RESEARCH ARTICLE

Embryonic stem cells derived from somatic cloned and fertilized blastocysts are post-transcriptionally indistinguishable: A MicroRNA and protein profile comparison

Junjun Ding[1,2], Yuanbiao Guo[3]*, Sheng Liu[1], Yujuan Yan[3], Gang Chang[1], Zhaohui Kou[1], Yu Zhang[1], Ying Jiang[3]**, Fuchu He[3]**, Shaorong Gao[1] and Jianli Sang[2]*

[1] National Institute of Biological Sciences (NIBS), Beijing, P. R. China
[2] College of Life Science, Beijing Normal University, Beijing, P. R. China
[3] Beijing Proteome Research Center, Beijing, P. R. China

Received: October 21, 2008
Revised: January 11, 2009
Accepted: January 14, 2009

Therapeutic cloning, whereby somatic cell nuclear transfer is used to generate customized embryonic stem cells (NT-ES) from differentiated somatic cells of specific individuals, has been successfully performed in mice and non-human primates. Safety concerns have prevented this technology from being potentially applied to humans, as severely abnormal phenotypes have been observed in cloned animals. Although it has been demonstrated that the transcriptional profiles and developmental potentials of ES cells derived from cloned blastocysts are identical to those of ES cells derived from fertilized blastocysts (F-ES), a systematic analysis of the post-transcriptional profiles of NT-ES cell lines has not yet been performed. To investigate whether NT-ES cells are comparable to F-ES cells post-transcriptionally, we compared the microRNA and protein profiles of five NT- and matching F-ES cell lines by microRNA microarray, 2-D DIGE and bioinformatic analyses. Stem-loop real-time PCR and MS/MS assays were further performed to verify the expression of specific microRNAs and characterize differentially expressed proteins. Our results demonstrate that the ES cell lines derived from cloned and fertilized mouse blastocysts have highly similar microRNA and protein expression profiles, consistent with their similar developmental potentials and transcriptional profiles.

Keywords:
ES cells / MicroRNA profiles / Proteomics / Somatic cell nuclear transfer

1 Introduction

Somatic cell nuclear transfer (SCNT) represents a remarkable process by which a differentiated somatic cell can be converted into a totipotent embryonic cell to subsequently generate a cloned animal [1–3]. More importantly, the derivation of customized NT-ES cells from SCNT embryos has been proposed to be one of the most attractive approaches for generating patient-specific pluripotent stem cells to treat many degenerative and genetic diseases [4–6]. The cloning efficiency is extremely low, however, with less than 5% of cloned embryos able to develop to term. Moreover, the cloned animals always exhibit severe fetal and postnatal

Correspondence: Dr. Shaorong Gao, #7 Science Park Road, Zhongguancun Life Science Park, Beijing 102206, P. R. China
E-mail: gaoshaorong@nibs.ac.cn
Fax: +86-10-80727535

Abbreviations: ES, embryonic stem; **F-ES,** fertilization-derived ES; **iPS,** induced pluripotent stem; **RT,** real time; **SCNT,** somatic cell nuclear transfer

* These authors contributed equally to this work.

** Additional corresponding author: Professor Fuchu He, e-mail: hefc@nic.bmi.ac.cn; Dr. Ying Jiang, e-mail: jiangy@hupo.org.cn

abnormalities, such as obesity [7] and premature death [8]. The expression patterns of imprinted genes appear normal in apparently ''normal'' cloned mice [9], but many epigenetic defects have been found [10–12]. The surviving clones always display severe transcriptional abnormalities [13–16]. The abnormalities that have been observed in cloned animals have raised doubt as to whether embryonic stem (ES) cells derived from SCNT embryos are identical to ES cells derived from early, normally fertilized embryos. If the NT-ES cells inherit these abnormalities from SCNT embryos, there may be risks associated with the therapeutic application of these cells [17–19].

The generation of autologous NT-ES cells has been successfully performed in mice [20] and more recently in non-human primates, namely rhesus monkeys [21]. These studies provided strong evidence that it may be possible for NT-ES cells to be derived from specific human individuals. More importantly, it has previously been shown that NT-ES cell lines are capable of differentiating into all three germ layers *in vitro* or into spermatozoa and oocytes in chimeric mice [3]. Furthermore, both NT- and fertilization-derived ES (F-ES) cell lines are functionally indistinguishable and can support the development of entire ES cell-derived mice after injection into tetraploid blastocysts [22–24]. However, the data for the developmental potency of NT-ES cells are not sufficient to address the significant safety concerns regarding gene expression and other abnormalities in cloned embryos. Therefore, it is important to complement the biological evidence with a molecular characterization of ES cells derived from fertilized and NT blastocysts. The transcriptional profiles of F-ES cells have already been investigated [25, 26]. The systematic comparison of NT-ES and F-ES cells showed that they have similar transcriptional profiles [27]. The NT-ES cells were identical to F-ES cells in terms of their expression of pluripotency markers, in the presence of tissue-dependent differentially DNA methylated regions, in DNA microarray profiles and in high-coverage gene expression profiling [28]. A systematic comparison of post-transcriptional regulation in NT-ES and F-ES cells has not yet been performed, however.

Proteins are the final products of mRNAs, but the expression profiles of mRNAs are not always equivalent to the final protein expression profiles. Translation from mRNAs to proteins is determined by post-transcriptional regulation. MicroRNAs have recently been proven to be important molecules involved in post-transcriptional regulation. In the present study, we investigated the microRNA and protein profiles of NT-ES and F-ES cells to address whether NT-ES cells are truly comparable to F-ES cells post-transcriptionally. By performing microRNA microarray and proteomics analyses, highly similar microRNA profiles and protein profiles in NT-ES and F-ES cells were observed. This result provides strong evidence to support the notion that NT-ES cells are post-transcriptionally indistinguishable from F-ES cells.

2 Materials and methods

2.1 ES cell derivation, pluripotency characterization and maintenance

The derivation of ES cell lines after nuclear transplantation was carried out as described previously by using testicular sertoli cells collected from B6D2F1 and B6129F1 mice respectively, as donor cells. The same strains of mice were used to collect fertilized embryos to derive the matching F-ES cell lines. All the established ES cell lines were cultured on mitomycin C-treated primary feeder fibroblasts in DMEM containing 15% FBS, 1000 U/mL leukemia inhibitory factor. To characterize the pluripotency of the established ES cell lines, a tetraploid blastocyst complementation experiment was performed. After characterization, two NT-ES cell lines, S5 (genetic background is B6D2F1) and S16 (genetic background is B6129F1), and three matching F-ES cell lines, CL11 (genetic background is B6D2F1), D1 (genetic background is B6129F1) and C8 (genetic background is B6129F1), were used for the following experiments.

To collect ES cells for microRNA microarray and protein analyses, three independent cultures of each cell line were assayed. After feeder cell depletion by pre-plating, ES cells were pelleted, snap frozen in liquid nitrogen and stored at $-80^{\circ}C$ until RNA and protein extraction.

2.2 RNA extraction and microRNA microarray analysis

To collect RNA for microRNA microarray and stem-loop real-time (RT)-PCR analyses, total RNA was extracted from cells using Trizol reagent (Invitrogen) as *per* the manufacturer's instructions. MicroRNA microarrays were performed using the miRCURY^{TM} LNA Array microRNA Profiling Service (Exiqon, Denmark). The total RNAs were labeled with Cy3 dye using miRCURY^{TM} LNA Array labeling kit (Exiqon) as *per* the manufacturer's instructions. The Cy3-labeled RNA molecules were hybridized with miRCURY^{TM} LNA Arrays (Exiqon), consisting of control probes, mismatch probes and more than 2000 capture probes, perfectly matched probes for human, mouse and rat microRNAs as registered and annotated in the miRBase release 9.2 at The Wellcome Trust Sanger Institute. Gene Pix 4000B scanner and GenePix Pro 6.0 software (Axon Instruments, Union city, CA) were used to scan images and for the analysis. Each group was hybridized with three miRCURY^{TM} LNA Arrays in triplicate with independent samples for ES cell lines. Each chip was normalized to the U6-2 signal intensity. The U6-2 signal intensity of each chip was adjusted to 13 000, which was approximately the median intensity of the U6-2. These normalized intensity values were then used to obtain the geometric means for each microRNA. The arithmetic means and standard errors for the two groups in triplicate were calculated, and the fold

changes were measured. Each microRNA signal was transformed to logarithm base 1.5, and a two-sample *t*-test was conducted. MicroRNAs with a significant value of 0.05 or lower and a fold change value of 1.5 or higher were listed and considered to be differentially expressed between NT- and F-ES cell lines. A statistical analysis was performed in MATLAB 7.5 (MathWorks, USA).

2.3 MicroRNA reverse transcriptase reactions and microRNA stem-loop RT-PCR

All primers and probes were designed based on microRNA sequences released by the Sanger Institute. The primer design and reverse transcriptase reactions were performed according to previously described methods [29]. RT PCR was performed using a standard SYBR Premix Ex Taq™_PCR kit protocol (Takara Bio) on the Mastercycler ep realplex RT PCR system (Eppendorf, http://www.eppendorf.com). The 20 μL PCR system contained 0.3 μL RT product, 1 × SYBR Premix Ex Taq™_PCR Mix, 0.2 μM forward primer and universal reverse primer. The reactions were incubated in a 96-well plate at 95°C for 2 min, followed by 40 cycles of 95°C for 15 s and 60°C for 40 s. All reactions were run in triplicate. The threshold cycle (C_T) was defined as the fractional cycle number at which the fluorescence passes the fixed threshold [29]. U6 was used as an internal control. The sequences of the reverse primer, forward primer and universal reverse primer are 5′-GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACAAAAATATG-3′, 5′-GCGCGTCGTGAAGCGTTC-3′ and 5′-GTGCAGGGTCCGAGGT-3′, respectively. The relative quantities (Q.rel) of each candidate miRNA gene were calculated from the cycle threshold (C_t) values scaled to a calibrator gene (U6) and corrected for efficiency of amplification (similar to 2) according to the formula $\text{Q.rel.} = 2^{-2\Delta C_t}$, where $\Delta C_t =$ average C_t test miRNA−average C_t U6. The errors were calculated as described previously [30].

2.4 Reagents for 2-D DIGE

IPG strips and IPG buffer were from GE Healthcare (Little Chalfont Bucks, UK); urea, CHAPS and DTT were from Bio-Rad (Hercules, CA, USA), and the complete proteinase inhibitor cocktail tablet was from Roche (Mannheim, Germany). Modified trypsin (sequencing grade) was obtained from Promega (WI, USA). Cy2, Cy3, Cy5, IPG strips and IPG buffer were from GE Healthcare. Thiourea was from Fluka (Buchs, Switzerland). All other chemical and biochemical reagents used were analytical grade.

2.5 Protein preparation

ES cells were ground into a fine powder in liquid nitrogen and homogenized in lysis buffer (7 M urea, 2 M thiourea, 4% CHAPS, 10 mM Tris, 5 mM magnesium acetate, one complete proteinase inhibitor cocktail tablet *per* 50 mL lysis buffer). For improved cell lysis, the solution was sonicated on ice for 1 min with a 1 s pulse ON and 1 s pulse OFF to prevent overheating. The samples were incubated for 30 min at room temperature with repeated vortexing and then centrifuged at 40 000 × *g* for 60 min at 20°C. The supernatant was stored in aliquots at −80°C. Protein concentration was determined with the Bradford assay kit (Bio-Rad), using albumin diluted in lysis buffer as a standard.

2.6 2-D DIGE

Lysates of embryonic stem cells were labeled with Cy2, Cy3 and Cy5 according to the protocols described in the Ettan DIGE User Manual (18-1164-40 Edition AA, GE Healthcare). The DIGE experimental design is shown in Supporting Information Table S1. Typically, 50 μg of lysates were labeled with 400 pmol of Cy3 or Cy5, and the same amount of the pool standard, which contained equal quantities of all the samples, was labeled with Cy2. To achieve statistical confidence, each sample was reloaded on three gels (Supporting Information Table S1). Labeling reactions were carried out on ice and in the dark for 30 min before being quenched with 1 μL of 10 mM lysine for 10 min on ice. These labeled samples were then combined for 2-D DIGE analysis. 2-D gel electrophoresis was performed as described earlier with some modifications [31]. IPG strips (24 cm, pH3-10 and NL) were rehydrated with labeled samples in the dark overnight with rehydrated buffer (8 M urea, 4% w/v CHAPS, 20 mM DTT, 1% v/v IPG buffer and trace amount of bromophenol blue). First-dimension IEF was performed using an Ettan IPGphor System (GE Healthcare) for a total of 87 kVh at 20°C. The strips were then treated with a two-step reduction and alkylation step prior to the second dimension (SDS-PAGE). After equilibration with a solution containing 6 M urea, 30% glycerol, 2% SDS, 50 mM Tris-HCl, pH 8.8 and 0.5% w/v DTT, the strips were treated with the same solution containing 4.5% w/v iodoacetamide instead of DTT. The strips were over-layered onto 12% polyacrylamide gels (20 × 24 cm), immobilized to a low-fluorescent glass plate and electrophoresed for ∼10 h at 30 mA *per* gel using an Ettan DALT Twelve System (GE Healthcare). The Cy2-, Cy3- and Cy5-labeled images were acquired on a Typhoon 9410 scanner (GE Healthcare).

2.7 Image analysis

DeCyder v.5.02 was used to analyze the DIGE images as described in the Ettan DIGE User Manual (GE Healthcare). Intra-gel spot detection and quantification and inter-gel matching and quantification were performed using differential in-gel analysis and biological variation analysis modules, respectively. Briefly, in differential in-gel analysis,

the Cy2, Cy3 and Cy5 images for each gel were merged, spot boundaries were automatically detected and normalized spot volumes (protein abundance) were calculated. The resulting spot maps were exported to biological variation analysis. The matching of the protein spots across eight gels was performed after several rounds of extensive land marking and automatic matching. Dividing each Cy3 or Cy5 spot volume by the corresponding Cy2 (internal standard) spot volume within each gel gave a standard abundance, thereby correcting for inter-gel variations. Each gel was firstly grouped into ''CL11'', ''S5'', ''C8'', ''S16'' and ''D1''. To test the significant differences in the expression of proteins between every two groups, a one-way ANOVA was performed using a statistical significance level of 0.05. The differentially expressed protein spots were filtered based on an average volume ratio of 1.5-fold with statistical significance ($p < 0.05$). Each protein spot signal was transformed to logarithm base 1.5. The KNN impute package in GenePattern was used to impute missing data if a protein spot had intensity values for at least half the samples.

2.8 Peptides extracted for identification

Separate preparative gels were run to obtain sufficient amounts of protein for MS analysis. These gels were fixed and stained with colloidal CBB. Protein spots of interest, as defined by the 2-D DIGE/DeCyder analysis, were excised from the colloidal CBB-stained gels using a modified in-gel tryptic digestion procedure. Gel pieces were first discolored in 50% ACN and 25 mM ammonium bicarbonate and then subjected to reduction and alkylation in 10 mM DTT and 55 mM iodoacetic acid, respectively. Following vacuum drying, the gel pieces were incubated with sequencing grade modified trypsin (Promega) at a final concentration of 0.01 μg/μL in 25 mM ammonium bicarbonate for 16 h at 37°C. Supernatants were collected, vacuum-dried and redissolved in 50% ACN and 0.1% TFA for MS analysis.

2.9 MALDI TOF/TOF analysis

Peptides were mixed with a MALDI matrix (7 mg/mL CHCA, 0.1% TFA and 50% ACN) and spotted on to the 192-well stainless steel MALDI target plates. Samples on the MALDI target plates were then analyzed using an ABI 4800 Proteomics Analyzer MALDI TOF/TOF mass spectrometer (Applied Biosystems). For MS analyses, typically 1000 shots were accumulated for each spot, while for MS/MS analysis, 2500 shots were accumulated. MS/MS analyses were performed using air, at a collision energy of 1 kV and a collision gas pressure of 2.0×10^{-8}–3.0×10^{-7} Torr. The MS together with MS/MS spectra were searched against the International Protein Index mouse database version 3.23 (http://www.ebi.ac.uk/IP/IPIhelp.html) using the software GPS Explorer™ Version 3.0 and MASCOT database search algorithms (version 2.0). Cysteine carbamidomethylation and methionine oxidation were selected as variable modifications. One missing cleavage was allowed. Precursor error tolerance was set to < 0.1 Da and MS/MS fragment error tolerance was set to < 0.25 Da. All the proteins identified should have protein scores greater than 59 ($p < 0.05$) and individual ion scores greater than 21 with expected values < 0.05. All the MS/MS spectra were further validated manually.

2.10 Data processing and statistical analysis

An unsupervised hierarchical data set clustering was performed using the parameters that the distance was a Euclidean distance and the linkage was the average. PCA was also performed to evaluate the similarity between different samples. All of the above analyses were performed with the Bioinfomatics toolbox 2.1/MATLAB 7.5 (MathWorks). The analysis of miRNA predicted targets was determined using the algorithms TargetScan v4.2 (http://www.targetscan.org/) and miRanda v5 (http://microrna.sanger.ac.uk/targets/).

3 Results

3.1 Developmental potency of ES cell lines

Tetraploid blastocyst complementation, the most stringent test for ES cell pluripotency, was performed to evaluate the pluripotency of the NT-ES cell lines and the matching F-ES cell lines derived from B6D2F1 and B6129F1 blastocysts, respectively. The verification of pluoripotency in S5, CL11, D1 and S16 ES cell lines has been reported in our previously published papers [32, 33]. The other ES cell line analyzed in this study, C8, generated live pups after tetraploid blastocyst complementation (Supporting Information Fig. S1). The rate of establishment of the ES cell lines was presented in Supporting Information Table S2. As summarized in Supporting Information Table S3, two NT-ES cell lines, S5 and S16, and three matching F-ES cell lines, CL11, C8 and D1, were selected for further analysis of the microRNA and protein expression profiles.

3.2 The microRNA profiles from NT- and F-ES cell lines are highly similar

Although similar transcriptional profiles have been observed for both NT- and F-ES cells [27, 28], we posited that aberrant epigenetic reprogramming might cause an incomplete reprogramming of microRNA expression in NT-ES cells; as a result, there could be aberrant post-transcriptional regulation and protein expression profiles that differed from F-ES cells. Therefore, microRNA microarray and stem-loop

RT-PCR technology were performed to compare the expression levels of 417 mouse microRNAs in the ES cell lines derived from NT and fertilized blastocysts.

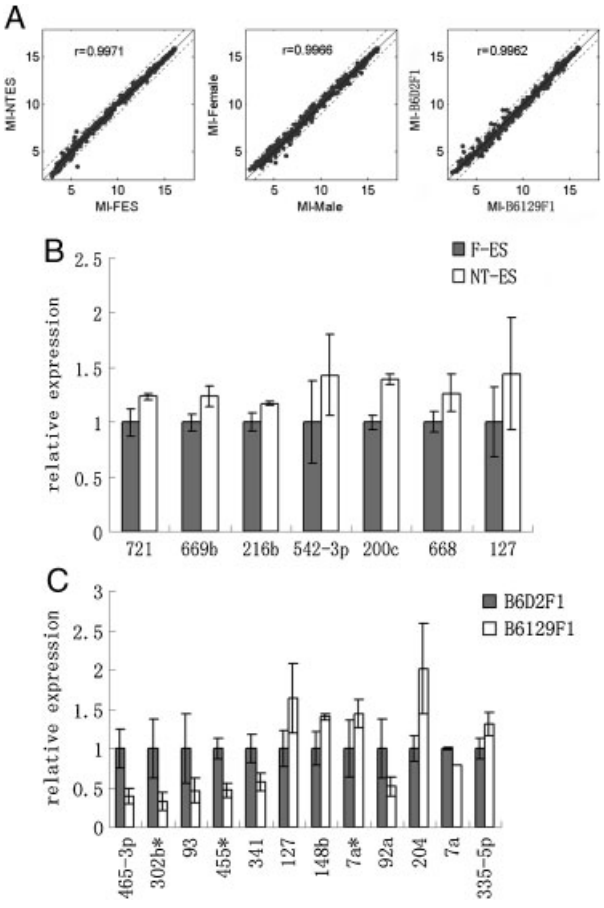

Figure 1. Analysis of miRNA expression profiles between F-ES/NT-ES, male-ES/female-ES and B6129F1-ES/B6D2F1-ES cell lines. (A) Analysis of microRNA expression profiles. Mean signal intensities (MI) of two NT-ESC lines were plotted against the corresponding MI of two F-ESC lines. Dotted lines indicate 1.5-fold regulation. (B) The differentially expressed microRNAs between NT- and F-ES cell lines (*t*-test, $p<0.1$). (C) The differentially expressed microRNAs between B6D2F1 and B6129F1 blastocysts-derived ES cell lines (*t*-test, $p<0.05$).

To assess whether systematic mis-expression of microRNA genes could be observed in NT-ES cells, the mean probe signal levels of all NT-ES cell lines were compared with the corresponding mean signal values of all matching F-ES cell lines (Fig. 1A). A high degree of similarity in the microRNA profiles was observed between NT- and matching F-ES cells (Pearson's coefficient of correlation: $r = 0.9971$). More importantly, no transcripts were found to be significantly deregulated in NT-ES cells, with mean signal changes of >1.5-fold for only two probes that displayed Student's *t*-test *p* values <0.05. Even when the student's *t*-test *p* value was changed from <0.05 to <0.1, no microRNAs had mean signal changes of >1.5-fold for all seven probes (Table 1 and Fig. 1B).

Stem-loop RT-PCR was used to detect the specific microRNA expression levels in the NT-ES, F-ES and MEF cells. In ES cells, miR-290 and miR-291-5p are highly expressed microRNAs, and miR-503 is a lowly expressed microRNA (data not shown). The data show that these microRNAs are expressed at significantly different levels in ES cells and MEF cells, but at similar levels in NT- and F-ES cells (Supporting Information Fig. S2).

These results indicate that there is no significant difference in the microRNA expression profiles of NT- and F-ES cell lines.

3.3 The differences in microRNA profiles are due to genetic background rather than due to the type of donor blastocyst

As shown in Fig. 1, the Pearson's coefficient of correlation between NT- and F-ES cells was 0.9971; between B6D2F1 and B6129F1 blastocyst-derived ES cell lines, it was 0.9962; between female and male blastocyst-derived ES cell lines, it was 0.9966. The highest Pearson's coefficient of correlation could be observed between NT- and F-ES cells, indicating that NT-ES cells are highly similar to F-ES cells in their global microRNA expression levels.

More importantly, no significantly deregulated microRNA transcripts were found in NT-ES cells with mean signal changes of >1.5-fold and Student's *t*-test *p* values

Table 1. The differentially expressed miRNAs between NT- and F-ES cell lines (*t*-test, $p<0.1$)

Name	Fold changes (F/NT)	*t*-Test *p* value	S5	CL11	D1	S16
mmu-miR-721	0.8098	0.0098	2347.203	1704.149	2034.959	2270.345
mmu-miR-669b	0.8088	0.0157	7209.665	5825.142	6483.761	8008.184
mmu-miR-216b	0.8529	0.0515	80.33384	65.06078	73.36755	81.97248
mmu-miR-542-3p	0.7000	0.0566	5551.53	2411.669	4155.857	3830.651
mmu-miR-200c	0.7184	0.0633	116.8028	89.69807	82.00305	122.2077
mmu-miR-688	0.7893	0.0665	6631.1	4472.141	5090.72	5484.082
mmu-miR-127	0.6927	0.0954	115.4606	82.63884	131.2803	193.3427

The values in the column: fold changes are the mean signal intensities (MSI) of two F-ES cell lines divided by MSI of two NT-ES cell lines for each miRNA analyzed. The values in columns S5, CL11, D1 and S16 indicate the standard signal intensities of different miRNAs detected by the miRNA microarray.

Table 2. The differentially expressed microRNAs between B6D2F1 and B6129F1-derived ES cell lines (*t*-test, $p<0.05$)

Name	Fold changes (B6D2F1/B6129F1)	*t*-Test *p* value	S5	CL11	D1	S16
mmu-miR-465a-3p/465b-3p/465c-3p	2.5247	9.09501E-05	286.5136	201.9918	114.1582	79.33228
mmu-miR-302b*	2.9981	0.0002	866.354	506.5487	283.7315	174.1927
mmu-miR-93	2.1750	0.0089	874.8247	457.5499	379.3302	233.2692
mmu-miR-455*	2.1392	0.0107	83.52247	70.15356	40.88771	30.95204
mmu-miR-341	1.7397	0.0142	23956.25	18708.93	10456.74	14067.05
mmu-miR-127	0.6102	0.0151	115.4606	82.63884	131.2803	193.3427
mmu-miR-148b	0.7110	0.0151	393.2497	291.4012	491.0211	471.9137
mmu-miR-7a*	0.6917	0.0194	133.8031	79.67145	167.5633	141.0635
mmu-miR-92a	1.9350	0.0251	304.6556	176.1988	146.0106	102.4938
mmu-miR-204	0.4970	0.0253	95.16596	75.45994	137.0125	206.3109
mmu-miR-7a	1.2735	0.0404	24419.6	25199.93	19415.59	19546.04
mmu-miR-335-5p	0.7623	0.0472	963.4336	797.9478	1065.141	1245.579

The values in the column: fold changes are the mean signal intensities (MSI) of two B6D2F1-ES cell lines divided by MSI of two B6129F1-ES cell lines for each miRNA analyzed. The values in columns S5, CL11, D1 and S16 indicate the standard signal intensities of different miRNAs detected by the miRNA microarray.

<0.05 (Table 1 and Fig. 1B). In striking contrast, our examination revealed eight significantly deregulated microRNA transcripts in B6D2F1-ES cells compared with B6129F1-ES cells, with mean signal changes of >1.5-fold for all 12 probes and Student's *t*-test *p* values <0.05 (Table 2 and Fig. 1C).

These results demonstrate that the most prominent differences in microRNA expression profiles between the ES cell lines that we analyzed can be attributed to genetic background rather than to the derivation of the respective ES cell line, from an NT or a fertilization-derived blastocyst.

3.4 Variability in microRNA profiles is similar in NT-ES and F-ES cells

To investigate whether the variability in the microRNA expression was elevated in NT- *versus* F-ES cells, the expression levels for a subset of microRNAs were measured. miR-203 promotes epidermal differentiation by repressing p63 expression during skin development [34]. In the immune system, miR-155 plays a key role in regulating B-cell and T-cell function [35]. The brain-specific miRNA, miR-9, plays an important role in neural lineage differentiation of ES cells [36]. The muscle-specific microRNA, miR-133, promotes mesoderm formation from ES cells [37], and miR-291-5p, miR-292-3p, miR-294 and miR-295 are ES cell-specific microRNAs [38]. These microRNAs are important for the differentiation and self-renewal of ES cell lines. As shown in Fig. 2A, their expression levels were not significantly different in NT- and F-ES cells.

The standard deviations of all probe signal levels obtained in the NT- and F-ES cell groups were sorted and plotted by

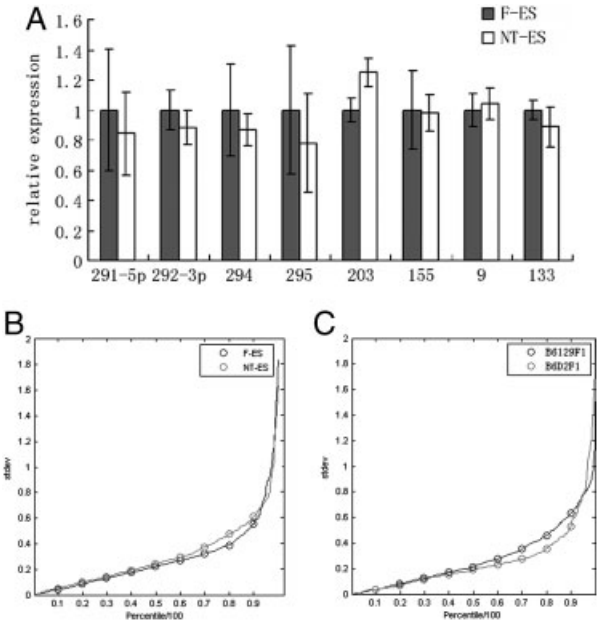

Figure 2. Analysis of variability in microRNA genes expression among different ES cell lines. (A) Comparison of gene expression levels in NT-ES (open bars) and F-ES (filled bars) cell lines. Columns display mean signal intensities; error bars display standard deviation. (B and C) Comparison of standard deviation levels across all probes in the data set. As a measure of miRNA gene expression variability, standard deviation levels were calculated for the log 2 probe signal values for each group. Probes in each group (NT-ES and F-ES cell lines) were ordered by their standard deviation levels, and then standard deviation levels were compared at different percentiles. Data sets display closely matched standard deviation levels for different percentiles, indicating highly similar variability in microRNA gene expression between the two groups.

calculating the standard deviation values for different percentiles of these sorted data sets. These plots show the global microRNA gene expression variability. The standard deviation levels between the data sets of NT- and F-ES cell lines are highly similar in the resulting diagram (Fig. 2B). The standard deviation values obtained from the 70th to the 90th percentile are slightly more variable in microRNA gene expression in the NT-ES cell lines, but no increase in overall variability was found. In contrast to the highly similar variability between NT- and F-ES cell lines, the standard deviation values from the 40th to the 100th percentile are significantly different between B6D2F1- and B6129F1-ES cell lines (Fig. 2C).

These data indicate that the individual and overall variability in microRNA gene expression among cell lines of the same genetic background is highly similar.

3.5 The protein profiles from NT- and F-ES cell lines are highly similar

Our microRNA expression data combined with previous results demonstrate that there are marked similarities between NT- and F-ES cells on the transcriptional and post-transcriptional levels [27, 28]. Further investigation was

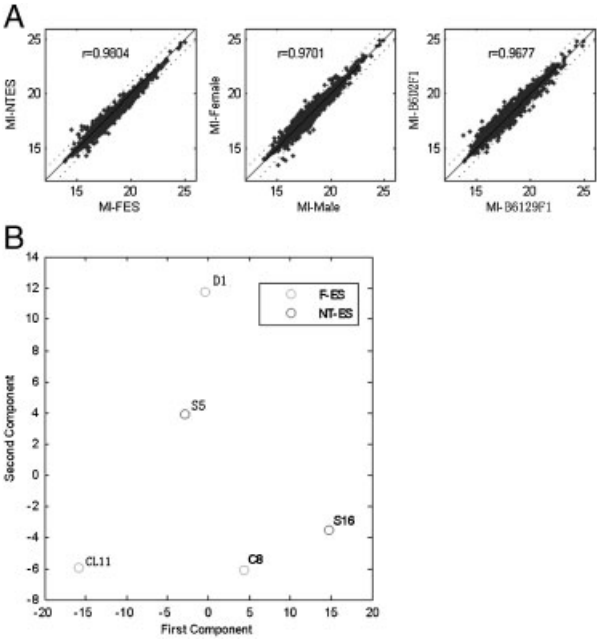

Figure 3. Analysis of expression profiles of all protein spots between F-ES/NT-ES, male-ES/female-ES and B6129F1-ES/B6D2F1-ES cell lines. (A) Pearson's coefficient of correlation analysis. (B) PCA.

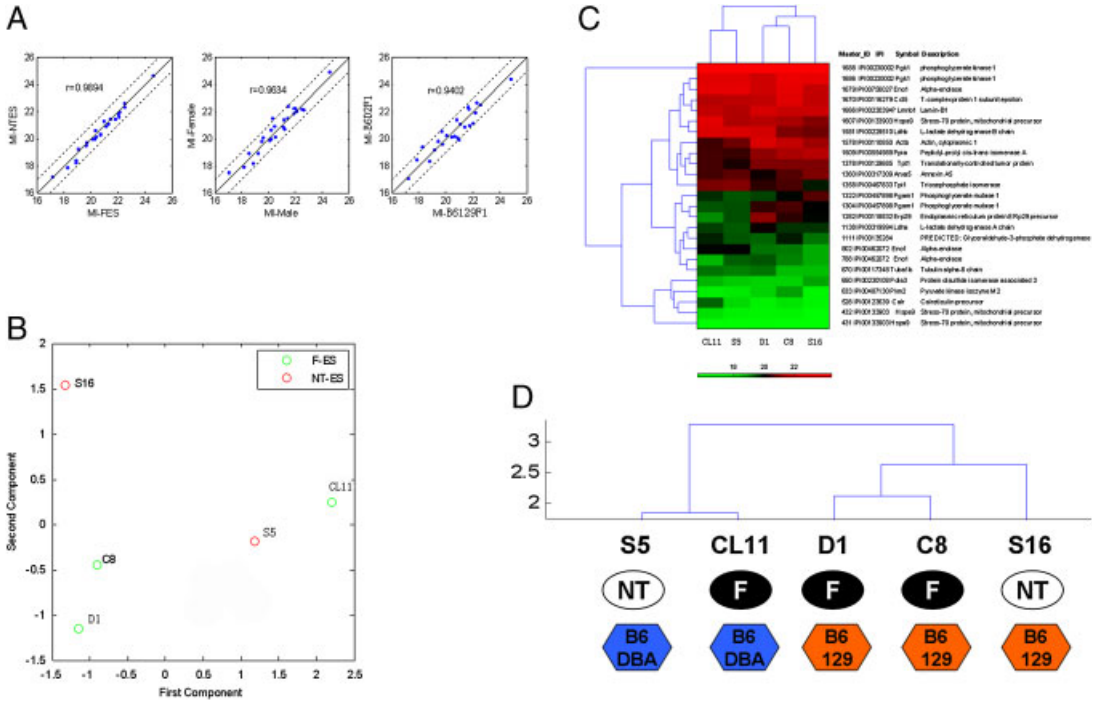

Figure 4. Analysis of expression profiles of the characterized proteins between F-ES/NT-ES, Male-ES/Female-ES and B6129F1-ES/B6D2F1-ES cell lines. (A) Pearson's coefficient of correlation analysis. (B) PCA. (C) Heat map of clustering results (blue, no or very low expression; black, low expression; red, high expression). (D) Sample tree obtained from hierarchical clustering. ES cell line expression profiles cluster by genetic background rather than by the type of donor blastocyst.

required to assess whether the protein expression profiles were comparable between NT- and F-ES cells. Thus, 2-D DIGE, MALDI-TOF/TOF and bioinformatic analyses were performed to determine whether NT- and F-ES cells are truly comparable.

On average, 1691 protein spots were detected in each of the eight three-color gels. A representative gel image was shown in Supporting Information Fig. S3. We identified 87 protein spots that were differentially expressed among all ES cell lines; these proteins were picked for mass spectrometry (Supporting Information Fig. S3). Twenty-five proteins were identified from the 87 differentially expressed protein spots by MALDI-TOF/TOF; these proteins originated from 19 unique genes (Supporting Information Table S4).

As shown in Fig. 3A, a high degree of similarity was observed in the protein profiles of NT- and F-ES cells based on the total protein spots detected (Pearson's coefficient of correlation: $r = 0.9804$). Only five protein spots were found to be differentially expressed between NT-ES cell line S5 and F-ES cell line CL11, and 20 protein spots were found to be differentially expressed between NT-ES cell line S16 and F-ES cell line C8, with Student's *t*-test p values <0.05 and mean signal changes of >1.5-fold (Supporting Information Table S5). Among these 25 protein spots, only one differentially expressed protein in NT- and F-ES cells, Ldhb (L-lactate dehydrogenase B chain), was identified by MALDI-TOF/TOF (Supporting Information Fig. S3A).

Based on the 25 differentially expressed proteins characterized by MALDI-TOF/TOF, the Pearson's coefficient of correlation between NT- and F-ES cells was found to be 0.9894. There was only one significantly deregulated protein, out of 25 protein spots, with a Student's *t*-test p value <0.05 and mean signal change of >1.5-fold (Fig. 4A and Supporting Information Fig. S3A).

These results suggest that the protein profiles of NT- and F-ES cell lines are highly similar.

3.6 The differences in protein profiles are due to genetic background rather than the type of donor blastocyst

For all 1691 protein spots in the 2-D DIGE, the Pearson's coefficient of correlation between NT- and F-ES cells was 0.9804; between B6D2F1- and B6129F1- ES cell lines, it was 0.9677; and between female and male blastocyst-derived ES cell lines, it was 0.9701 (Fig. 3A). For all 25 differentially expressed proteins characterized by MALDI-TOF/TOF, the Pearson's coefficient of correlation between NT- and F-ES cells was 0.9894; between B6D2F1 and B6129F1 blastocyst-derived ES cell lines, it was 0.9402; and between female and male blastocyst-derived ES cell lines, it was 0.9634 (Fig. 4A). The Pearson's coefficient of correlation between NT- and the matching F-ES cells was the highest, and that between B6D2F1 and B6129F1-ES cells was the lowest, proving that the NT-ES cells and the matching F-ES cells are highly similar.

A total of 25 protein spots were found to be differentially expressed between NT- and F-ES cell lines, with Student's *t*-test p values <0.05 and mean signal changes of >1.5-fold. However, more differentially expressed protein spots were found between B6D2F1- and B6129F1-ES cells. For example, 60 differentially expressed protein spots were found in F-ES cell lines CL11 and C8, and 41 in NT-ES cell lines S5 and S16 with different genetic backgrounds (Supporting Information Table S5). Furthermore, only one differentially expressed protein in NT- and F-ES cell lines was characterized by MALDI-TOF/TOF, whereas 18 differentially expressed proteins were identified by MALDI-TOF/TOF in B6D2F1 and B6129F1 ES cell lines (Fig. 4A and Supporting Information Figs. S3B and S3C). More differentially expressed protein spots could be observed in the ES cell lines with different genetic backgrounds than in those with the same genetic background.

PCA is a vector space transform that is often used to reduce the number of dimensions in a multidimensional data set and the similar samples will be close in the distance [39]. For all of the protein spots in the 2-D DIGE, the ES cells with the same genetic background, F-ES cell line C8 and NT-ES cell line S16 were the closest in the PCA (Fig. 3B). For the characterized protein spots in the 2-D DIGE, S5/CL11 and C8/D1 were the closest in the PCA (Fig. 4B). Given all of the characterized protein spots, the closest ES cell pair was that derived from blastocysts of the same genetic background. The difference in the ES cells with different genetic backgrounds was more conspicuous than in the NT- and F-ES cells with the same genetic background.

Unsupervised hierarchical data set clustering was performed to assess the differences and similarities in the single ES cell line protein expression profiles in an unbiased way (Figs. 4C and D). The 25 differentially expressed proteins characterized by MALDI-TOF/TOF were analyzed with this assay. In the resulting sample tree, there was no separation of NT- and F-ES cell lines or direct clustering of the two NT-ES cell lines (Fig. 4C). The NT-ES cell lines were clustered with the F-ES cell lines with the same genetic background; *i.e.* S5 clustered with CL11 and S16 with C8 and D1. These results show that the NT-ES cell lines are more similar in their protein profiles to the F-ES cell lines with the same genetic background than to other NT-ES cell lines with the different genetic background. The genetic background annotation revealed that there was a separation of different genetic backgrounds between clusters, with the S5/CL11 cluster including the only two B6D2F1 lines and the C8/D1/S16 cluster containing all three B6129F1 ES cell lines (Fig. 4D).

These results suggest that the most prominent differences in the protein expression profiles of the ES cell lines that we analyzed can be attributed to genetic background rather than to the derivation of the respective ES cell line from an NT or a fertilization-derived blastocyst.

4 Discussion

In the present study, we compared the microRNA and protein expression profiles of NT-ES and matching F-ES cell lines by microRNA microarray, 2-D DIGE, MALDI-TOF/TOF and related bioinformatics analyses. Our results show that there are no marked differences in the microRNA and protein expression profiles of NT- and F-ES cell lines; these data support the notion that the NT-ES cell lines are post-transcriptionally indistinguishable from the F-ES cell lines. Our results provide strong evidence that patient-specific NT-ES cells can be applied for therapeutic purposes if such cell lines could be successfully and efficiently generated.

The microRNA microarray has emerged as a powerful tool for investigating the global expression of microRNAs among different organisms, and differentially expressed microRNAs can be detected for further functional characterization. Several important microRNAs have been identified using this approach. In a recent study, miR-21 was shown to be differentially expressed in the uterus between the implantation site and the inter-implantation site; it was found to play an important role in embryo implantation [40]. Here, we used a microRNA microarray containing 417 mouse microRNAs to distinguish whether microRNAs are differentially expressed in SCNT-derived ES cells and normal F-ES cells. Our microRNA microarray data and RT-PCR results indicate that the microRNA expression profiles of NT-ES cells and F-ES cells with the same genetic background are indistinguishable. We did detect differentially expressed microRNAs in ES cell lines with different genetic backgrounds, however, regardless of whether these cells were derived from normally fertilized embryos or SCNT-produced embryos.

2-D DIGE and MS/MS have been widely used to identify the differentially expressed proteins in different samples. We found it important to investigate whether the NT-ES cell lines are truly comparable to F-ES lines in their protein expression profiles. By performing the 2-D DIGE and MS/MS analyses, we found that only one protein, Ldha, is differentially expressed in NT- and F-ES cells. In striking contrast, 18 characterized proteins were found to be differentially expressed in ES cells of different genetic backgrounds. Moreover, most of their encoded mRNAs were computationally predicted to be targets of the differentially expressed microRNAs (Supporting Information Table S6). Therefore, our data show that NT-ES cells are highly similar to F-ES cell lines with the same genetic background based on a post-transcriptional analysis.

Many lines of evidence support the notion that ES cell derivation is a rigorously selective process for the rare cells that are able to proliferate under tissue culture conditions. The efficiency of deriving NT-ES cell lines is significantly lower than that of F-ES cell lines (Supporting Information Table S2). Nonetheless, the established NT- and F-ES cell lines are highly similar, both transcriptionally as demonstrated previously and post-transcriptionally as demonstrated in the present study. These results indicate that cellular differentiation influences the epigenetic state of the donor cell nucleus, which in turn determines the efficiency with which an enucleated oocyte can reprogram a donor cell into a pluripotent embryonic stem cell fate [41]. Secondly, only a small fraction of the explanted inner cell mass cells maintain Oct4 expression and proliferate to generate immortal cell lines; these cells are designated ''embryonic stem cells'' [42]. This selective process possibly leads to the loss of the ''epigenetic memory'' of the donor nucleus.

An extensive analysis of chimeric animals over the past two decades has not revealed any obvious defects or tumor-forming potential for ES cell-derived somatic cells. Many therapeutic applications of NT-ES cells have been established in mice [5, 20]. Recently, it has been shown that cloned blastocysts can be successfully generated by transferring differentiated somatic cell nuclei into enucleated human oocyte. These attempts may accelerate the attainment of NT-ES cells from specific individuals [43].

Induced pluripotent stem (iPS) cells have been generated recently by introducing several transcriptional factors into somatic cells; however, the derivation of NT-ES cells remains an attractive alternative way to generate patient-specific pluripotent stem cells [44]. Although recent studies have shown that insertional mutagenesis is not required for *in vitro* reprogramming and that iPS cells can be generated without viral integration, the variation in pluripotency of the iPS cells still needs to be stringently evaluated. To date, not even one iPS cell line can generate live pups after tetraploid blastocyst complementation; they exhibit a different developmental potential than the NT-ES cell lines that have been reported [45–47].

In conclusion, our data indicate that NT-ES cells are highly similar to F-ES cells post-transcriptionally, and the application of NT-ES cells to regenerative medicine is practically feasible.

The study was supported by National High Technology Project 863 (2005AA210930 and 2008AA022311) to S.G. We thank the lab members for helpful comments on the paper and Shanghai SensiChip Tech&Infor Company for the assistance with bioinformatics.

The authors have declared no conflict of interest.

5 References

[1] Wilmut, I., Schnieke, A. E., McWhir, J., Kind, A. J., Campbell, K. H., Viable offspring derived from fetal and adult mammalian cells. *Nature* 1997, *385*, 810–813.

[2] Munsie, M. J., Michalska, A. E., O'Brien, C. M., Trounson, A. O. *et al.*, Isolation of pluripotent embryonic stem cells from reprogrammed adult mouse somatic cell nuclei. *Curr. Biol.* 2000, *10*, 989–992.

[3] Wakayama, T., Tabar, V., Rodriguez, I., Perry, A. C. *et al.*, Differentiation of embryonic stem cell lines generated from adult somatic cells by nuclear transfer. *Science* 2001, *292*, 740–743.

[4] Hochedlinger, K., Jaenisch, R., Nuclear transplantation, embryonic stem cells, and the potential for cell therapy. *N. Engl. J. Med.* 2003, *349*, 275–286.

[5] Rideout, W. M., III, Hochedlinger, K., Kyba, M., Daley, G. Q., Jaenisch, R., Correction of a genetic defect by nuclear transplantation and combined cell and gene therapy. *Cell* 2002, *109*, 17–27.

[6] Karra, R., Wu, S. M., Multipotent stem cells in cardiac regenerative therapy. *Regen. Med.* 2008, *3*, 189–198.

[7] Tamashiro, K. L., Wakayama, T., Akutsu, H., Yamazaki, Y. *et al.,* Cloned mice have an obese phenotype not transmitted to their offspring. *Nat. Med.* 2002, *8*, 262–267.

[8] Ogonuki, N., Inoue, K., Yamamoto, Y., Noguchi, Y. *et al.*, Early death of mice cloned from somatic cells. *Nat. Genet.* 2002, *30*, 253–254.

[9] Inoue, K., Kohda, T., Lee, J., Ogonuki, N. *et al.*, Faithful expression of imprinted genes in cloned mice. *Science* 2002, *295*, 297.

[10] Ohgane, J., Wakayama, T., Kogo, Y., Senda, S. *et al.*, DNA methylation variation in cloned mice. *Genesis* 2001, *30*, 45–50.

[11] Senda, S., Wakayama, T., Yamazaki, Y., Ohgane, J. *et al.*, Skewed X-inactivation in cloned mice. *Biochem. Biophys. Res. Commun.* 2004, *321*, 38–44.

[12] Humpherys, D., Eggan, K., Akutsu, H., Hochedlinger, K. *et al.*, Epigenetic instability in ES cells and cloned mice. *Science* 2001, *293*, 95–97.

[13] Ng, R. K., Gurdon, J. B., Epigenetic memory of active gene transcription is inherited through somatic cell nuclear transfer. *Proc. Natl. Acad. Sci. USA* 2005, *102*, 1957–1962.

[14] Yanagimachi, R., Cloning: experience from the mouse and other animals. *Mol. Cell. Endocrinol.* 2002, *187*, 241–248.

[15] Rhind, S. M., Taylor, J. E., De Sousa, P. A., King, T. J. *et al.*, Human cloning: can it be made safe? *Nat. Rev. Genet.* 2003, *4*, 855–864.

[16] Ogura, A., Inoue, K., Ogonuki, N., Lee, J. *et al.,* Phenotypic effects of somatic cell cloning in the mouse. *Cloning Stem Cells* 2002, *4*, 397–405.

[17] Kohda, T., Inoue, K., Ogonuki, N., Miki, H. *et al.*, Variation in gene expression and aberrantly regulated chromosome regions in cloned mice. *Biol. Reprod.* 2005, *73*, 1302–1311.

[18] Fulka, J., Jr., Miyashita, N., Nagai, T., Ogura, A., Do cloned mammals skip a reprogramming step? *Nat. Biotechnol.* 2004, *22*, 25–26.

[19] Armstrong, L., Lako, M., Dean, W., Stojkovic, M., Epigenetic modification is central to genome reprogramming in somatic cell nuclear transfer. *Stem Cells* 2006, *24*, 805–814.

[20] Tabar, V., Tomishima, M., Panagiotakos, G., Wakayama, S. *et al.*, Therapeutic cloning in individual parkinsonian mice. *Nat. Med.* 2008, *14*, 379–381.

[21] Byrne, J. A., Pedersen, D. A., Clepper, L. L., Nelson, M. *et al.*, Producing primate embryonic stem cells by somatic cell nuclear transfer. *Nature* 2007, *450*, 497–502.

[22] Rideout, W. M., III Eggan, K., Jaenisch, R., Nuclear cloning and epigenetic reprogramming of the genome. *Science* 2001, *293*, 1093–1098.

[23] Eggan, K., Akutsu, H., Loring, J., Jackson-Grusby, L. *et al.*, Hybrid vigor, fetal overgrowth, and viability of mice derived by nuclear cloning and tetraploid embryo complementation. *Proc. Natl. Acad. Sci. USA* 2001, *98*, 6209–6214.

[24] Li, J., Ishii, T., Feinstein, P., Mombaerts, P., Odorant receptor gene choice is reset by nuclear transfer from mouse olfactory sensory neurons. *Nature* 2004, *428*, 393–399.

[25] Ramalho-Santos, M., Yoon, S., Matsuzaki, Y., Mulligan, R. C. ''Stemness'': transcriptional profiling of embryonic and adult stem cells. *Science* 2002, *298*, 597–600.

[26] Ivanova, N. B., Dimos, J. T., Schaniel, C., Hackney, J. A. *et al.*, A stem cell molecular signature. *Science* 2002, *298*, 601–604.

[27] Brambrink, T., Hochedlinger, K., Bell, G., Jaenisch, R., ES cells derived from cloned and fertilized blastocysts are transcriptionally and functionally indistinguishable. *Proc. Natl. Acad. Sci. USA* 2006, *103*, 933–938.

[28] Wakayama, S., Jakt, M. L., Suzuki, M., Araki, R. *et al.*, Equivalency of nuclear transfer-derived embryonic stem cells to those derived from fertilized mouse blastocysts. *Stem Cells* 2006, *24*, 2023–2033.

[29] Chen, C., Ridzon, D. A., Broomer, A. J., Zhou, Z. *et al.*, Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic. Acids. Res.* 2005, *33*, e179.

[30] Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B. *et al.,* Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome. Biol.* 2002, *3*, RESEARCH0034.

[31] Liang, C. R., Leow, C. K., Neo, J. C., Tan, G. S. *et al.*, Proteome analysis of human hepatocellular carcinoma tissues by two-dimensional difference gel electrophoresis and mass spectrometry. *Proteomics* 2005, *5*, 2258–2271.

[32] Chang, G., Liu, S., Wang, F., Zhang, Y. *et al.*, Differential methylation status of imprinted genes in nuclear transfer derived ES (NT-ES) cells. *Genomics* 2009, *93*, 112–119.

[33] Lu, S., Liu, S., He, W., Duan, C. *et al.*, Bioreactor cultivation enhances NTEB formation and differentiation of NTES cells into cardiomyocytes. *Cloning Stem Cells* 2008, *10*, 363–370.

[34] Yi, R., Poy, M. N., Stoffel, M., Fuchs, E., A skin microRNA promotes differentiation by repressing 'stemness'. *Nature* 2008, *452*, 225–229.

[35] Rodriguez, A., Vigorito, E., Clare, S., Warren, M. V. *et al.*, Requirement of bic/microRNA-155 for normal immune function. *Science* 2007, *316*, 608–611.

[36] Krichevsky, A. M., Sonntag, K. C., Isacson, O., Kosik, K. S., Specific microRNAs modulate embryonic stem cell-derived neurogenesis. *Stem Cells* 2006, *24*, 857–864.

[37] Ivey, K. N., Muth, A., Arnold, J., King, F. W., Yeh, R. F. *et al.*, MicroRNA regulation of cell lineages in mouse and human embryonic stem cells. *Cell Stem Cell* 2008, *2*, 219–229.

[38] Houbaviy, H. B., Murray, M. F., Sharp, P. A., Embryonic stem cell-specific MicroRNAs. *Dev. Cell* 2003, *5*, 351–358.

[39] Raychaudhuri, S., Stuart, J. M., Altman, R. B., Principal components analysis to summarize microarray experiments: application to sporulation time series, in: Altman, R. B., Dunleer, A. K., Hunter, L., Lauderla, K., Klein, T.E.D. (Eds.) *Pacific Symposium on Biocomputing 2000, World Scientific, New Jersey* 2000, 455–466.

[40] Hu, S. J., Ren, G., Liu, J. L., Zhao, Z. A. *et al.*, MicroRNA expression and regulation in mouse uterus during embryo implantation. *J. Biol. Chem.* 2008, *283*, 23473–23484.

[41] Blelloch, R., Wang, Z., Meissner, A., Pollard, S. *et al.,* Reprogramming efficiency following somatic cell nuclear transfer is influenced by the differentiation and methylation state of the donor nucleus. *Stem Cells* 2006, *24*, 2007–2013.

[42] Buehr, M., Nichols, J., Stenhouse, F., Mountford, P. *et al.*, Rapid loss of Oct-4 and pluripotency in cultured rodent blastocysts and derivative cell lines. *Biol. Reprod.* 2003, *68*, 222–229.

[43] French, A. J., Adams, C. A., Anderson, L. S., Kitchen, J. R. *et al.,* Development of human cloned blastocysts following somatic cell nuclear transfer with adult fibroblasts. *Stem Cells* 2008, *26*, 485–493.

[44] Takahashi, K., Yamanaka, S., Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 2006, *126*, 663–676.

[45] Shi, Y., Desponts, C., Do, J. T., Hahm, H. S. *et al.,* Induction of pluripotent stem cells from mouse embryonic fibroblasts by Oct4 and Klf4 with small-molecule compounds. *Cell Stem Cell* 2008, *3*, 568–574.

[46] Stadtfeld, M., Nagaya, M., Utikal, J., Weir, G. *et al.,*Induced pluripotent stem cells generated without viral integration. *Science* 2008, *322*, 945–949.

[47] Okita, K., Nakagawa, M., Hyenjong, H., Ichisaka, T., Yamanaka, S., Generation of Mouse induced pluripotent stem cells without viral vectors. *Science* 2008, *322*, 949–953.

