Comprehensive MicroRNA Profiling Reveals a Unique Human Embryonic Stem Cell Signature Dominated by a Single Seed Sequence

LOUISE C. LAURENT,[a,b] JING CHEN,[c] IGOR ULITSKY,[d] FRANZ-JOSEF MUELLER,[b,e] CHRISTINA LU,[a,b] RON SHAMIR,[d] JIAN-BING FAN,[c] JEANNE F. LORING[b]

[a]Department of Reproductive Medicine, University of California San Diego, San Diego, California, USA; [b]The Scripps Research Institute, La Jolla, California, USA; [c]Illumina, Inc., San Diego, California, USA; [d]School of Computer Science, Tel Aviv University, Tel Aviv, Israel; [e]Zentrum für Integrative Psychiatrie, Universitätsklinikums Schleswig-Holstein, Kiel, Germany

ABSTRACT

Embryonic stem cells are unique among cultured cells in their ability to self-renew and differentiate into a wide diversity of cell types, suggesting that a specific molecular control network underlies these features. Human embryonic stem cells (hESCs) are known to have distinct mRNA expression, global DNA methylation, and chromatin profiles, but the involvement of high-level regulators, such as microRNAs (miRNA), in the hESC-specific molecular network is poorly understood. We report that global miRNA expression profiling of hESCs and a variety of stem cell and differentiated cell types using a novel microarray platform revealed a unique set of miRNAs differentially regulated in hESCs, including numerous miRNAs not previously linked to hESCs. These hESC-associated miRNAs were more likely to be located in large genomic clusters, and less likely to be located in introns of coding genes. hESCs had higher expression of oncogenic miRNAs and lower expression of tumor suppressor miRNAs than the other cell types. Many miRNAs upregulated in hESCs share a common consensus seed sequence, suggesting that there is cooperative regulation of a critical set of target miRNAs. We propose that miRNAs are coordinately controlled in hESCs, and are key regulators of pluripotence and differentiation. STEM CELLS *2008;26:1506–1516*

Disclosure of potential conflicts of interest is found at the end of this article.

INTRODUCTION

Embryonic stem cells (ESCs) possess three features that in combination set them apart from all other cell types: the ability to self-renew indefinitely, the potential to generate every differentiated cell type, and a normal genetic complement. In mice, these properties can be demonstrated by the ability of the cells to develop into whole animals by germline transmission. As a proxy for a germline assay, human embryonic stem cells (hESCs) have been shown to be capable of differentiation into all three germ layers, both in culture by embryoid body formation and in vivo by teratoma formation. It is our goal to characterize the regulatory processes underlying these properties of hESCs on a molecular level.

MicroRNAs (miRNAs) are small (19–25 nucleotides) endogenous noncoding RNAs that have been shown to influence the abundance and translational efficiency of cognate mRNAs. Discovered in *Caenorhabditis elegans*, miRNAs are known to control critical time points in development of plants and lower animals. However, the roles of miRNAs in the development of higher animals are less well understood. Details of the biogenesis and mechanisms of action of miRNAs continue to be the subjects of intense investigation [1–9].

There is evidence in mouse that miRNAs may be implicated in ESC self-renewal and differentiation. Murine ESCs with either reduced *Dicer1* or absent *Dgcr8*, enzymes necessary for miRNA processing, displayed proliferation defects. In addition, the *Dgcr8* knockouts showed accumulation of cells in G1, which may point to alterations in regulation of cell cycle in these mutant cells [10, 11]. Both mutant murine ESC lines retained expression of pluripotency markers but were not able to differentiate normally [11, 12]. Of note, *Dicer1*-null homozygous mouse embryos appeared to be unable to produce normal ESCs [13].

Previous reports on miRNAs in ESCs include two studies describing isolation and cloning of novel miRNAs, one in murine ESCs [14] and one in hESCs [15]. These authors confirmed differential expression of a subset of the cloned miRNAs in ESCs by Northern blot. An additional four studies measured miRNA expression in murine ESCs using quantitative reverse transcription (qRT)-polymerase chain reaction (PCR) [16, 17] and using a microarray-based platform [18, 19]. In all six studies, two clusters of miRNAs were found to be strongly

expressed in ESCs (mir-302, mmu-mir-290/hsa-mir-371/372/373). The mir-290 cluster has also been noted to be expressed in trophoblast stem cells, suggesting that it may play a role in cellular self-renewal [14, 20].

A recent study reporting the largest miRNA cloning and sequencing effort to date included two samples of murine ESCs [21]. This study involved sequencing ~330,000 clones from 256 small RNA libraries from a wide variety of organs from human, mouse, and rat. The limited sample replication and low clone counts (only ~1,000 clones per library were sequenced) make it difficult to glean statistically significant differential expression information from this data set, but the murine ESC data are generally consistent with the miRNA expression results generated by the other methods discussed above.

The unique biology of miRNAs, as well as limitations in detection and quantitation methods for these small RNAs, has made it difficult to understand their functions in higher animals. It appears that there are likely to be more than 1,000 miRNAs in animals. Overexpression experiments indicate that each miRNA can downregulate 100–200 transcripts [22]. Also, transcripts may contain multiple miRNA target sequences in their 3′-untranslated regions and hence be regulated by more than one miRNA. Furthermore, there are classes of closely related but not identical miRNAs that differ at only one or a few nucleotides.

The small size of miRNAs and the existence of closely related types create technical difficulties for detection methods. Traditional methods, such as cloning and Northern blot, are time-consuming and are limited by the low abundance of some miRNAs. Direct hybridization methods are neither sensitive nor specific enough for this application. qRT-PCR methods are sensitive, specific, and quantitative but are impractical for profiling large numbers of genes in multiple samples.

Here, we describe the application of a novel, robust, and highly reproducible microarray method to generate global miRNA profiles of hESCs, neural stem cells (NSCs)/neural progenitor cells (NPCs), mesenchymal stem cells (MSCs), and differentiated cells (including a cell line differentiated from an hESC line) and the identification of cell-type-specific differences in miRNA usage that may regulate self-renewal and pluripotency.

MATERIALS AND METHODS

Cell Culture

Extraembryonic endoderm (XE) cells were differentiated from WA09 cells in hESC medium [23] with 20 ng/ml basic fibroblast growth factor (bFGF) on Matrigel (BD Biosciences, San Diego, http://www.bdbiosciences.com) in the absence of feeders. By immunofluorescent antibody staining and gene expression profiling, XE cells do not express ESC-specific markers and do express markers that are associated with primitive endoderm (R. Gonzalez, unpublished).

XE cells were differentiated from WA09 cells in hESC medium [23] with 20 ng/ml bFGF on Matrigel in the absence of feeders. XE cells are predominantly euploid (supplemental online Fig. 8), polygonal, flat cells that grow in monolayer and resemble fibroblasts. By immunofluorescent antibody staining and gene expression profiling, XE cells do not express ESC-specific markers (POU5F1/OCT4, LIN28, EBAF, UTF1, and ZFP42/REX) and do express markers that are associated with primitive endoderm (GATA6, DAB2, SPARC/osteonectin, PLAT, and PLAU) (R. Gonzalez, unpublished). The XE cells are genotypically identical to the parent WA09 cells by SNP genotyping using the Illumina Hap550 platform. The SNP genotyping results between the XE cells and two WA09 samples were 99.994% and 99.996% identical, whereas the results between the two WA09 samples were 99.997% identical. These results are within the error of the platform. Unrelated samples

are typically ~75% identical. All other cell types were derived and propagated as described in the references listed in Table 1.

RNA Purification

Total RNA, including miRNA, was purified from all cell types using the mirVANA miRNA Isolation Kit (Ambion, Austin, TX, http://www.ambion.com). Total RNA quantitation was performed using a NanoDrop N-1000 spectrophotometer (NanoDrop, Wilmington, DE, http://www.nanodrop.com). RNA quality was demonstrated using the Bio-Rad Experion Automated Electrophoresis System (RNA Standard Sensitivity Kit; Bio-Rad, Hercules, CA, http://www.bio-rad.com).

DNA Purification

Genomic DNA was purified from WA09 and XE cells using the DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany, http://www1.qiagen.com).

Microarray Quantitation of miRNA Expression

Microarray-based miRNA expression profiling was performed using a novel method (J.-B. Fan et al., manuscript in preparation). The method was a modification of the high-throughput gene expression profiling assay, the cDNA-mediated Annealing, Selection, Extension, and Ligation Assay, developed previously [24]. It applied a solid-phase primer extension (after target hybridization) to enhance the discrimination among homologous miRNA sequences. In addition, PCR with universal primers was used to amplify all targets prior to array hybridization.

One specific assay oligonucleotide was designed for each miRNA, consisting of three parts: at the 5′ end was a universal PCR priming site; in the middle was an address sequence, complementary to a corresponding capture sequence on the array; and at the 3′ end was a miRNA-specific sequence. Seven hundred assay probes were designed, corresponding to 397 well-annotated human miRNA sequences (miRBase, version 9.0 [October 2006]; The Wellcome Trust Sanger Institute, Cambridgeshire, England, http://microrna.sanger.ac.uk) and 303 miRNAs identified recently from human and chimpanzee brain [25].

Pooled assay oligonucleotides corresponding to the 700 human miRNAs are first annealed to cDNA. An allele-specific primer extension step is then carried out; the assay oligonucleotides are extended only if their 3′ bases are complementary to their cognate sequence in the cDNA template. The extended products are then amplified by PCR using common primers, of which one is fluorescently labeled, and hybridized to a microarray bearing the complementary address sequences. The DASL process, array image processing, and signal extraction were as described previously [24].

miRNA Microarray Data Processing

Data preprocessing was performed in BeadStudio version 2.0 (Illumina, Inc., San Diego, http://www.illumina.com). Data from each microarray was quantile-normalized using Expander (Ron Shamir, http://acgt.cs.tau.ac.il/expander/expander.html) [26]. miRNAs undetectable in all samples were removed. Technical replicates were averaged, and then biological replicates were averaged. Details on further analysis are given in supplemental online data, part 2.

Data Analysis

Additional details are given in supplemental online data, part 3.

t Test.   For the hESC versus non-hESC analysis, Welch's t test was performed with a p value cutoff of .05 and multiple testing correction by false discovery rate (implemented in GeneSpring [27]).

Consensus Clustering.   Consensus clustering was performed using Pearson distance and average linkage [28] (implemented in GenePattern (Broad Institute of MIT and Harvard, http://www.broad.mit.edu/cancer/software/genepattern) [29, 30]). For each value of k from 2 to 10, 100 iterations were performed. The consensus cumulative distribution function (cdf) and $\triangle$ area plots

Table 1. Cell lines analyzed, description of the cell lines, number of biological replicates, contributing collaborators, and relevant citations

Sample name	Description	No. of biological replicates	Source	Reference
HUES7	Undifferentiated human embryonic stem cell	2	CJL	[38]
HUES13	Undifferentiated human embryonic stem cell	2	CJL	[38]
WA09	Undifferentiated human embryonic stem cell	2	LCL	[39]
WA01	Undifferentiated human embryonic stem cell	2	PHS	[39]
HSF6	Undifferentiated human embryonic stem cell	2	HSP	[40]
SM2	Primary fetal neural progenitor cells	2	SRM[a]	[44]
SM3	Primary fetal neural progenitor cells	2	SRM[a]	[44]
HFT13	Fetal neural stem cell line	2	DRW	[42]
2050	Fetal neural stem cell line	2	DRW	[43]
SC23	Neural stem cell line from 28-week gestation	2	PHS	[44], [45]
HANSE2	Primary adult neural progenitor cells	3	NOS	[44]
HANSE3	Primary adult neural progenitor cells	3	NOS	[44]
HANSE4	Primary adult neural progenitor cells	3	NOS	[44]
HANSE5	Primary adult neural progenitor cells	3	NOS	[44]
SC01	Primary glial cell line	2	PHS	[44], [45]
SC11	Primary glial cell line	2	PHS	[44], [45]
SC30	Primary dermal fibroblast cell line	2	PHS	Sly 1979
SC31	Bone marrow mesenchymal stem cell line, CD105+, CD34−	2	PHS	Sly 1979
SC33	Primary dermal fibroblast cell line	2	PHS	Sly 1979
SC41	Bone marrow mesenchymal stem cell line, CD105+, CD34−	2	PHS	Sly 1979
XE	Extraembryonic endoderm phenotype, differentiated from WA09	3	LCL	R. Gonzalez, unpublished
HS27	Neonatal foreskin fibroblast	3	LCL	ATCC CRL-1634
HUVEC-AA	Primary human umbilical vein endothelial cells, black patients	4	DC	[47]
HUVEC-Cauc	Primary human umbilical vein endothelial cells, white patients	4	DC	[47]
BEWO	Choriocarcinoma cell line	2	DC	[48]
JEG3	Choriocarcinoma cell line	2	DC	[49]

[a]Tissue source: Advanced Bioscience Resources, Inc., Alameda, CA, http://www.abr-inc.com.
Abbreviations: ATCC, American Type Culture Collection, Manassas, VA, http://www.atcc.org; CJL, Christina J. Lu, Department of Reproductive Medicine, The Burnham Institute, University of California San Diego; DC, Dongbao Chen, Department of Reproductive Medicine, University of California San Diego; DRW, Dustin R. Wakeman, Department of Biomedical Sciences, The Burnham Institute, University of California San Diego; HSP, Hyun-Sook Park, Mizmedi Hospital, Seoul National University; LCL, Louise C. Laurent, Department of Reproductive Medicine, The Burnham Institute, University of California San Diego; NOS, Nils O. Schmidt, Department of Neurosurgery, Universitätsklinikum Hamburg-Eppendorf; PHS, Phillip H. Schwartz, Children's Hospital of Orange County; SRM, Scott R. McKercher, The Burnham Institute.

were examined, and $k = 6$ was determined to be the model with the smallest k for which the consensus cdf plot approximated the ideal step function, with insignificant proportional increases in the $\triangle$ area with increasing k values above 6 (supplemental online Fig. 3).

miRNA Grouping. miRNA grouping was performed using the Cluster Identification via Connectivity Kernels algorithm [26] via the Expander software [31]. Computation of p values to determine significance of overlaps between miRNA groups and annotations were performed by computing the tail of the hypergeometric distribution [32].

miRNA Clustering Analysis. miRNAs were considered to belong to the same genomic cluster if the genomic locations of the first nucleotides of the predicted pre-miRNA hairpins were within 50 kilobases (kb) (as suggested previously [33]).

Seed Similarity Graph. miRNA seed sequences were aligned using the Needleman-Wunsch algorithm [34]. A similarity graph was constructed, where edges connected miRNA pairs with six or seven identical positions in the alignment. The graph was subsequently clustered using Cluster Affinity Search Technique [35]. The clustering results were displayed using Cytoscape (http://www.cytoscape.org) [36].

Consensus Seed Sequence Identification. Consensus seed sequences for groups of miRNAs with related seed sequences upregulated in hESCs relative to non-hESCs were calculated using ClustalW [37].

Spectral Karyotyping

Cells were harvested and karyotyped [23]. Karyotyping was done using SkyPaint and SkyView software according to the manufacturer's instructions (Applied Spectral Imaging, Vista, CA, http://www.spectral-imaging.com.).

RESULTS

We used a novel microarray-based method (described in Materials and Methods) to determine the expression of 397 mature human miRNAs listed in the Sanger database (version 9.0 [October 2006]) and of 303 miRNAs recently identified in human brain [25] in 62 samples representing 26 cell lines, including hESCs, NSCs, NPCs, MSCs, and differentiated cells [38–49] (Table 1). There were two to four biological replicates per cell line and two technical replicates per biological replicate (details are given in supplemental online data, part 4). Raw data are given in supplemental online Table 1. After preprocessing and filtering, bioinformatic analysis techniques were applied to the data (diagram of experimental design is given in supplemental online Fig. 1). We verified that reproducibility of technical and biological replicates was excellent and that the reported results are robust to the number of biological replicates used (supplemental online data parts 1, 2; supplemental online Fig. 2).

miRNAs Differentially Expressed Between hESCs and Differentiated Cells Are Spatially Coregulated

We initially focused on miRNAs differentially expressed in hESCs compared with the other cell types. miRNA genes occur in the genome as independently transcribed units, in introns of coding
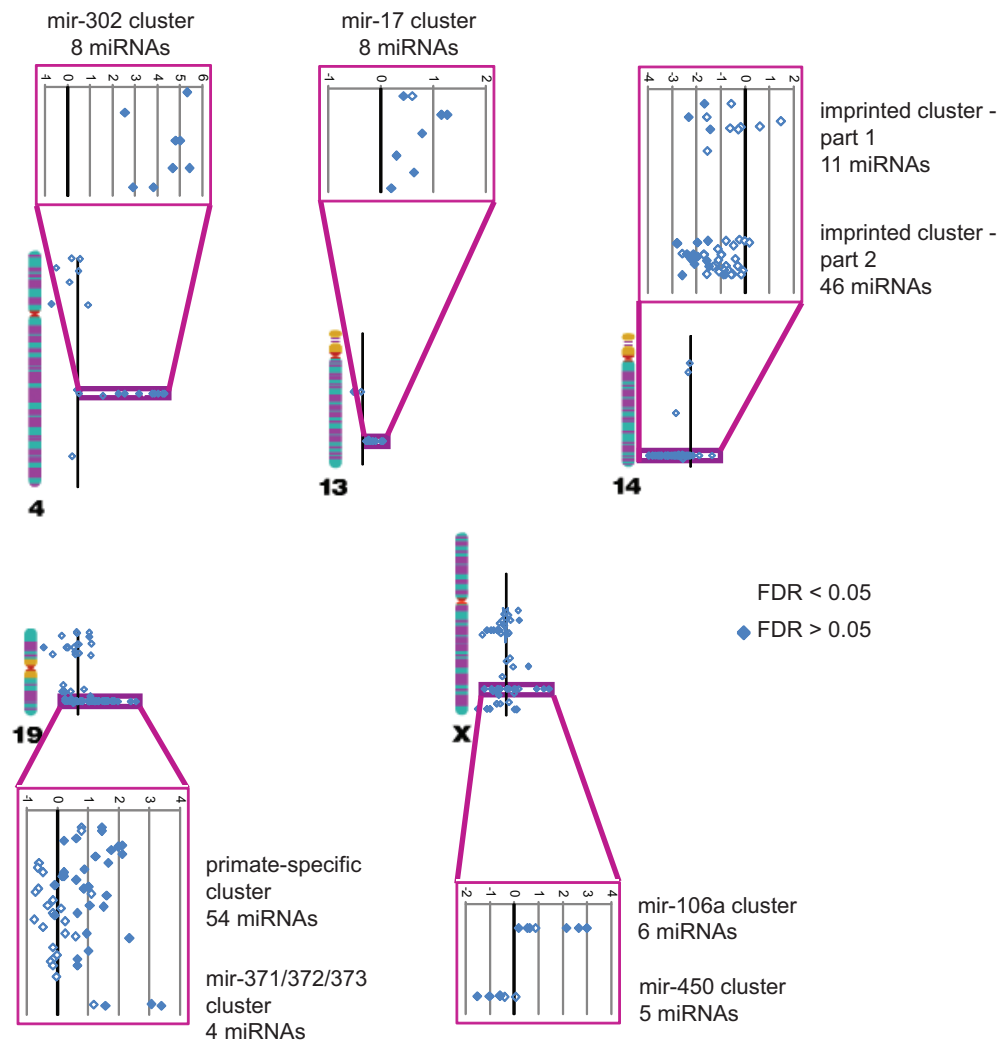
Figure 1. Prominent clusters of miRNAs showing differential expression in human embryonic stem cells (hESCs) compared with non-hESCs. The log2 ratio between the average miRNA expression in the hESC samples and that in the non-hESC samples is mapped by genomic location. Points to the left of the black lines have lower relative expression in hESCs, whereas points to the right of the black lines have higher relative expression in hESCs. Solid blue diamonds: miRNAs differentially expressed at an FDR <0.05, open blue diamonds: the rest of the miRNAs. The x-axis indicating the log2 ratio is the same scale for all chromosomes; only the x-axis for chromosome 1 is shown. Highlighted clusters: chromosome 4 (mir-302 cluster), chromosome 13 (mir-17 cluster), chromosome 14 (bipartite imprinted cluster), chromosome 19 (primate-specific cluster and mir-371/372/373 cluster), and X chromosome (mir-106a cluster). For the highlighted clusters, the log2 ratios are shown on the x-axes. The images of the chromosome are from the U.S. Department of Energy Genome Programs (http://genomics.energy.gov). Note that the t test takes variance into account, and therefore genes with higher log2 expression ratios do not necessarily have a more significant differential expression. Abbreviations: FDR, false discovery rate; miRNA, microRNA.

genes, and in clusters that are transcribed as polycistrons [50, 51]. When the differential expression was plotted against the genomic location, it was apparent that a large proportion of the differentially expressed miRNAs in hESCs occurred in clusters (the most prominent clusters are shown in Fig. 1; data plotted across all chromosomes are shown in supplemental online Fig. 3).

The most prominent upregulated clusters are found on chromosomes 4, 13, 19, and X. The mir-302 cluster, located on chromosome 4, has been associated with murine and human ESCs [14–16]. Chromosome 19 contains two subclusters located 25 kb apart, an ESC-associated cluster consisting of hsa-mir-371/372/373 [14–16] and a large primate-specific placenta-associated cluster containing 54 miRNAs spanning 96 kb [52]. Two paralogous clusters occur on chromosome 13 (mir-17 cluster) and the X chromosome (mir-106a cluster). The chromosome 13 cluster is associated with a number of cancers [50] and has been shown to be upregulated by *MYC* and to downregulate *E2F1* [53]. Interestingly,

in mouse, *Myc* has been shown, in combination with *Sox2*, *Pou5f1/Oct4*, and *Klf4*, to be sufficient for transforming somatic cells into ESC-like induced pluripotent stem cells capable of germline transmission [54–56].

A large bipartite cluster on chromosome 14 (11 and 46 miRNAs spanning 59 kb and 45 kb, respectively) is downregulated in hESCs. This cluster is located downstream of the reciprocally expressed imprinted *DLK1* and *GTL2/MEG3* genes. This cluster was first identified in mouse [57] and noted to be a maternally expressed imprinted cluster, with expression controlled by an intergenic differentially methylated region located between the *DLK1* and *GTL2/MEG3* genes.

Identification of a Large Number of miRNAs Not Previously Associated with ESCs

To identify hESC-specific expression of miRNAs, we extracted a list of 150 miRNAs that were significantly differentially
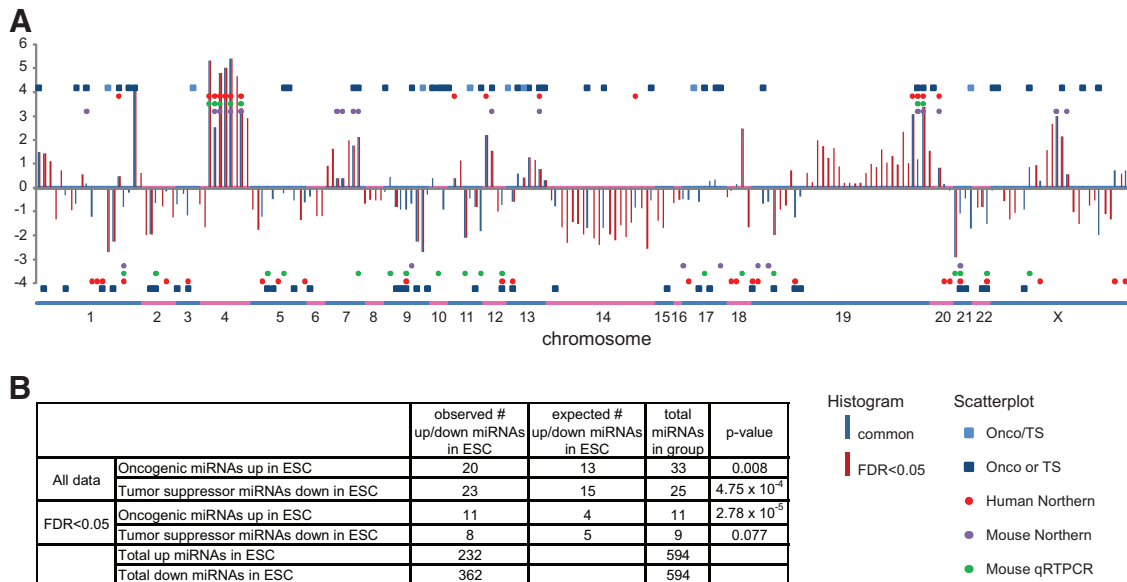
A

B

		observed # up/down miRNAs in ESC	expected # up/down miRNAs in ESC	total miRNAs in group	p-value
All data	Oncogenic miRNAs up in ESC	20	13	33	0.008
	Tumor suppressor miRNAs down in ESC	23	15	25	4.75×10^{-4}
FDR<0.05	Oncogenic miRNAs up in ESC	11	4	11	2.78×10^{-5}
	Tumor suppressor miRNAs down in ESC	8	5	9	0.077
	Total up miRNAs in ESC	232		594	
	Total down miRNAs in ESC	362		594	

Histogram
common
FDR<0.05

Scatterplot
Onco/TS
Onco or TS
Human Northern
Mouse Northern
Mouse qRTPCR

Figure 2. Data from this report compared with results from previous studies on miRNAs in ESCs or cancer, aligned by genomic location. **(A):** The log2 ratio between the average miRNA expression in the human embryonic stem cell (hESC) samples and in non-hESC samples is presented in the histogram, with the log2 ratio indicated on the y-axis. Only miRNAs for which there is statistically significant differential expression (FDR <0.05) and/or data from previous reports are shown. A linear representation of the genome below the graph is mirrored on the x-axis. miRNAs are evenly spaced on the x-axis, so genomic distances are not scaled. Red bars: miRNAs with significant differential expression. Blue bars: miRNAs with previously published data. Data from previous studies are shown by squares and circles; the distance from the x-axis for these points is arbitrary, as previous reports were largely qualitative. Light blue squares: miRNAs previously described as both Onco and TS. Dark blue squares: miRNAs previously designated as Onco (above the x-axis) or TS (below the x-axis). Red, green, and lavender circles: data from hESCs by Northern blot [15], mouse ESCs by qRTPCR [16, 17], and mouse ESCs by Northern blot [14], respectively. Circles above the x-axis have higher expression in ESCs, and circles below the x-axis have lower expression in ESCs, in relation to control cells used in those reports. **(B):** Table showing the number of Onco miRNAs upregulated and the number of TS miRNAs downregulated in the hESCs. Expected values were rounded to the nearest integer. *p* values are according to one-tailed *t* test. Abbreviations: FDR, false discovery rate; miRNA, microRNA; Onco, oncogenic; qRTPCR, quantitative reverse transcription polymerase chain reaction; TS, tumor suppressor.

expressed in hESCs (false discovery rate [FDR] <0.05). Seventy-six were upregulated, and 74 were downregulated (Fig. 2). Previous reports have identified 37 of these to be differentially regulated in murine or human ESCs [14–16].

Figure 2 is a plot of our differential expression results compared with findings from the previous reports. For the miRNAs with data available from earlier reports, the concordance in assignment of upregulation and downregulation is good, particularly when the findings have been reported in more than one study. However, we have discovered novel hESC-specific differences in miRNA expression. These include the large downregulated cluster on chromosome 14 and the large upregulated cluster on chromosome 19. We have confirmed these findings by qRT-PCR (not shown).

Oncogenic miRNAs Are Upregulated and Tumor Suppressor miRNAs Are Downregulated in hESCs

A number of miRNAs have been associated with human cancers [58, 59]. On the basis of the literature, cancer-related miRNAs can be categorized as oncogenic or tumor suppressor. In addition to experimental evidence pointing to the roles of these miRNAs in human cancers, these miRNAs have been shown to target mRNAs whose products have significant roles in cancer [60, 61]. Our data indicated that the oncogenic miRNAs were significantly upregulated ($p = .008$), and the tumor suppressor miRNAs were significantly downregulated ($p = 4.75 \times 10^{-4}$), in hESCs (Fig. 2).

hESCs Possess a Distinct miRNA Profile

To investigate the potential utility of miRNA profiling in classifying diverse cell types, we performed unsupervised consensus

clustering of cell samples using an agglomerative hierarchical clustering algorithm, which showed that there were four major cell sample clusters (Fig. 3A; supplemental online Fig. 4). The hESC samples were all found in a single cluster, containing two subclusters. The neural lineage cells, on the other hand, partitioned into an adult neural progenitor cell (aNPC) cluster and a fetal neural stem cell (fNSC) cluster. The MSCs and differentiated cell types, including the fibroblasts, human umbilical vein endothelial cells (HUVECs), glial cells, and XE cells, formed a fourth major cluster. These four major clusters were also found when the data were analyzed using two other unsupervised clustering methods, non-negative matrix factorization and K-means (supplemental online data, part 5; supplemental online Figs. 5, 6).

The two choriocarcinoma cell lines (BEWO and JEG3) were not closely associated with any of the four major clusters or with each other. However, the JEG3 line shares characteristics with members of the fNSC cluster (Fig. 3). These results demonstrate the utility of unsupervised classification in discovering differences among ostensibly closely related cell types.

miRNA Expression Profiles Distinguish Categories of Cell Types

To better understand the relationships between cell types, we also identified groups of miRNAs that share patterns of expression (we are using the term "grouping" here to avoid confusion with the term "miRNA cluster" used previously because the latter indicates spatial clustering of miRNA genes in the genome). Using an algorithm that maximizes both the within-group homogeneity and the between-group separation [26], we formed seven miRNA groups, each having a distinct expression
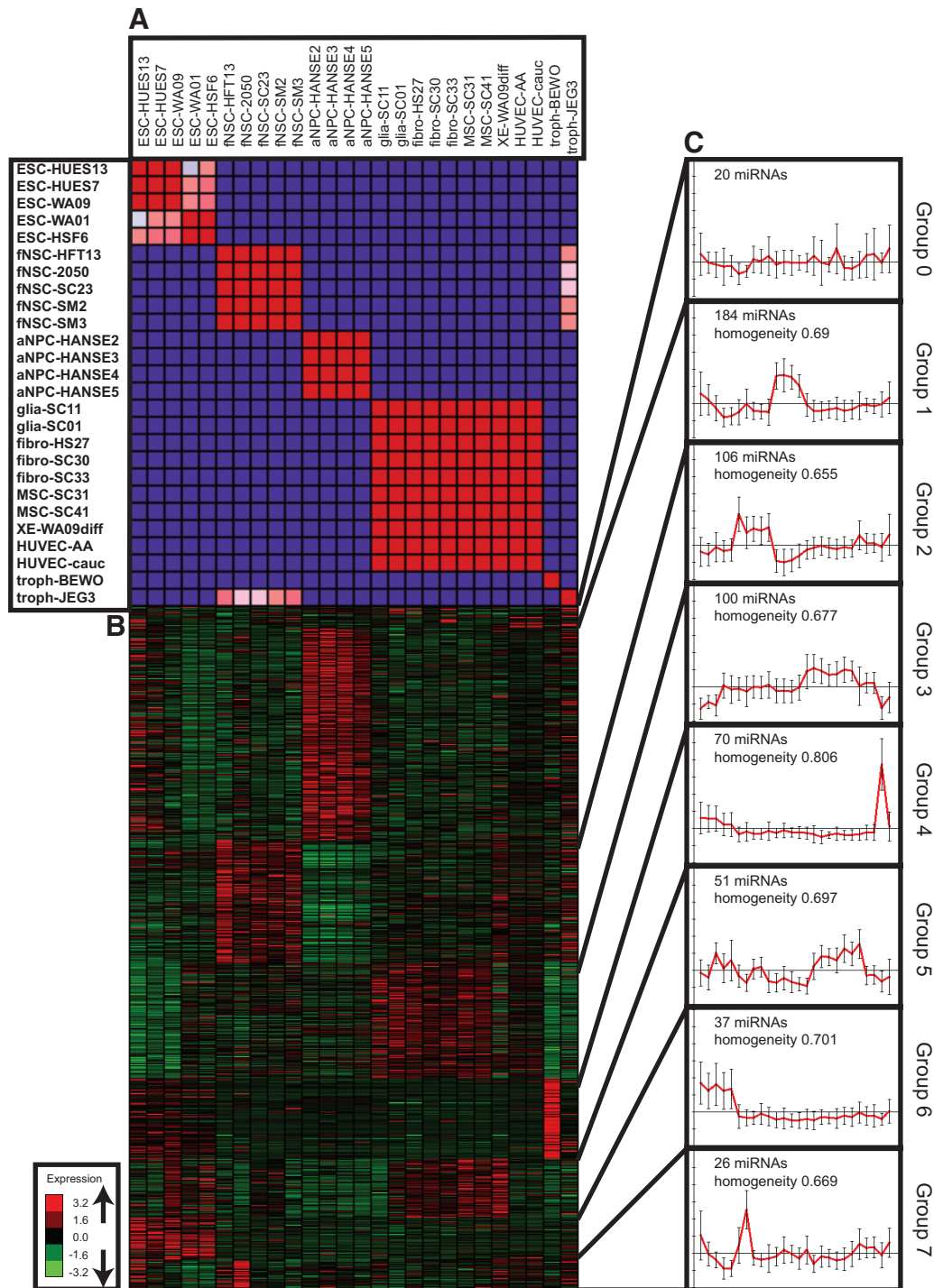
Figure 3. Consensus clustering and miRNA grouping results. **(A):** Consensus clustering matrix for the 26 cell lines. Results are based on 100 iterations of an unsupervised hierarchical clustering algorithm [28] (implemented in GenePattern [29, 30]). **(B):** Heat map showing expression profiles for the 26 cell lines, ordered according to the consensus clustering matrix in **(A)**. miRNAs in the heat map are grouped according the miRNA grouping results in **(C)**. **(C):** Average expression patterns of miRNA groups in the 26 cell lines. The order of cell lines is the same as in **(A)**. The number of miRNAs in each group and the homogeneity values are shown. Bars indicate ± 1 SD. Abbreviations: aNPC, adult neural progenitor cell; fibro, fibroblast; fNSC, fetal neural stem cell; HUVEC, human umbilical vein endothelial cell; miRNA, microRNA; MSC, mesenchymal stem cell; troph, trophoblast; XE, hESC-derived extraembryonic endoderm-like cell.

pattern (Fig. 3B, 3C). Only 20 miRNAs did not fall into one of the seven expression pattern groups (Fig. 3C, group 0; Table 2).

The chromosomal locations of the miRNAs in each group were nonrandom (supplemental online Fig. 7). For instance, the groups that were differentially expressed in hESCs were en-

riched for members that were clustered in the genome and depleted of members located in introns of coding genes (supplemental online Fig. 7).

Table 2 summarizes the composition of each miRNA group and surveys the available knowledge about the component

Table 2. Composition of the eight miRNA groups indicated in Figure 3

Groups 0–3	Number detected	Group 0: 20 miRNAs			Group 1: 184 miRNAs		
		Observed	Expected	p value[a]	Observed	Expected	p value[a]
Total	594						
Human brain-expressed	208	5	7	0.006	109	64	2.52×10^{-16b}
CNS-related	51	0	2	1	3	16	1
REST-related	13	0	0	1	1	4	1
Oncogenic	33	1	1	0.688	1	10	1
Tumor suppressor	25	2	1	0.204	3	8	1
Chromosome 14 cluster	44	0	1	1	8	14	1
Chromosome 19 cluster	44	0	1	1	1	14	1

Groups 4–7	Number detected	Group 4: 20 miRNAs			Group 5: 184 miRNAs		
		Observed	Expected	p value[a]	Observed	Expected	p value[a]
Total	594						
Human brain expressed	208	7	25	1	32	18	2.25×10^{-5b}
CNS-related	51	1	6	0.999	2	4	0.948
REST-related	13	0	2	1	1	1	0.693
Oncogenic	33	6	4	0.181	4	3	0.312
Tumor suppressor	25	1	3	0.959	0	2	1
Chromosome 14 cluster	44	0	5	1	1	4	0.984
Chromosome 19 cluster	44	42	5	4.70×10^{-43b}	1	4	0.984

(Continued)

miRNAs. miRNAs have been associated in previous studies with organs, chromosome locations, activation by specific transcription factors, and other properties. We focused on seven of these associations. Our data set contained 208 miRNAs originally identified by deep sequencing of small RNAs from human brain (human brain-expressed) [25]; 51 miRNAs that have been described in the literature as central nervous system (CNS)-related [62–64]; 13 miRNAs regulated by REST (REST-regulated), which is involved in repressing neural-specific genes in non-neural cells [65]; 33 oncogenic miRNAs [59]; 25 tumor suppressor miRNAs [59]; 44 miRNAs in the chromosome 14 cluster; and 44 miRNAs in the chromosome 19 cluster.

Groups 1 and 2 accounted for the bulk of the differences between the aNPCs and fNSCs. Group 1 contained 184 miRNAs that were upregulated in aNPCs relative to the other cell types. This group was significantly enriched for human brain-expressed miRNAs. Group 2 consisted of 106 miRNAs upregulated in fNSCs and mildly downregulated in the aNPCs. The CNS-related and REST-regulated miRNAs were overrepresented in this group.

Group 3 contained 100 miRNAs and had a more complex profile. In general, the expression of the miRNAs in this group was correlated with the degree of differentiation, with lowest expression in the hESCs, intermediate expression in the NSCs/ NPCs, and highest expression in the differentiated cells. The tumor suppressor and CNS-related miRNAs were overrepresented in this group. Notably, most of the miRNAs in the large cluster on chromosome 14 also belonged to this group.

The 70 miRNAs in group 4 were mildly upregulated in the hESCs and very strongly expressed in the BEWO choriocarcinoma cells. Almost all of the members of the chromosome 19 cluster were found in group 4. Interestingly, this cluster of miRNAs has been strongly associated with the placenta [52].

Group 5 was a group of 51 miRNAs that had a profile of highest expression in the fibroblasts, MSCs, XE cells, and some of the hESC samples. A possible interpretation of this group is that it contains the miRNAs that are upregulated in fibroblast/ MSC-like cells. The sporadic upregulation of this group of miRNAs in some hESC samples may be due to a degree of heterogeneity in those cultures. The human brain-expressed miRNAs were overexpressed in this group as well. Group 6

contained 37 miRNAs that are uniquely upregulated in hESCs and included all of the members of the mir-302 cluster, which has been identified as ESC-specific in prior studies [14–16].

Group 7 consisted of 26 miRNAs that were upregulated in a diverse set of cell types, including one of the glial cell lines, the HUVECs, and the JEG3s. It was enriched in human brain-expressed miRNAs.

One Seed Sequence Dominates the Population of miRNAs Upregulated in hESCs

A number of miRNA target prediction models are based on the assumption that nucleotides 2–8 in the mature miRNA sequence constitute a seed sequence that contributes heavily to target mRNA specificity by Watson-Crick base-pairing [66–69]. We compared the seed sequences of hESC-upregulated miRNAs and observed that members of the mir-302 cluster (chromosome 4), the chromosome 19 cluster, the mir-17 cluster (chromosome 13), and the mir-106a cluster (X chromosome) have similar seed sequences. By grouping the miRNAs significantly upregulated (FDR <0.05) in hESCs relative to non-hESCs according to seed sequence, we identified four seed similarity clusters that share near-identical seed sequences (Fig. 4; only clusters containing at least five miRNAs were considered). The largest cluster has a consensus miRNA seed sequence (AAGTGC) that is dramatically overrepresented in hESCs (Fig. 4A; $p = 1.2 \times 10^{-14}$). In fact, 18 of the 21 miRNAs containing this seed sequence were among the 76 miRNAs upregulated in hESCs.

We also examined the cognate target mRNAs for the major (Fig. 4A) as well as the three minor seed similarity clusters (Fig. 4B–4D), as predicted by five miRNA target prediction algorithms, TargetScan [68], MirZ [70–72], PicTar [66], RNA22 [72], and Miranda [71]. With four of the prediction methods, and particularly with the methods using sequence/target sequence matching together with cross-species conservation (TargetScan, MirZ, and PicTar), the targets were significantly enriched for genes with promoters bound by the ESC-associated transcription factors *NANOG*, *SOX2*, and *POU5F1/OCT4* [73] (supplemental online Table 1). In many cases, the Gene Ontology categories associated with transcription were also overrepresented in the predicted mRNA targets (supplemental online

Table 2. (Continued)

	Group 2: 106 miRNAs			Group 3: 100 miRNAs	
Observed	Expected	p value[a]	Observed	Expected	p value[a]
6	37	1	17	35	1
20	9	1.30×10^{-4b}	51	9	4.81×10^{-46b}
6	2	0.016	5	2	0.051
8	6	0.220	9	6	0.084
3	4	0.855	16	4	7.77×10^{-8b}
0	8	1	34	7	1.61×10^{-20b}
0	8	1	0	7	1

	Group 6: 106 miRNAs			Group 7: 100 miRNAs	
Observed	Expected	p value[a]	Observed	Expected	p value[a]
10	13	0.893	19	9	6.00×10^{-5b}
4	3	0.395	0	2	1
0	1	1	0	1	1
4	2	0.142	0	1	1
0	2	1	0	1	1
1	3	0.947	0	2	1
0	3	1	0	2	1

Properties of the miRNAs in each group are described in the text. For each category of miRNA (human brain-expressed, CNS-related, etc.), the number of miRNAs with detectable expression in at least one cell type, the number observed in each group, and the number expected to be in each group are listed. p values were calculated according to a one-tailed test, assuming a hypergeometric distribution. The Bonferroni corrected α is $0.05/56 = 9 \times 10^{-4}$.

[a]p value: one-tailed test, hypergeometric distribution. Bonferroni corrected $\alpha < 0.05/56 = 9 \times 10^{-4}$.

[b]Significant p values.

Abbreviations: CNS, central nervous system; miRNA, microRNA.

Table 1). The zebrafish embryo expresses a very abundant miRNA family, the mir-430 family, which contains the same consensus seed sequence as the major hESC-upregulated seed similarity cluster (Fig. 4A). Remarkably, the targets of this cluster, as predicted by TargetScan and MirZ, were significantly enriched for homologs of mRNAs found to be upregulated in the zebrafish embryo when the mir-430 family was knocked out [74] (supplemental online Table 1).
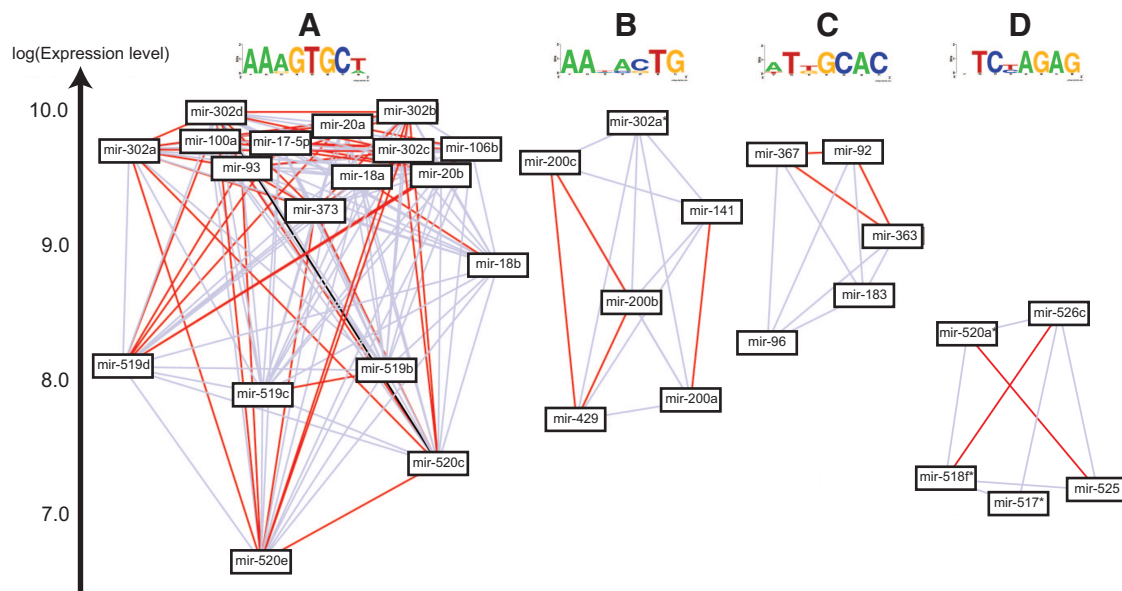

Figure 4. Seed similarity graph and consensus seed sequences. MicroRNAs (miRNAs) significantly upregulated (false discovery rate <0.05) in human embryonic stem cells (hESCs) relative to non-hESCs are grouped by seed sequence. Nodes are labeled with the miRNA name. The average expression of the miRNAs in hESCs is represented by the log intensity on the y-axis. Red lines connect miRNAs that have identical seed sequences (mature miRNA positions 2–8). Blue lines connect miRNAs with six of seven matches in the seed sequence. Consensus seed motifs are shown for each cluster. miRNAs that appeared in a similarity cluster with fewer than four other upregulated miRNAs are not shown (23 singleton miRNAs that do not have seed sequence similarity with any other upregulated miRNAs, five groups consisting of three miRNAs, and one group of four miRNAs). Lines (edges) between miRNAs in different similarity clusters are not shown. **(A):** The major seed similarity cluster upregulated in hESCs. **(B–D):** Three minor seed similarity clusters upregulated in hESCs.

DISCUSSION

Molecular profiling of hESCs in the context of a diverse collection of somatic stem cells and differentiated cells can be used to identify a unique hESC molecular signature, to identify potentially shared pathways between cell types that otherwise are quite dissimilar, and to identify candidate molecular species that warrant functional analysis. The results we report here constitute the most comprehensive determination of miRNA expression in hESCs to date. The strengths of our approach include a highly specific, reproducible, high-throughput platform for measurement of miRNA expression and a well-curated collection of diverse cell types, including hESCs, NSCs/NPCs, MSCs, and differentiated cells. We report on a large number of novel miRNA-hESC associations, as well as additional evidence of hESC-specific expression for miRNAs previously described to be upregulated in hESCs. Our results indicate that a large proportion of miRNAs differentially expressed in hESCs occur in clusters in the genome, including two very large clusters that have not been associated previously with hESCs. One of these clusters, located in an imprinted region of chromosome 14, is of particular interest. Given the imprinted nature of this region [75] and the fact that aneuploidies and subchromosomal deletions commonly occur during propagation of hESC lines, it is clearly important to examine this region in a number of hESC lines to look for changes in methylation status and for evidence of loss of heterozygosity.

We have found that oncogenic miRNAs, including the mir-17 cluster, which has been genetically and functionally associated with human cancers, are overrepresented in the miRNAs upregulated in hESCs, whereas tumor suppressor miRNAs are overrepresented in the miRNAs downregulated in hESCs. We note that the two choriocarcinoma lines do not cluster closely with the hESC lines or with each other. It has been observed that the profiles of different cancer cells are quite diverse [76], varying by diagnosis, stage, and prognosis (reviewed in [77]). For an individual cancer type, specific oncogenic miRNAs are overexpressed (and specific tumor suppressor miRNAs are downregulated), whereas others are not. In contrast, the large majority of oncogenic and tumor suppressor miRNAs were found to be differentially expressed in hESCs. This suggests that although hESCs are not closely related to any one type of cancer cell, they may share some mechanisms for self-renewal with cancer cells as a class.

We observed a large set of miRNAs upregulated in hESCs, whose sequences have near-identical seeds. This overrepresentation of a class of highly similar seed sequences points to a critical role for repressing a group of cognate mRNAs in maintaining the stem cell state. The mir-430 family, which appears to be responsible for rapid clearance of maternal mRNAs in the zebrafish embryo, has a seed sequence (AAGTGCT) that matches this consensus seed sequence [74]. The recently discovered mmu-mir-467 family, 10 copies of which are present in the mmu-mir-297b cluster on chromosome 2, also contains this seed sequence and was observed to have the highest clone counts in the mouse embryonic stem (ES) cells [21].

Previous studies from this and other laboratories have shown that hESCs have mRNA and DNA methylation profiles that are highly uniform across different hESC preparations (even when those preparations originate from different hESC lines derived and propagated under different conditions in different laboratories) and that are distinct from the profiles from a variety of other cell types [40, 78–80]. We have shown here that there is also a characteristic hESC miRNA profile that can be used to distinguish hESCs from all other cell types. We have also shown that in contrast to hESCs, cells that have been classified as neural stem and progenitor cells by classic marker studies can be categorized into two distinct subgroups by miRNA profiling. This is consistent with our comprehensive analysis of mRNA profiles in many of the same cell types (F.-J. Mueller et al., manuscript in review).

Grouping miRNAs by coherence in expression pattern across conditions revealed several specific miRNA expression patterns that highlighted a number of subtleties in the expression of related miRNAs. As noted above, the mir-302, hsa-mir-371/372/373, and large chromosome 19 clusters are all upregulated in hESCs and have similar seed sequences, suggesting that they target similar pools of mRNAs. However, we have seen that although the mir-302 cluster is expressed in a hESC-specific pattern, two of the other clusters upregulated in hESCs (hsa-mir-371/372/373 and the large chromosome 19 clusters) are also highly expressed in one of the trophoblast-type cell lines.

At least two regulatory mechanisms have been shown to result in coregulation of clustered miRNAs. There is evidence that two of the RNA polymerase II-transcribed miRNA clusters differentially expressed in ES cells are transcribed polycistronically. The mir-17-5p cluster was shown to have a primary transcript miRNA species that contained all the members of the cluster [50], whereas the mouse counterpart to the downregulated cluster on chromosome 14 was shown to have an initiating histone mark in the promoter region and elongation histone marks throughout the remainder of the cluster [81]. In contrast, the large miRNA cluster on chromosome 19 appears to be transcribed by RNA polymerase III, with RNA polymerase III promoters present in Alu repeats occurring between miRNA genes. Group-specific regulatory elements may be responsible for these different expression patterns. However, identifying such regulatory sequences will be challenging, as direct experimental evidence on the miRNA composition of many of the presumably polycistronic miRNA transcripts is not available, and the locations of most miRNA transcriptional start sites are unknown.

CONCLUSION

These findings significantly extend our knowledge and suggest potential roles for miRNAs in regulating cellular pluripotency and self-renewal. As noted above, there have been a number of reports on miRNA expression in embryonic stem cells [14–19]. Ours is the first study based on a well-replicated data set including samples from a large collection of homogeneous, well-characterized cells run on a novel high-throughput miRNA expression platform. The nonpluripotent cells in this study encompass a wide variety of cells, including somatic stem cells, tissue-derived differentiated cells, hESC-derived differentiated cells, and choriocarcinoma cells, which provide a broad context in which to place the hESC samples. The miRNA expression platform is extremely sensitive and specific, and it is the most comprehensive high-throughput platform available.

We conclude that miRNA profiling can be used for robust classification of diverse cells types and that hESCs possess a unique miRNA signature, with the upregulated miRNAs dominated by a single seed sequence. Our results point to specific individual miRNAs and families of miRNAs that are good candidates for future functional studies of the potential roles of miRNAs in the maintenance of the undifferentiated hESC state and in hESC differentiation.

DISCLOSURE OF POTENTIAL CONFLICTS OF INTEREST

L.C.L., J.C., J.-B.F., and J.F.L. own stock in Illumina.

REFERENCES

1 Bao N, Lye KW, Barton MK. MicroRNA binding sites in *Arabidopsis* class III HD-ZIP mRNAs are required for methylation of the template chromosome. Dev Cell 2004;7:653–662.
2 Llave C, Xie Z, Kasschau KD et al. Cleavage of Scarecrow-like mRNA targets directed by a class of *Arabidopsis* miRNA. Science 2002;297:2053–2056.
3 Mallory AC, Vaucheret H. Functions of microRNAs and related small RNAs in plants. Nat Genet 2006;38(suppl):S31–S36.
4 Mette MF, Aufsatz W, Kanno T et al. Analysis of double-stranded RNA and small RNAs involved in RNA-mediated transcriptional gene silencing. Methods Mol Biol 2005;309:61–82.
5 Olsen PH, Ambros V. The lin-4 regulatory RNA controls developmental timing in *Caenorhabditis elegans* by blocking LIN-14 protein synthesis after the initiation of translation. Dev Biol 1999;216:671–680.
6 Palatnik JF, Allen E, Wu X et al. Control of leaf morphogenesis by microRNAs. Nature 2003;425:257–263.
7 Rhoades MW, Reinhart BJ, Lim LP et al. Prediction of plant microRNA targets. Cell 2002;110:513–520.
8 Seggerson K, Tang L, Moss EG. Two genetic circuits repress the *Caenorhabditis elegans* heterochronic gene lin-28 after translation initiation. Dev Biol 2002;243:215–225.
9 Wightman B, Ha I, Ruvkun G. Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in *C. elegans*. Cell 1993;75:855–862.
10 Murchison EP, Partridge JF, Tam OH et al. Characterization of Dicer-deficient murine embryonic stem cells. Proc Natl Acad Sci U S A 2005;102:12135–12140.
11 Wang Y, Medvid R, Melton C et al. DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. Nat Genet 2007;39:380–385.
12 Kanellopoulou C, Muljo SA, Kung AL et al. Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. Genes Dev 2005;19:489–501.
13 Bernstein E, Kim SY, Carmell MA et al. Dicer is essential for mouse development. Nat Genet 2003;35:215–217.
14 Houbaviy HB, Murray MF, Sharp PA. Embryonic stem cell-specific MicroRNAs. Dev Cell 2003;5:351–358.
15 Suh MR, Lee Y, Kim JY et al. Human embryonic stem cells express a unique set of microRNAs. Dev Biol 2004;270:488–498.
16 Strauss WM, Chen C, Lee CT et al. Nonrestrictive developmental regulation of microRNA gene expression. Mamm Genome 2006;17:833–840.
17 Chen C, Ridzon D, Lee CT et al. Defining embryonic stem cell identity using differentiation-related microRNAs and their potential targets. Mamm Genome 2007;18:316–327.
18 Lakshmipathy U, Love B, Goff LA et al. MicroRNA expression pattern of undifferentiated and differentiated human embryonic stem cells. Stem Cells Dev 2007;16:1003–1016.
19 Josephson R, Ording CJ, Liu Y et al. Qualification of embryonal carcinoma 2102Ep as a reference for human embryonic stem cell research. STEM CELLS 2007;25:437–446.
20 Houbaviy HB, Dennis L, Jaenisch R et al. Characterization of a highly variable eutherian microRNA gene. RNA 2005;11:1245–1257.
21 Landgraf P, Rusu M, Sheridan R et al. A mammalian microRNA expression atlas based on small RNA library sequencing. Cell 2007;129:1401–1414.
22 Lim LP, Lau NC, Garrett-Engele P et al. Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. Nature 2005;433:769–773.
23 Loring JF, Wesselschmidt R, Schwartz P, eds. Human Stem Cell Manual: A Laboratory Guide. 1st ed. New York: Elsevier, 2007.
24 Fan JB, Yeakley JM, Bibikova M et al. A versatile assay for high-throughput gene expression profiling on universal array matrices. Genome Res 2004;14:878–885.
25 Berezikov E, Thuemmler F, van Laake LW et al. Diversity of microRNAs in human and chimpanzee brain. Nat Genet 2006;38:1375–1377.
26 Sharan R, Shamir R. CLICK: A clustering algorithm with applications to gene expression analysis. Proc Int Conf Intell Syst Mol Biol 2000;8:307–316.
27 Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. J R Stat Soc 1995;57:289–300.
28 Eisen MB, Spellman PT, Brown PO et al. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 1998;95:14863–14868.
29 Monti S, Tamayo P, Mesirov J et al. Consensus clustering: A resampling-based method for class discovery and visualization of gene expression microarray data. Machine Learning J 2003;52:91–118.
30 Reich M, Liefeld T, Gould J et al. GenePattern 2.0. Nat Genet 2006;38:500–501.
31 Shamir R, Maron-Katz A, Tanay A et al. EXPANDER—An integrative program suite for microarray data analysis. BMC Bioinformatics 2005;6:232.
32 Tavazoie S, Hughes JD, Campbell MJ et al. Systematic determination of genetic network architecture. Nat Genet 1999;22:281–285.
33 Baskerville S, Bartel DP. Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. RNA 2005;11:241–247.
34 Needleman S, Wunsch C. A general method applicable to the search for similarities in the amino acid sequence of two proteins. J Mol Biol 1970;48:443–453.
35 Ben-Dor A, Shamir R, Yakhini Z. Clustering gene expression patterns. J Comput Biol 1999;6:281–297.
36 Shannon P, Markiel A, Ozier O et al. Cytoscape: A software environment for integrated models of biomolecular interaction networks. Genome Res 2003;13:2498–2504.
37 Thompson JD, Higgins, DG, Gibson, TJ. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994;22:4673–4680.
38 Cowan CA, Klimanskaya I, McMahon J et al. Derivation of embryonic stem-cell lines from human blastocysts. N Engl J Med 2004;350:1353–1356.
39 Thomson JA, Itskovitz-Eldor J, Shapiro SS et al. Embryonic stem cell lines derived from human blastocysts. Science 1998;282:1145–1147.
40 Abeyta MJ, Clark AT, Rodriguez RT et al. Unique gene expression signatures of independently-derived human embryonic stem cell lines. Hum Mol Genet 2004;13:601–608.
41 White MG, Hammond RR, Sanders VJ et al. Neuron-enriched second trimester human cultures: Growth factor response and in vivo graft survival. Cell Transplant 1999;8:59–73.
42 Sidman RL, Li J, Stewart GR et al. Injection of mouse and human neural stem cells into neonatal Niemann-Pick A model mice. Brain Res 2007;1140:195–204.
43 Flax JD, Aurora S, Yang C et al. Engraftable human neural stem cells respond to developmental cues, replace neurons, and express foreign genes. Nat Biotechnol 1998;16:1033–1039.
44 Palmer TD, Schwartz PH, Taupin P et al. Cell culture. Progenitor cells from human brain after death. Nature 2001;411:42–43.
45 Schwartz PH, Bryant PJ, Fuja TJ et al. Isolation and characterization of neural progenitor cells from post-mortem human cortex. J Neurosci Res 2003;74:838–851.
46 Sly WS, Grubb J. Isolation of fibroblasts from patients. Methods Enzymol 1979;58:444–450.
47 Jaffe EA, Nachman RL, Becker CG et al. Culture of human endothelial cells derived from umbilical veins. Identification by morphologic and immunologic criteria. J Clin Invest 1973;52:2745–2756.
48 Pattillo RA, Gey GO. The establishment of a cell line of human hor-

mone-synthesizing trophoblastic cells in vitro. Cancer Res 1968;28: 1231–1236.

49 Kohler PO, Bridson WE. Isolation of hormone-producing clonal lines of human choriocarcinoma. J Clin Endocrinol Metab 1971;32:683–687.

50 He L, Thomson JM, Hemann MT et al. A microRNA polycistron as a potential human oncogene. Nature 2005;435:828–833.

51 Zhang B, Wang Q, Pan X. MicroRNAs and their regulatory roles in animals and plants. J Cell Physiol 2007;210:279–289.

52 Bentwich I, Avniel A, Karov Y et al. Identification of hundreds of conserved and nonconserved human microRNAs. Nat Genet 2005;37: 766–770.

53 O'Donnell KA, Wentzel EA, Zeller KI et al. c-Myc-regulated microRNAs modulate E2F1 expression. Nature 2005;435:839–843.

54 Okita K, Ichisaka T, Yamanaka S. Generation of germline-competent induced pluripotent stem cells. Nature 2007;448:313–317.

55 Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 2006;126:663–676.

56 Wernig M, Meissner A, Foreman R et al. In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. Nature 2007;448: 318–324.

57 Seitz H, Royo H, Bortolin ML et al. A large imprinted microRNA gene cluster at the mouse Dlk1-Gtl2 domain. Genome Res 2004;14: 1741–1748.

58 Kent OA, Mendell JT. A small piece in the cancer puzzle: MicroRNAs as tumor suppressors and oncogenes. Oncogene 2006;25:6188–6196.

59 Wiemer EA. The role of microRNAs in cancer: No small matter. Eur J Cancer 2007;43:1529–1544.

60 Johnson SM, Grosshans H, Shingara J et al. RAS is regulated by the let-7 microRNA family. Cell 2005;120:635–647.

61 Lee YS, Dutta A. The tumor suppressor microRNA let-7 represses the HMGA2 oncogene. Genes Dev 2007;21:1025–1030.

62 Cao X, Yeo G, Muotri AR et al. Noncoding RNAs in the mammalian central nervous system. Annu Rev Neurosci 2006;29:77–103.

63 Kawasaki H, Taira K. Functional analysis of microRNAs during the retinoic acid-induced neuronal differentiation of human NT2 cells. Nucleic Acids Res Suppl 2003;243–244.

64 Lagos-Quintana M, Rauhut R, Yalcin A et al. Identification of tissue-specific microRNAs from mouse. Curr Biol 2002;12:735–739.

65 Wu J, Xie X. Comparative sequence analysis reveals an intricate network among REST, CREB and miRNA in mediating neuronal gene expression. Genome Biol 2006;7:R85.

66 Krek A, Grun D, Poy MN et al. Combinatorial microRNA target predictions. Nat Genet 2005;37:495–500.

67 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.

68 Lewis BP, Shih IH, Jones-Rhoades MW et al. Prediction of mammalian microRNA targets. Cell 2003;115:787–798.

69 Stark A, Brennecke J, Bushati N et al. Animal MicroRNAs confer robustness to gene expression and have a significant impact on 3′UTR evolution. Cell 2005;123:1133–1146.

70 Gaidatzis D, van Nimwegen E, Hausser J et al. Inference of miRNA targets using evolutionary conservation and pathway analysis. BMC Bioinformatics 2007;8:69.

71 John B, Enright AJ, Aravin A et al. Human MicroRNA targets. PLoS Biol 2004;2:e363.

72 Miranda KC, Huynh T, Tay Y et al. A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. Cell 2006;126:1203–1217.

73 Boyer LA, Lee TI, Cole MF et al. Core transcriptional regulatory circuitry in human embryonic stem cells. Cell 2005;122:947–956.

74 Giraldez AJ, Mishima Y, Rihel J et al. Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. Science 2006;312: 75–79.

75 Wylie AA, Murphy SK, Orton TC et al. Novel imprinted DLK1/GTL2 domain on human chromosome 14 contains motifs that mimic those implicated in IGF2/H19 regulation. Genome Res 2000;10:1711–1718.

76 Lu J, Getz G, Miska EA et al. MicroRNA expression profiles classify human cancers. Nature 2005;435:834–838.

77 Calin GA, Croce CM. MicroRNA signatures in human cancers. Nat Rev Cancer 2006;6:857–866.

78 Bibikova M, Chudin E, Wu B et al. Human embryonic stem cells have a unique epigenetic signature. Genome Res 2006;16:1075–1083.

79 Cai J, Chen J, Liu Y et al. Assessing self-renewal and differentiation in human embryonic stem cell lines. STEM CELLS 2006;24:516–530.

80 Liu Y, Shin S, Zeng X et al. Genome wide profiling of human embryonic stem cells (hESCs), their derivatives and embryonal carcinoma cells to develop base profiles of U.S. Federal government approved hESC lines. BMC Dev Biol 2006;6:20.

81 Mikkelsen TS, Ku M, Jaffe DB et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature 2007; 448:553–560.

See www.StemCells.com for supplemental material available online.

Comprehensive MicroRNA Profiling Reveals a Unique Human Embryonic Stem Cell Signature Dominated by a Single Seed Sequence

Louise C. Laurent, Jing Chen, Igor Ulitsky, Franz-Josef Mueller, Christina Lu, Ron Shamir, Jian-Bing Fan and Jeanne F. Loring

Stem Cells 2008;26;1506-1516; originally published online Apr 10, 2008;

DOI: 10.1634/stemcells.2007-1081

This information is current as of November 30, 2008

Updated Information & Services	including high-resolution figures, can be found at: http://www.StemCells.com/cgi/content/full/26/6/1506
Supplementary Material	Supplementary material can be found at: http://www.StemCells.com/cgi/content/full/2007-1081/DC1

⍺ AlphaMed Press