Ambros V: **Expression profiling of mammalian microRNAs uncovers a subset of brain-expressed microRNAs with possible roles in murine and human neuronal differentiation.** *Genome Biol* 2004, **5**:R13.
6. Suh MR, Lee Y, Kim JY, Kim SK, Moon SH, Lee JY, Cha KY, Chung HM, Yoon HS, Moon SY, et al: **Human embryonic stem cells express a unique set of microRNAs.** *Dev Biol* 2004, **270**:488-498.
7. Lee E, Gusev Y, Jiang J, Nuovo G, Lerner M, Frankel W, Morgan D, Postier R, Brackett D, Schmittgen T: **Expression profiling identifies microRNA signature in pancreatic cancer.** *International Journal of Cancer* 2007, **120**:1046-1054.
8. Volinia S, Calin G, Liu C, Ambs S, Cimmino A, Petrocca F, Visone R, Iorio M, Roldo C, Ferracin M: **A microRNA expression signature of human solid tumors defines cancer gene targets.** *Proc Natl Acad Sci USA* 2006, **103**:2257.
9. Calin G, Croce C: **MicroRNA signatures in human cancers.** *Nature Reviews Cancer* 2006, **6**:857-866.
10. Murchison EP, Partridge JF, Tam OH, Cheloufi S, Hannon GJ: **Characterization of Dicer-deficient murine embryonic stem cells.** *Proc Natl Acad Sci USA* 2005, **102**:12135-12140.
11. Kanellopoulou C, Muljo SA, Kung AL, Ganesan S, Drapkin R, Jenuwein T, Livingston DM, Rajewsky K: **Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing.** *Genes Dev* 2005, **19**:489-501.
12. Wang Y, Medvid R, Melton C, Jaenisch R, Blelloch R: **DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal.** *Nat Genet* 2007, **39**:380-385.
13. Houbaviy HB, Murray MF, Sharp PA: **Embryonic stem cell-specific MicroRNAs.** *Dev Cell* 2003, **5**:351-358.
14. 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.
15. Melton C, Judson RL, Blelloch R: **Opposing microRNA families regulate self-renewal in mouse embryonic stem cells.** *Nature* 2010, **463**:621-626.
16. Xu N, Papagiannakopoulos T, Pan G, Thomson JA, Kosik KS: **MicroRNA-145 regulates OCT4, SOX2, and KLF4 and represses pluripotency in human embryonic stem cells.** *Cell* 2009, **137**:647-658.
17. Sengupta S, Nie J, Wagner RJ, Yang C, Stewart R, Thomson JA: **MicroRNA 92b controls the G1/S checkpoint gene p57 in human embryonic stem cells.** *Stem cells* 2009, **27**:1524-1528.
18. Lee NS, Kim JS, Cho WJ, Lee MR, Steiner R, Gompers A, Ling D, Zhang J, Strom P, Behlke M, et al: **miR-302b maintains "stemness" of human embryonal carcinoma cells by post-transcriptional regulation of Cyclin D2 expression.** *Biochem Biophys Res Commun* 2008, **377**:434-440.
19. Zovoilis A, Smorag L, Pantazi A, Engel W: **Members of the miR-290 cluster modulate in vitro differentiation of mouse embryonic stem cells.** *Differentiation* 2009, **78**:69-78.
20. Wang Y, Keys DN, Au-Young JK, Chen C: **MicroRNAs in embryonic stem cells.** *J Cell Physiol* 2009, **218**:251-255.
21. Allegrucci C, Young L: **Differences between human embryonic stem cell lines.** *Human reproduction update* 2007, **13**:103.
22. Tsai ZY, Singh S, Yu SL, Kao LP, Chen BZ, Ho BC, Yang PC, Li SS: **Identification of microRNAs regulated by activin A in human embryonic stem cells.** *J Cell Biochem* 2010, **109**:93-102.
23. Thomson J, Parker J, Perou C, Hammond S: **A custom microarray platform for analysis of microRNA gene expression.** *Nature Methods* 2004, **1**:47-53.
24. Shendure J, Ji H: **Next-generation DNA sequencing.** *Nature biotechnology* 2008, **26**:1135-1145.
25. Li R, Yu C, Li Y, Lam T, Yiu S, Kristiansen K, Wang J: **SOAP2: an improved ultrafast tool for short read alignment.** *Bioinformatics* 2009, **25**:1966.
26. Winston K, Chen S, Betty T, Qian L, Lim K, Vivek T: **Analysis of deep sequencing microRNA expression profile from human embryonic stem cells derived mesenchymal stem cells reveals possible role of let-7 microRNA family in downstream targeting of Hepatic Nuclear Factor 4 Alpha.** *BMC Genomics* 11.
27. Fraser HB, Hirsh AE, Steinmetz LM, Scharfe C, Feldman MW: **Evolutionary rate in the protein interaction network.** *Science* 2002, **296**:750-752.
28. Babiarz J, Ruby J, Wang Y, Bartel D, Blelloch R: **Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs.** *Genes dev* 2008, **22**:2773.
29. Marson A, Levine S, Cole M, Frampton G, Brambrink T, Johnstone S, Guenther M, Johnston W, Wernig M, Newman J: **Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells.** *Cell* 2008, **134**:521-533.
30. Joglekar MV, Parekh VS, Mehta S, Bhonde RR, Hardikar AA: **MicroRNA profiling of developing and regenerating pancreas reveal post-transcriptional regulation of neurogenin3.** *Dev Biol* 2007, **311**:603-612.
31. Goff LA, Davila J, Swerdel MR, Moore JC, Cohen RI, Wu H, Sun YE, Hart RP: **Ago2 immunoprecipitation identifies predicted microRNAs in human embryonic stem cells and neural precursors.** *PLoS One* 2009, **4**:e7192.
32. Zhang R, Wang Y, Su B: **Molecular evolution of a primate-specific microRNA family.** *Molecular biology and evolution* 2008, **25**:1493.

33. Mineno J, Okamoto S, Ando T, Sato M, Chono H, Izu H, Takayama M, Asada K, Mirochnitchenko O, Inouye M: **The expression profile of microRNAs in mouse embryos.** *Nucleic acids research* 2006, **34**:1765.
34. Ahn H, Morin R, Zhao H, Harris R, Coarfa C, Chen Z, Milosavljevic A, Marra M, Rajkovic A: **MicroRNA transcriptome in the newborn mouse ovaries determined by massive parallel sequencing.** *Molecular Human Reproduction* 2010, **16**:463.
35. Navara C, Mich-Basso J, Redinger C, Ben-Yehudah A, Jacoby E, Kovkarova-Naumovski E, Sukhwani M, Orwig K, Kaminski N, Castro C: **Pedigreed primate embryonic stem cells express homogeneous familial gene profiles.** *Stem cells* 2007, **25**:2695-2704.
36. Burdon T, Tracey C, Chambers I, Nichols J, Smith A: **Suppression of SHP-2 and ERK signalling promotes self-renewal of mouse embryonic stem cells.** *Dev Biol* 1999, **210**:30-43.
37. Ying QL, Nichols J, Chambers I, Smith A: **BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3.** *Cell* 2003, **115**:281-292.
38. Pardal R, Molofsky A, He S, Morrison S: **Stem cell self-renewal and cancer cell proliferation are regulated by common networks that balance the activation of proto-oncogenes and tumor suppressors.** *Cold Spring Harbor Laboratory Press* 2005, 177.
39. Bar M, Wyman SK, Fritz BR, Qi J, Garg KS, Parkin RK, Kroh EM, Bendoraite A, Mitchell PS, Nelson AM, *et al*: **MicroRNA discovery and profiling in human embryonic stem cells by deep sequencing of small RNA libraries.** *Stem cells* 2008, **26**:2496-2505.
40. Liu L, Luo GZ, Yang W, Zhao X, Zheng Q, Lv Z, Li W, Wu HJ, Wang L, Wang XJ, Zhou Q: **Activation of the imprinted Dlk1-Dio3 region correlates with pluripotency levels of mouse stem cells.** *J Biol Chem* 2010, **285**:19483-19490.
41. Hanna J, Cheng A, Saha K, Kim J, Lengner C, Soldner F, Cassady J, Muffat J, Carey B, Jaenisch R: **Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs.** *Proc Natl Acad Sci USA* 2010, **107**:9222.
42. Lin J: **Divergence measures based on the Shannon entropy.** *Information Theory, IEEE Transactions on* 1991, **37**:145-151.
43. Lewis BP, Burge CB, Bartel DP: **Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets.** *Cell* 2005, **120**:15-20.
44. Rhead B, Karolchik D, Kuhn RM, Hinrichs AS, Zweig AS, Fujita PA, Diekhans M, Smith KE, Rosenbloom KR, Raney BJ, *et al*: **The UCSC Genome Browser database: update 2010.** *Nucleic Acids Res* 38:D613-619.
45. Edgar RC: **MUSCLE: a multiple sequence alignment method with reduced time and space complexity.** *BMC Bioinformatics* 2004, **5**:113.
46. Huang da W, Sherman BT, Lempicki RA: **Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources.** *Nat Protoc* 2009, **4**:44-57.

doi:10.1186/1471-2164-12-276

**Cite this article as:** Sun *et al*: MicroRNA profiling of rhesus macaque embryonic stem cells. *BMC Genomics* 2011 **12**:276.

**Submit your next manuscript to BioMed Central  
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

