STEM CELLS AND DEVELOPMENT
Volume 21, Number 8, 2012

DOI: 10.1089/scd.2011.0243

Comparative Proteomic Analysis of Human Somatic Cells, Induced Pluripotent Stem Cells, and Embryonic Stem Cells

Sun Young Kim,[1,2] Min-Jeong Kim,[3] Hyeyun Jung,[1] Won Kon Kim,[1] Sang Oh Kwon,[4] Myung Jin Son,[3] Ik-Soon Jang,[4] Jong-Soon Choi,[4] Sung Goo Park,[1] Byoung Chul Park,[1] Yong-Mahn Han,[2] Sang Chul Lee,[1] Yee Sook Cho,[3] and Kwang-Hee Bae[1]

Induced pluripotent stem cells (iPSCs) are somatic cells that have been reprogrammed to a pluripotent state via introduction of defined transcription factors. iPSCs are a valuable resource for regenerative medicine, but whether iPSCs are identical to embryonic stem cells (ESCs) remains unclear. In this study, we performed comparative proteomic analyses of human somatic cells [human newborn foreskin fibroblasts (hFFs)], human iPSCs (hiPSCs) derived from hFFs, and H9 human ESCs (hESCs). We reprogrammed hFFs to a pluripotent state using 4 core transcription factors: Oct4 (O), Sox2 (S), Klf4 (K), and c-Myc (M). The proteome of hiPSCs induced by 4 core transcription factors was relatively similar to that of hESCs. However, several proteins, including dUTPase, GAPDH, and FUSE binding protein 3, were differentially expressed between hESCs and hiPSCs, implying that hiPSCs are not identical to hESCs at the proteomic level. The proteomes of iPSCs induced by introducing 3, 5, or 6 transcription factors were also analyzed. Our proteomic profiles provide valuable insight into the factors that contribute to the similarities and differences between hESCs and hiPSCs and the mechanisms of reprogramming.

Introduction

Induced pluripotent stem cells (iPSCs) are somatic cells that have been reprogrammed to a pluripotent state. In 2006, Yamanaka and coworkers reported that mouse embryonic fibroblasts (MEFs) could be reprogrammed to a pluripotent state by introducing 4 transcription factors Oct4, Sox3, Klf4, and c-Myc via retroviral delivery [1]. These cells exhibit many of the features that are characteristic of embryonic stem cells (ESCs), such as testing positive for alkaline phosphatase (ALP) and ES cell-specific surface markers, expressing Nanog, differentiating into all 3 germ layers, exhibiting transcriptional and epigenetic similarities with ESCs, and forming teratomas in immunodeficient mice [2–6]. Since then, iPSC technology has been received with great excitement in the medical world because of the potential to generate patient-derived pluripotent stem cells as sources for cell therapies for a variety of disorders, including many degenerative diseases.

Although iPSCs possess enormous therapeutic potential, there are several hurdles that must be overcome before iPSCs can be used for patient-specific cell therapies, including the tumorigenic potential of c-Myc, the use of retroviral infection, and the low efficiency of iPSC derivation [7,8]. The critical issue is whether iPSCs are identical to ESCs. It is generally accepted that iPSCs closely resemble ESCs morphologically, molecularly, and developmentally [9–12]. Similarities between iPSCs and ESCs have been identified by microarray studies, high-throughput sequencing, and analysis of DNA methylation status. However, there have been no detailed reports about comparative analyses at the proteomic level [13].

Proteomics can provide a global, systematic, and comprehensive approach to the identification of the biochemical processes, pathways, and networks involved in various physiological states at the protein level. Although nongel-based approaches, such as stable isotope labeling with amino acids in cell culture (SILAC) and surface-enhanced laser desorption/ionization (SELDI), have been used mainly for quantitative proteomic analysis [14–16], 2-dimensional electrophoresis (2-DE) coupled with mass spectrometry (MS) is a well-established and reliable method for investigation of differential protein expression.

The purpose of this study was to perform a detailed analysis of the proteomes of somatic donor cells, human iPSCs (hiPSCs) derived from the corresponding somatic cells, and human ESCs (hESCs) to validate the usefulness of hiPSCs at the proteomic level. From the data of comparative proteomic analysis, we suggest that hiPSCs and hESCs are very similar at the proteomic level, strengthening the usefulness of hiPSCs. However, several proteins with differential expression pattern between hiPSCs and hESCs were detected.

[1]Medical Proteomics Research Center, KRIBB, Daejeon, South Korea.
[2]Department of Biological Sciences, KAIST, Daejeon, South Korea.
[3]Development and Differentiation Research Center, KRIBB, Daejeon, South Korea.
[4]Proteome Research Team, Korea Basic Science Institute, Daejeon, South Korea.

Materials and Methods

Generation of hiPSCs

Human newborn foreskin fibroblasts (hFFs; CRL-2097; ATCC) were grown in Dulbecco's modified Eagle's medium (DMEM) containing 10% fetal bovine serum (Invitrogen), 1% nonessential amino acids (NEAA; Invitrogen), 1 mM L-glutamine (Invitrogen), and 0.1 mM β-mercaptoethanol (Sigma, St. Louis, MO). For reprogramming, hFFs (1×10^5 cells per well) were transduced with pMX-based retroviruses encoding human Oct4, Sox2, Klf4, and c-Myc (Addgene, Inc.) in 6-well culture dishes. Five days after transduction, the hFFs were replated on gelatin-coated 6-well dishes ($5–6 \times 10^4$ cells per well) that had been preplated with MEF feeder cells. On the following day, the medium was replaced with hESC medium supplemented with 10 ng/mL basic fibroblast growth factor (bFGF). The medium was changed every other day. Colonies that exhibited an hES-like morphology were picked 20–23 days after transduction and transferred to 12-well dishes that had been preplated with MEF feeder cells. Selected hiPSC colonies were expanded under standard hESC culture conditions and used for further analyses.

Maintenance of hiPSCs and hESCs

Undifferentiated hESCs (H9; WiCell Res. Ins.) and established hiPSCs were routinely maintained on γ-irradiated MEFs in hESC culture medium consisting of 80% DMEM/F12 medium, 20% knockout serum replacement (Invitrogen), 1% nonessential amino acids (NEAA; Invitrogen), 1 mM L-glutamine (Invitrogen), 0.1 mM β-mercaptoethanol (Sigma), and 6 ng/mL bFGF (Invitrogen). The cells were passaged once per week using mechanical or collagenase-based enzymatic methods as previously described [17].

ALP staining

ALP staining was performed using a commercially available ALP kit according to the manufacturer's instructions (Sigma). Images of ALP-positive cells were recorded using an HP Scanjet G4010. Bright-field images were obtained using an Olympus microscope (IX51; Olympus).

Polymerase chain reaction analysis

Total RNA was isolated from cells using the RNeasy Mini Kit (Qiagen) and reverse-transcribed using the SuperScript First-strand Synthesis System Kit (Invitrogen) according to the manufacturers' protocols. Semiquantitative reverse transcriptase (RT)-polymerase chain reaction (PCR) was performed using the Platinum Taq SuperMix kit (Invitrogen) under the following conditions: 3 min at 94°C, followed by 30 cycles of 30 s at 94°C, 30 s at 60°C, and 30 s at 72°C, and a final extension for 10 min at 72°C. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was used as an endogenous control. Quantitative real-time PCR was performed using the Power SYBR Green PCR Master Mix (Applied Biosystems) and the ABI 7500 Real-Time PCR System (Applied Biosystems). After activating Taq polymerase at 95°C for 15 min, the reactions were denatured at 95°C for 15 s and annealed and elongated at 60°C for 1 min; this process was repeated for 50 cycles. The PCR products were separated using a 1.5% agarose gel containing ethidium bromide and visualized with a Gel Doc EQ system (Bio-Rad). All experiments were performed in triplicate, and the cycle threshold (CT) value for each target gene was determined using the software provided by the manufacturer and normalized to the expression level of GAPDH. The primer sequences used in the present study are shown in Supplementary Table S1 (Supplementary Data are available online at www.liebertonline.com/scd).

Immunocytochemistry

Cells that had been cultured on gelatin-coated 4-well Lab-Tek chamber slides (Nunc) were fixed with 4% paraformaldehyde, permeabilized in phosphate-buffered saline (PBS)/0.2% bovine serum albumin (BSA)/0.1% Triton X-100, and blocked with 4% normal donkey serum (Molecular Probes) in PBS/0.2% BSA for 1 h at room temperature. After blocking, the cells were incubated with the respective primary antibodies [anti-Oct4 (Santa Cruz Biotechnology), anti-Nanog, anti-SSEA4, and anti-Tra1-60 (R&D Systems)] diluted in PBS/0.2% BSA. After washing, the cells were incubated with FITC- or Alexa 594-conjugated secondary antibodies (Invitrogen) in PBS/0.2% BSA for 1 h at room temperature. The chamber slides were analyzed using an Olympus microscope or an Axiovert 200M microscope (Carl Zeiss).

Karyotype analysis

Long-term-maintained hEBs were processed for chromosomal G-band analysis by GenDix Inc. A representative image was captured by ChIPS-Karyo (Chromosome Image Processing System; GenDix, Inc.).

Isoelectric focusing and 2-D gel electrophoresis

The cell lysates were obtained from hFFs (passage 3), hiPSC lines (passages 15–20), and H9 hESCs (passage 33). Cell lysates (150 μg of protein) were mixed with rehydration buffer [9 M urea, 4% 3-[(3-cholamidopropyl)-dimethylammonio]-1-propane sulfonate (CHAPS), 2 M thiourea, 40 mM dithiothreitol (DTT), and 2% immobilized pH gradient (IPG) buffer]. Protein samples were directly applied to IPG strips (pH 3–10, 13 cm) and rehydrated for 14 h at room temperature. Next, isoelectric focusing was performed using the Multiphor II (GE Healthcare) apparatus. The initial voltage was maintained at 300 V for 1 min and then linearly increased from 300 to 3,500 V within 1.5 h. The voltage was maintained at 3,500 V for 8 h. The plate temperature was kept constant at 25°C during isoelectric focusing. The focused IPG strips were briefly equilibrated for 15 min with equilibration solution [50 mM Tris-HCl (pH 8.8), 6 M urea, 2% sodium dodecyl sulfate (SDS), and 30% glycerol] containing 1% DTT, followed by equilibration for 15 min in the same solution containing 5% iodoacetamide instead of DTT. The equilibrated strips were directly loaded onto 13% polyacrylamide gels ($150 \times 150 \times 1.5\ mm^3$) or stored at −80°C until use in subsequent experiment. Polyacrylamide gels loaded with IPG strips were run at a constant current of 20 mA per gel with the PROTEAN II Xi/XL system (Bio-Rad).

Staining and image analysis

After electrophoresis, the gels were fixed, and the protein spots were visualized by silver staining (PlusOne Silver

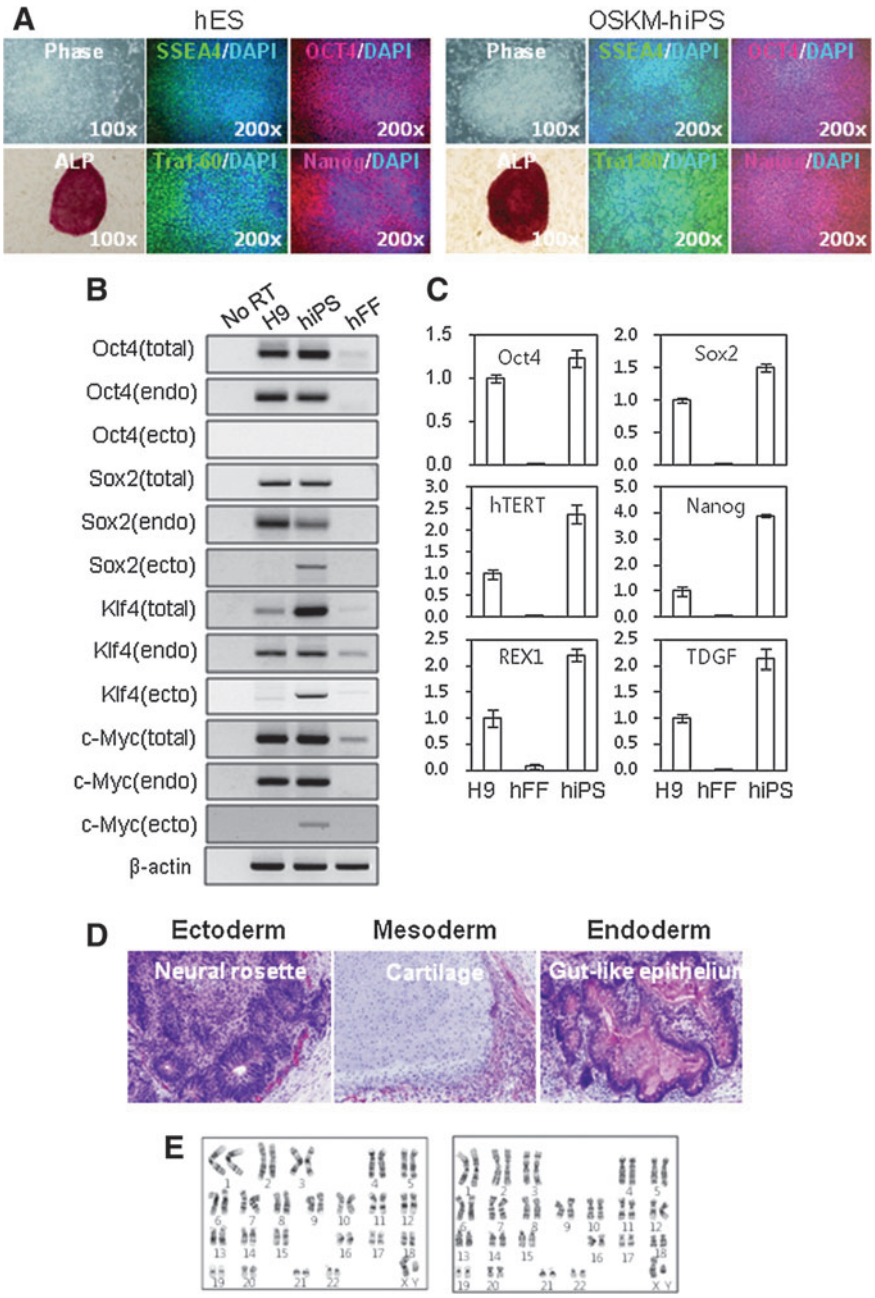


FIG. 1. Characterization of hiPSCs derived from hFFs. **(A)** Morphology of representative H9 hESC and OSKM-derived hiPSC colonies with high levels of ALP and positive immunostaining for the pluripotency markers: OCT4, Nanog, SSEA4, and TRa-1–60. **(B)** Semiquantitative RT-PCR analysis of pluripotency markers (Oct4, Sox2, Klf4, and c-Myc) in hFF, H9 hESCs (H9), and hiPSCs (hiPS). β-Actin was used as a loading control. **(C)** Quantitative real-time PCR analysis of pluripotency markers (Oct4, Sox2, Nanog, hTERT, Rex1, and TDGF) in hFFs, H9 hESCs (H9), and hiPSCs (hiPS). **(D)** Histological analysis of a teratoma derived from hiPSCs. Hematoxylin-eosin-stained sections from teratomas generated from OSKM-derived hiPSCs are shown. Differentiation into multiple derivatives of the 3 germ layers is shown (original magnification, ×200). **(E)** Normal karyotype of hFFs (*left*) and hiPSCs induced by 4 core transcription factors (OSKM-hiPSCs) (*right*). iPSCs, induced pluripotent stem cells; hiPSC, human iPSCs; hFFs, human newborn foreskin fibroblasts; ESCs, embryonic stem cells; hESCs, human ESCs; ALP, alkaline phosphatase; RT-PRC, reverse transcriptase-polymerase chain reaction.

Staining Kit; GE Healthcare). The 2-DE images were scanned and processed with Progenesis SameSpots v3.0 software (Nonlinear Dynamics Ltd.). Spot volumes were normalized based on the total spot volume of each gel. Protein spot intensity was defined as the normalized spot volume, that is, the ratio of the single spot volume to the total of the spot volumes on the 2-DE gel (total spot normalization). Computer analysis facilitated the automatic detection and quantification of protein spots and matches among gels of somatic cells, hiPSCs, and hESCs. Spots displaying reliable and significant differences (greater than 2.0-fold, $P < 0.05$) were selected for MS analysis.

In-gel digestion and identification by liquid chromatography–MS/MS

Spots of interest were manually excised from 2-DE gels and destained with chemical reducers to remove the silver [18–20]. Briefly, 50–100 mL of the freshly prepared reducing solution (1:1 mixture of 30 mM potassium ferricyanide and 100 mM sodium thiosulfate) was added to the gel plugs and mixed. After the brown color disappeared, the gel plugs were rinsed with water and incubated in 200 mM ammonium bicarbonate for 20 min. Subsequently, the gel plugs were cut into small pieces, washed with water, and dehydrated repeatedly with acetonitrile (ACN) until the pieces appeared opaque and white. Next, the gel pieces were dried in a vacuum centrifuge for 30 min, and the proteins were digested with 20 ng/mL of sequencing grade-modified trypsin (Promega) for 16–24 h at 37°C. Digested peptides were extracted with extraction solution (50% ACN and 5% trifluoroacetic acid [TFA]), and the extracted peptides were dried using a vacuum drier. Samples were then subjected to MS analysis.

Peptides were analyzed using a Synapt High Definition Mass Spectrometer (Waters) equipped with a nanoACQUITY Ultra

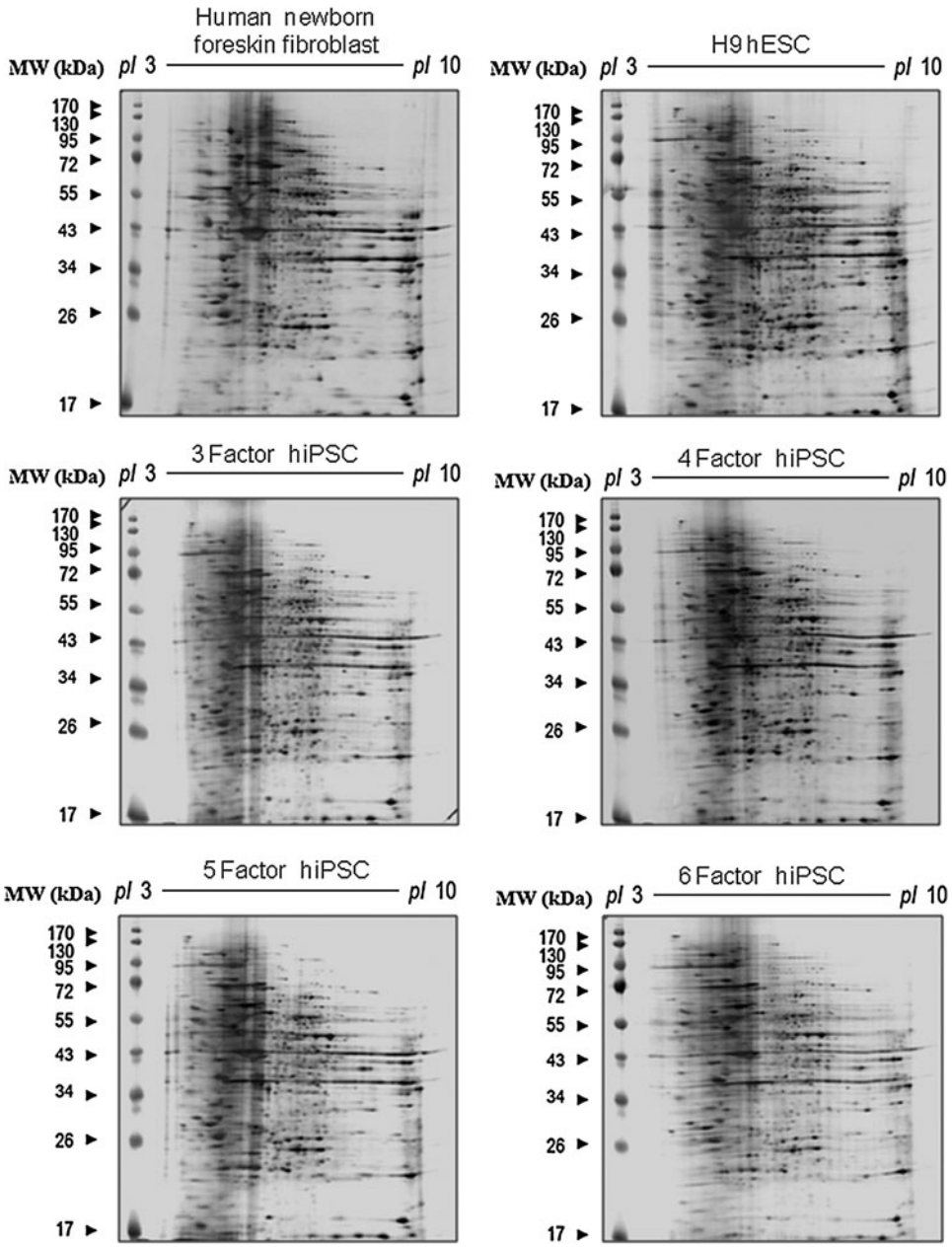


FIG. 2. Representative 2-DE gel images of proteins from donor cells, hESCs, and OSKM-hiPSCs. Total protein lysates were separated on pH 3–10 nonlinear immobilized pH gradient (IPG) strips in the first dimension, followed by 13% sodium dodecyl sulfate–polyacrylamide gel electrophoresis in the second dimension and subsequent visualization by silver staining. Protein spots that were differentially expressed compared with donor are indicated with circles (downregulated) or rectangles (upregulated). Spots were identified by liquid chromatography–mass spectrometry/mass spectrometry as outlined in Table 1. Negative control, hFF; positive control, H9 hESC; 3 Factor hiPSC, hiPSCs induced by Oct4, Sox2, and Klf4; 4 Factor hiPSC, hiPSCs induced by Oct4, Sox2, Klf4, and c-Myc; 5 Factor hiPSC, hiPSCs induced by Oct4, Sox2, Klf4, c-Myc, and hTERT; 6 Factor hiPSC, hiPSCs induced by Oct4, Sox2, Klf4, c-Myc, hTERT, and SV40 large T antigen; IPG, immobilization pH gradient.

Performance LC system (Waters). A volume of 2 μL of the peptide solution was injected onto a 75 m × 100 mm Atlantis dC18 column (Waters). Solvent A consisted of 0.1% formic acid in water, and Solvent B consisted of 0.1% formic acid in ACN. Peptides were initially separated using 100 min gradients and electrosprayed into the mass spectrometer (fitted with a nanoLockSpray source) at a flow rate of 300 nL/min. Mass spectra were acquired from m/z 300–1,600 for 1 s, followed by 4 data-dependent MS/MS scans from m/z 50–1,900 for 1 s each. The collision energy used to perform MS/MS was varied according to the mass and charge state of the eluting peptide. (Glu1)-Fibrinopeptide B was infused at a rate of 350 nL/min, and an MS scan was acquired for 1 s every 30 s throughout the run. A database search was performed with MASCOT (Matrix Science) using the following parameters: NCBInr.08.03.26 database, *Homo sapiens* species, and maximum number of missed cleavages by trypsin set to 1. The mass tolerance ranged from ± 50 to ± 100 ppm. The peptide modification allowed was carbamidomethylation in the fixed modification mode.

Target validation using western blot analysis

Protein samples (20 μg) were separated on a 13% SDS–polyacrylamide gel electrophoresis gel and transferred to a nitrocellulose membrane using standard procedures. The membrane was blocked with 5% (v/v) skim milk in TBS-T buffer [20 mM Tris-HCl (pH 7.6), 0.1369 M NaCl, and 0.1% Triton X-100] and then incubated with the primary antibody for 12 h on a rocking platform at 4°C. The membrane was then washed 3 times with TBS-T buffer for 15 min and incubated with 5% skim milk in TBS-T buffer containing horseradish peroxidase-conjugated secondary antibody (diluted to 1:3,000) for 1 h. The hybridized membrane was washed in TBS-T buffer and visualized using a chemiluminescent ECL detection kit (GE Healthcare).

Bioinformatic data analysis

Differentially expressed proteins were evaluated by Ingenuity Pathway Analysis (IPA; Ingenuity System). IPA is a software application that enables the identification of the biological mechanisms, pathways, and functions matching a particular dataset of proteins. IPA is based on a database obtained by abstracting and interconnecting a large fraction of the existing biomedical literature according to a strict algorithm. This database integrates protein functions, cellular localizations, small molecules, and disease interrelationships. The networks are displayed graphically as nodes, representing individual proteins, and edges, representing the biological relationships between nodes. Networks are ordered by score and optimized to include as many differentially expressed proteins as possible. A P score [ie, $-\log(P$ value)] for each possible network is computed. Therefore, networks with scores of 2 or higher have at least 99% confidence of having not been generated by random chance alone.

Statistical analysis

Experimental differences were tested for statistical significance using analysis of variance and Student's t-test. P values less than 0.05 were regarded as statistically significant.

Results and Discussion

Generation of hiPSCs

The hiPSCs used in this study were derived from hFFs by retroviral overexpression of Oct4 (O), Sox2 (S), Klf4 (K), and c-Myc (M). Reprogrammed hiPSCs exhibited an hESC-like morphology and strong ALP activity (Fig. 1A). hiPSCs were positive for Oct4, Nanog, SSEA-3, and Tra-1–61 proteins, as determined by immunohistochemistry (Fig. 1B). Quantitative real-time RT-PCR analysis showed that the expression of pluripotency markers, including Oct4, Sox2, Nanog, hTERT, Rex1 and TDGF, were markedly increased relative to donor somatic cells, which had pluripotency marker expression levels that were comparable to those of H9 hESCs (Fig. 1C). Semiquantitative RT-PCR analysis indicated that the hiPSCs silenced the expression of the retroviral transgenes (Oct4, Sox2, Klf4, and c-Myc) (Fig. 1C). The pluripotency of hiPSCs induced by 4 core transcription factors (OSKM-hiPSCs) was further confirmed by in vivo teratoma formation assays. As shown in Fig. 1D, the teratomas contained well-defined structures arising from all 3 germ layers. In addition, the hiPSCs retained a normal karyotype during in vitro culture (Fig. 1F).

Proteomic analysis of donor cells, hiPSCs, and hESCs by 2-DE

To establish the proteomic profiles of hFFs, OSKM-derived hiPSCs, and hESCs, protein lysates from these cells were separated by 2-DE (Fig. 2). Experiments were performed in triplicate. More than 1,800 protein spots were detected on gels after silver staining, automatic spot detection, background subtraction, and volume normalization. In 3 sets of experiments, protein spots displaying significant changes (greater than 2-fold changes when comparing somatic fibroblast cells to hESCs, somatic fibroblast cells to hiPSCs, and hiPSCs to hESCs) were scored and identified (Supplementary Fig. S1 and Table 1). In total, 65 spots were scored. Forty-three of which exhibited changes between the hFF donor cells and OSKM-hiPSCs (Figs. 2 and 3A), demonstrating that the proteomic pattern of the donor cells could be altered by introducing 4-core pluripotency transcription factors. Among the 43 spots, 18 also changed in the same pattern in hESCs. For spot no. 299 (identified as serpin peptidase inhibitor, clade B, member 9 by MS analysis), decreased expression was detected in OSKM-hiPSCs, whereas increased expression was detected in hESCs. Spots showing differential expression patterns between hESCs and OSKM-hiPSCs were also detected (Figs. 2 and 3A). Of 15 spots showing differential expression patterns, 7 spots were increased and 8 spots were decreased in OSKM-hiPSCs compared with hESCs. In addition to OSKM-hiPSCs, we further analyzed the proteomic profiles of hiPSCs induced by 3 factors (Oct4, Sox2, and Klf4), 5 factors (Oct4, Sox2, Klf4, c-Myc, and SV40 large T antigen), and 6 factors (Oct4, Sox2, Klf4, c-Myc, SV40 large T antigen, and hTERT) (Fig. 2). The global expression intensity map generated from the proteomic data clearly demonstrated that hiPSCs induced by the introduction of 5 factors possess the proteomic pattern closest to that of hESCs (Fig. 3B). On the other hand, OSKM-hiPSCs showed the proteomic pattern with lowest similarity to hESCs. iPSCs have been reportedly obtained from primary skin fibroblasts derived from a healthy adult donor via the aid of 2 additional factors, hTERT and SV40 large T antigen [10]. However, the detailed roles of hTERT (the catalytic subunit of human telomerase) and SV40 large T antigen in reprogramming remain unclear. Thus, identifying proteins that are differentially expressed in hiPSCs that are induced by 5 and 6 factors when compared with other hiPSCs, donor cells, and hESCs would be valuable in understanding the detailed roles played by hTERT and SV40 large T antigen in the generation of iPSCs (Table 2). For example, like H9 ESCs, Far upstream element binding protein 3 (FUBP-3; spot no. 1257) and Cu/Zn-superoxide dismutase (spot no. 416) are highly expressed (>2-fold) in only hiPSCs induced by the introduction of 6 factors among iPSCs. This may be caused by the introduction of hTERT.

Identification and classification of differentially expressed proteins by liquid chromotography–MS/MS

The spots with differential expression patterns among hFFs, hESCs, and OSKM-hiPSCs were identified. A total of 65 proteins were identified by MS analysis (Table 1). The proteins identified were classified according to their biological process, molecular function, and cellular location (Fig. 3C). The proteins were classified into various functional groups, implying that significant changes occur during reprogramming and many proteins may contribute to the pluripotency of hESCs and hiPSCs.

Among the identified proteins, a number of proteasome subunit proteins were differentially expressed. The proteasome has been implicated in the dynamic control of

Table 1. List of Proteins Identified from Comparative Proteomics Using Mass Spectrometry

Spot No.	Accession No.	Protein name	Function	Matched peptides	Moscot score	Coverage (%)	Fold change: H9 ESC/hFF	Fold change: 4F iPSC/hFF	Fold change: 4F iPSC/H9 ESC
230	gi|3041664	Deoxyuridine 5′-triphosphate nucleotidohydrolase, mitochondrial (DUT)	Nucleotide metabolism	2	131	6	3.24 ↑	9.44 ↑	2.91 ↑
299	gi|4758906	Serpin peptidase inhibitor, clade B, member 9 (SERPIN 9)	Catalytic protein	27	580	64	2.88 ↑	2.66 ↓	7.66 ↓
392	gi|5031981	26S proteasome-associated pad1 homolog (PSMD 14)	Proteasome activator activity	14	594	47	5.88 ↑	2.51 ↑	2.34 ↓
435	gi|5453880	Acidic (leucine-rich) nuclear phosphoprotein 32 family, member A (ANP32A)	Transcription regulation	16	591	32	4.72 ↑	2.03 ↑	2.33 ↓
15	gi|154355000	KHSRP protein (FUBP 2)	Transcription regulation	22	700	31	3.50 ↓	2.89 ↓	1.21 ↑
36	gi|609342	Nucleophosmin-anaplastic lymphoma kinase fusion protein (NPM-ALK)	Nervous system development	11	534	9	3.95 ↑	4.49 ↑	1.14 ↑
44	gi|4757774	ADP-ribosylation factor-like 3 (ARL 3)	Cytokinesis	6	210	49	3.03 ↑	3.98 ↑	1.32 ↑
64	gi|5453549	Thioredoxin peroxidase (PRDX 4)	Cell redox homeostasis	14	497	44	2.28 ↓	2.51 ↓	1.10 ↓
81	gi|4507357	Transgelin 2 (TAGLN 2)	Muscle organ development	4	173	15	1.32 ↑	2.67 ↑	2.02 ↑
356	gi|31645	GAPDH	Glycolysis	2	72	8	1.84 ↑	3.93 ↑	2.14 ↑
389	gi|54648253	KHSRP protein (FUBP 2)	Transcription regulation	10	212	18	2.97 ↓	3.52 ↓	1.19 ↓
416	gi|1237406	Cu/Zn-superoxide dismutase (SOD 1)	Oxidation reduction	10	347	40	5.59 ↑	1.54 ↑	3.64 ↓
601	gi|4506179	Proteasome alpha 1 subunit isoform 2 (PSMA 1)	Proteasome activator activity	32	1081	79	4.18 ↑	3.42 ↑	1.22 ↓
730	gi|26051237	Nucleoporin 54 kDa (NUP 54)	Nuclear pore transport	7	160	15	2.70 ↓	2.07 ↓	1.30 ↑
756	gi|18379349	Vesicle amine transport protein 1 (VAT 1)	Oxidation reduction	12	364	35	3.46 ↑	2.60 ↑	1.33 ↓
767	gi|2081622	UMP synthase (UMPS)	UMP biosynthetic process	26	820	48	3.41 ↑	2.05 ↑	1.66 ↓
806	gi|32189392	Peroxiredoxin 2 isoform a (PRDX 2)	Cell redox homeostasis	14	292	36	3.27 ↑	1.57 ↑	2.08 ↓
820	gi|54648253	KHSRP protein (FUBP2)	Transcription regulation	20	741	31	2.32 ↓	2.22 ↓	1.04 ↑
854	gi|338039	Set	Regulation of histone acetylation	8	408	16	2.39 ↑	2.77 ↑	1.16 ↑
861	gi|5453880	Acidic (leucine-rich) nuclear phosphoprotein 32 family, member A (ANP32A)	Transcription regulation	14	309	32	2.07 ↑	2.34 ↑	1.13 ↑
904	gi|671527	Gamma subunit of CCT chaperonin (CCT 3)	Protein folding	23	546	40	3.00 ↑	2.79 ↑	1.07 ↓
922	gi|35068	Nonmetastatic cells 1 protein (NME 1)	Nucleotide metabolism	23	644	59	2.72 ↑	1.08 ↓	2.94 ↓
1021	gi|976227	26S proteasome subunit p45 (PSMC 5)	Transcription cofactor activity	14	367	37	2.71 ↑	2.03 ↑	1.34 ↓
1240	gi|5031981	26S proteasome-associated pad1 homolog (PSMD 14)	Proteasome activator activity	16	588	47	2.19 ↑	1.04 ↓	2.27 ↓
1257	gi|1575609	FUSE binding protein 3 (FUBP 3)	Transcription regulation	5	115	4	2.32 ↓	1.12 ↓	2.07 ↑
1312	gi|4505763	Phosphoglycerate kinase 1 (PGK 1)	Glycolysis	7	149	35	2.26 ↑	2.16 ↑	1.05 ↓
57	gi|16950633	Argininosuccinate synthetase 1 (ASS 1)	Arginine biosynthesis	10	332	19	1.55 ↓	2.84 ↓	1.83 ↓
61	gi|33803	Set	Regulation of histone acetylation	7	240	20	1.71 ↑	2.38 ↑	1.39 ↑
75	gi|5803076	Heterochromatin protein 1-beta (CBX 1)	Chromatin assembly or disassembly	3	147	9	2.70 ↑	1.79 ↑	1.51 ↓
88	gi|9910244	Mitochondrial ribosomal protein S22 (MRPS 22)	Glycolysis	27	782	53	1.72 ↑	2.07 ↑	1.21 ↑
594	gi|13899241	Ionized calcium binding adapter	Cytoskeletal protein binding	5	99	31	1.93 ↑	2.86 ↑	1.48 ↑

(*continued*)

Table 1. (Continued)

Spot No.	Accession No.	Protein name	Function	Matched peptides	Moscot score	Coverage (%)	Fold change: H9 ESC/hFF	Fold change: 4F iPSC/hFF	Fold change: 4F iPSC/H9 ESC
		molecule 2 isoform 1 (AIFL 1)							
844	gi|22538465	Proteasome beta 3 subunit (PSMB 3)	Proteasome activator activity	20	553	53	2.17 ↑	1.47 ↑	1.48 ↓
860	gi|5174447	Guanine nucleotide binding protein (G protein), beta polypeptide 2-like 1 (GNB2L1)	Protein binding	15	339	36	1.71 ↑	2.91 ↑	1.70 ↑
892	gi|251370	Acid phosphatase isoenzyme Af [human, erythrocytes, Peptide, 157 aa]	Catalytic protein	3	242	24	1.87 ↑	2.94 ↑	1.57 ↑
903	gi|21758578	Nucleoside phosphorylase (PNP)	Catalytic protein	26	829	66	1.81 ↑	3.00 ↑	1.66 ↑
909	gi|119617360	Prostaglandin E synthase 3 (cytosolic), isoform CRA_b (PTGES 3)	Signal transduction	4	233	13	1.39 ↑	2.31 ↑	1.66 ↑
920	gi|4506217	Proteasome 26S non-ATPase subunit 10 isoform 1 (PSMD 10)	Transcription regulation	14	435	44	2.95 ↑	1.98 ↑	1.49 ↓
926	gi|17402893	Phosphoserine aminotransferase isoform 1 (PSAT 1)	Amino-acid biosynthesis	10	204	26	1.26 ↑	2.29 ↑	1.81 ↑
930	gi|5031851	Stathmin 1 isoform a (STMN 1)	Cell differentiation	7	213	32	1.66 ↑	2.66 ↑	1.60 ↑
972	gi|1160963	Transmembrane protein (IMMT)	Protein binding	25	942	34	2.38 ↓	1.66 ↓	1.43 ↑
994	gi|178277	*S*-Adenosylhomocysteine hydrolase (AHCY)	1-carbon metabolism	30	774	38	2.48 ↑	1.41 ↑	1.77 ↓
1037	gi|14790115	Caspase 3 preproprotein (CASP 3)	Apoptosis	2	61	14	1.65 ↑	2.68 ↑	1.62 ↑
1040	gi|2781202	Chain A, 3-dimensional structure of human electron transfer flavoprotein to 2.1 A resolution	Protein transport	2	101	7	1.71 ↑	2.68 ↑	1.56 ↑
1057	gi|16878077	FUBP1 protein (FUBP 1)	Transcription regulation	13	559	28	1.24 ↑	2.33 ↑	1.33 ↑
1072	gi|4929769	CGI-150 protein (GLOD 4)	Metabolism	7	190	14	2.07 ↓	1.88 ↓	1.10 ↑
1095	gi|16878077	FUBP1 protein (FUBP 1)	Transcription regulation	8	167	13	1.50 ↑	2.57 ↑	1.72 ↑
1112	gi|16924265	Enoyl coenzyme A hydratase 1, peroxisomal (ECH 1)	Lipid metabolism	8	268	28	2.55 ↑	1.69 ↑	1.51 ↓
1124	gi|13027378	Glucosamine-6-phosphate deaminase 1 (GNPDA 1)	Carbohydrate metabolism	4	233	13	1.34 ↑	2.52 ↑	1.87 ↑
1181	gi|48255905	Transgelin (TAGLN)	Muscle organ development	19	605	64	2.43 ↓	1.73 ↓	1.40 ↑
1193	gi|4092058	Proteasome subunit HSPC (PSMA 7)	Proteasome activator activity	13	567	55	1.34 ↓	1.55 ↑	2.08 ↑
1194	gi|4504035	Guanine monophosphate synthetase (GMPS)	Purine biosynthesis	33	657	48	1.21 ↑	2.41 ↑	1.99 ↑
1209	gi|32129199	Cytokine induced protein 29 kDa (SARNP)	Transcription regulation	4	163	18	1.37 ↑	2.18 ↑	1.59 ↑
1281	gi|5453595	Adenylyl cyclase-associated protein (CAP 1)	Actin binding protein	11	306	24	1.61 ↓	2.01 ↓	1.25 ↓
1290	gi|546901	Cytosolic acetoacetyl-coenzyme A thiolase (ACAT 2)	Lipid metabolic process	26	746	43	2.06 ↑	1.91 ↑	1.08 ↓
1320	gi|4758484	Glutathione-*S*-transferase omega 1	Catalytic protein	10	193	28	1.30 ↓	2.26 ↓	1.74 ↓

(*continued*)

TABLE 1. (CONTINUED)

Spot No.	Accession No.	Protein name	Function	Matched peptides	Moscot score	Coverage (%)	Fold change		
							H9 ESC/hFF	4F iPSC/hFF	4F iPSC/H9 ESC
		(GSTO 1)							
1332	gi|847724	Methylthioadenosine phosphorylase (MTAP)	Nucleoside metabolic process	3	170	19	1.65 ↑	2.19 ↑	1.32 ↑
1358	gi|10863927	Peptidylprolyl isomerase A (PPIA)	Protein folding	18	381	54	1.16 ↑	2.22 ↑	1.92 ↑
1376	gi|4504035	Guanine monophosphate synthetase (GMPS)	Purine biosynthesis	6	164	7	1.13 ↑	2.20 ↑	1.94 ↑
1380	gi|5031635	Cofilin 1 (non-muscle) (CFL1)	Signal transduction	4	110	21	1.26 ↓	1.74 ↑	2.20 ↑
1392	gi|4557032	L-Lactate dehydrogenase B (LDHB)	Glycolysis	63	1795	60	2.17 ↑	1.16 ↑	1.88 ↓
1406	gi|3646128	Thioredoxin-like protein (GLRX 3)	Cell redox homeostasis	14	507	41	1.95 ↑	1.10 ↓	2.15 ↓
1411	gi|19743875	Fumarate hydratase precursor (FH)	Fumarate metabolic process	8	200	17	2.15 ↑	1.49 ↑	1.44 ↓
1446	gi|4826659	F-actin capping protein beta subunit (CAPZB)	Actin cytoskeleton organization	18	540	56	1.80 ↓	1.17 ↑	2.11 ↑
1471	gi|4507645	Triosephosphate isomerase 1 (TPI 1)	Glycolysis	13	320	63	2.08 ↓	1.52 ↓	1.37 ↑
1488	gi|4506195	Proteasome beta 2 subunit (PSMB 2)	Proteasome activator activity	7	235	28	1.59 ↑	2.06 ↑	1.30 ↑

ESC, embryonic stem cell; hFF, human newborn foreskin fibroblast; iPSC, induced pluripotent stem cell.

factor binding to transcriptionally active promoters [21]. Tissue-specific gene loci are maintained in a state that is competent for future expression but remains inactive at the pluripotent stage. At that stage, stable binding of transcription factors and RNA polymerase II to specific sequence elements is inhibited by the proteasome [22]. The upregulation of several proteasome subunits in hESCs and hiPSCs compared with donor cells would therefore be expected, and this observation further supports our experimental design in profiling the proteomes of hFFs, hESCs, and hiPSCs.

Validation of proteins by western blot analysis

Western blotting was used to validate the 2-DE results and assess the expression changes of several proteins that showed differential patterns (Fig. 4). With the exception of molecular chaperones and proteins involved in the cytoskeleton, we tested all possible identified proteins by western blot analysis using commercially available antibodies. Generally, the western blot results correlated well with the 2-DE data (Table 1 and Fig. 4).

Among the identified proteins, heterochromatin protein 1-β (HP1β) was significantly upregulated in hESCs and hiPSCs, compared with donor cells (Fig. 4). HP1β is a highly conserved nonhistone protein and a member of the heterochromatin protein family. The N-terminal chromodomain of HP1β can bind histone proteins, and the C-terminal chromo shadow-domain is involved in heterodimerization and interactions with a variety of chromatin-associated nonhistone proteins. This protein plays a crucial role in the epigenetic regulation of chromatin structure and DNA repair [23,24]. A variety of epigenetic changes take place during the reprogramming process [25,26]. The differential expression pattern of HP1β in our experiments suggests that HP1β may be involved in chromatin remodeling during reprogramming.

ESCs and iPSCs have lower levels of mitochondrial mass and oxidative phosphorylation and higher levels of lactate production than mature or differentiated cells. Further, hypoxia significantly enhances reprogramming efficiency [27], implying that ESCs and iPSCs preferentially use nonoxidative glycolysis as a main energy source [28–30]. In our experiments, several glycolytic enzymes, including GAPDH, phosphoglycerate kinase 1, triose phosphate isomerase 1, and lactate dehydrogenase B, were differentially expressed in hESCs and hiPSCs relative to hFFs, indicating that glycolytic metabolism is the main energy generation system in hESCs and hiPSCs. Notably, a more dramatic increase in GAPDH was detected in OSKM-hiPSCs than in hESCs. GAPDH is a multifunctional protein with multiple intracellular localizations and plays key roles in diverse cellular processes independent of its traditional role in glycolysis; these processes include DNA repair, membrane fusion and transport, cytoskeletal dynamics, and cell death [31–33]. Therefore, the detailed role of GAPDH in the reprogramming process is likely to be complex.

Decreased expression of nucleoporin p54 (Nup54) was detected in both hESCs and hiPSCs compared with hFFs. Developmental signal transduction involves nucleocytoplasmic transport and occurs through nuclear pore

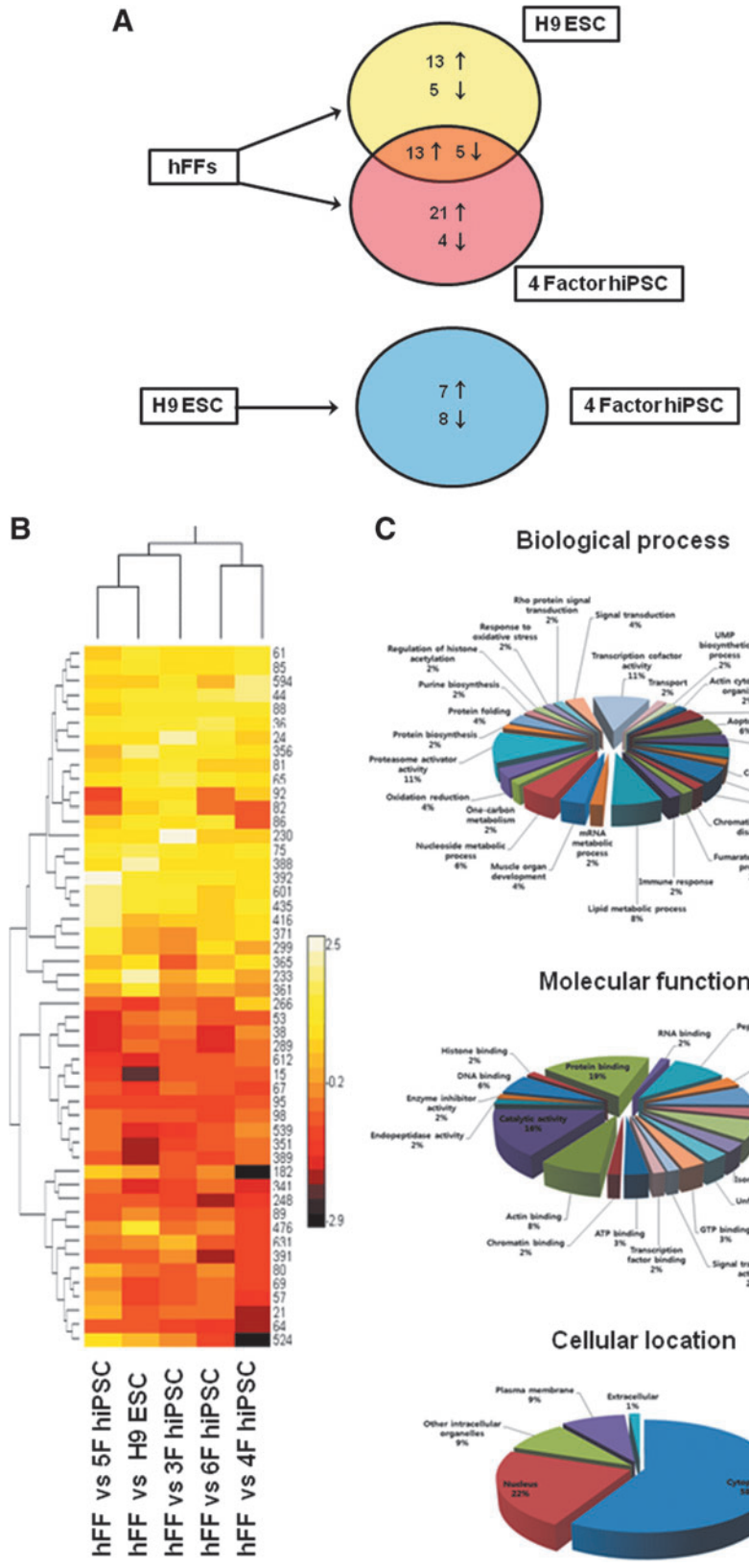


FIG. 3. Bioinformatic analysis of the identified proteins. **(A)** Venn diagrams showing the degree of overlap of differentially expressed proteins between hESCs and OSKM-hiPSCs (4 Factor iPSC). Seventeen proteins were common between H9 hESCs and hiPSCs compared with somatic donor cells. In addition, 15 differentially expressed proteins were detected between H9 hESCs and hiPSCs. The identified proteins are listed in Tables 1 and 2. **(B)** Heat map and hierarchical clustering of the proteome across samples from somatic cells, hESCs, and hiPSCs. **(C)** Functional classification of the identified proteins according to their biological process, molecular function, and cellular location.

complexes. A nuclear pore complex consists of approximately 30 distinct proteins (nucleoporins or Nups). Recent studies have shown distinct roles for different FG-Nups, which have repetitive stretches of Phe-Gly residues [34,35]. Further, it was reported that the composition of the nuclear pore complex is also important during cell differentiation in the developing mouse embryo [36]. In conjunction with our proteomic data, these reports suggest that the composition of the nuclear pore complex may be critical to the reprogramming process. More extensive experiments will be

Table 2. List of Proteins Identified from Comparative Proteomics Using Mass Spectrometry

Spot No.	Accession No.	Protein name	Fold change				
			H9 ESC/hFF	3 F hiPSC/hFF	4 F hiPSC/hFF	5 F hiPSC/hFF	6 F hiPSC/hFF
15	gi|154355000	KHSRP protein (FUBP2)	3.50 ↓	6.31 ↓	2.89 ↓	2.91 ↓	2.86 ↓
36	gi|609342	Nucleophosmin-anaplastic lymphoma kinase fusion protein	3.95 ↑	4.12 ↑	4.49 ↑	4.94 ↑	3.34 ↑
44	gi|4757774	ADP-ribosylation factor-like 3	3.03 ↑	3.93 ↑	3.98 ↑	4.34 ↑	4.51 ↑
57	gi|16950633	Argininosuccinate synthetase 1	1.55 ↓	3.23 ↓	2.84 ↓	2.55 ↓	3.61 ↓
61	gi|33803	Set	1.71 ↑	3.42 ↑	2.38 ↑	2.10 ↑	3.17 ↑
64	gi|5453549	Thioredoxin peroxidase	2.28 ↓	2.44 ↓	2.51 ↓	2.86 ↓	3.27 ↓
75	gi|5803076	Heterochromatin protein 1β	2.70 ↑	2.84 ↑	1.79 ↑	1.62 ↑	1.89 ↑
81	gi|4507357	Transgelin 2	1.32 ↑	1.95 ↑	2.67 ↑	1.52 ↑	1.56 ↑
85	gi|14249382	Abhydrolase domain containing 14B	1.87 ↑	2.25 ↑	1.92 ↑	1.85 ↑	2.38 ↑
86	gi|239552	Squamous cell carcinoma antigen; SCC antigen	1.32 ↑	2.01 ↑	1.77 ↑	2.35 ↑	1.03 ↓
88	gi|9910244	Mitochondrial ribosomal protein S22	1.72 ↑	2.18 ↑	2.07 ↑	2.30 ↑	2.26 ↑
230	gi|3041664	Deoxyuridine 5′-triphosphate nucleotidohydrolase, mitochondrial (dUTPase)	3.24 ↑	3.07 ↑	9.44 ↑	2.32 ↑	2.26 ↑
299	gi|4758906	Serpin peptidase inhibitor, clade B, member 9	2.88 ↑	1.69 ↓	2.66 ↓	1.13 ↑	1.51 ↓
356	gi|31645	GAPDH	1.84 ↑	6.39 ↑	3.93 ↑	4.11 ↑	3.44 ↑
389	gi|54648253	KHSRP protein (FUBP2)	2.97 ↓	5.92 ↓	3.52 ↓	3.83 ↓	2.73 ↓
392	gi|5031981	26S proteasome-associated pad1 homolog	5.88 ↑	2.88 ↑	2.51 ↑	2.48 ↑	2.03 ↑
416	gi|1237406	Cu/Zn-superoxide dismutase	5.59 ↑	1.04 ↑	1.54 ↑	1.35 ↑	2.25 ↑
435	gi|5453880	Acidic (leucine-rich) nuclear phosphoprotein 32 family, member A	4.72 ↑	2.79 ↑	2.03 ↑	1.14 ↓	1.92 ↑
594	gi|13899241	Ionized calcium binding adapter molecule 2 isoform 1	1.93 ↑	2.19 ↑	2.86 ↑	1.57 ↑	4.22 ↑
601	gi|4506179	Proteasome alpha 1 subunit isoform 2	4.18 ↑	2.77 ↑	3.42 ↑	1.04 ↑	1.71 ↑
730	gi|26051237	Nucleoporin 54kDa	2.70 ↓	3.35 ↓	2.07 ↓	1.39 ↓	1.06 ↑
756	gi|18379349	Vesicle amine transport protein 1	3.46 ↑	2.11 ↑	2.60 ↑	2.13 ↑	2.60 ↑
806	gi|32189392	Peroxiredoxin 2 isoform a	3.27 ↑	2.02 ↑	1.57 ↑	1.34 ↑	1.52 ↑
816	gi|1545813	Keio novel protein I (KNP-I a)	1.32 ↑	3.23 ↑	1.42 ↑	1.52 ↑	1.14 ↑
820	gi|54648253	KHSRP protein (FUBP2)	2.32 ↓	3.23 ↓	2.22 ↓	2.91 ↓	2.45 ↓
844	gi|22538465	Proteasome β3 subunit	2.17 ↑	3.15 ↑	1.47 ↑	1.23 ↑	1.21 ↑
854	gi|338039	Set	2.39 ↑	3.13 ↑	2.77 ↑	2.21 ↑	1.57 ↑
860	gi|5174447	Guanine nucleotide binding protein (G protein), β polypeptide 2-like 1 (GNB2L1)	1.71 ↑	2.23 ↑	2.91 ↑	3.12 ↑	2.46 ↑
861	gi|5453880	Acidic (leucine-rich) nuclear phosphoprotein 32 family, member A	2.07 ↑	3.11 ↑	2.34 ↑	2.57 ↑	2.39 ↑
892	gi|251370	Acid phosphatase isoenzyme Af [human, erythrocytes, Peptide, 157 aa]	1.87 ↑	3.03 ↑	2.94 ↑	2.43 ↑	2.37 ↑
900	gi|189054178	Intermediate filament protein	1.75 ↑	3.01 ↑	1.83 ↑	1.84 ↑	1.74 ↑
903	gi|21758578	Nucleoside phosphorylase	1.81 ↑	2.14 ↑	3.00 ↑	2.47 ↑	2.32 ↑
904	gi|671527	Gamma subunit of CCT chaperonin	3.00 ↑	2.42 ↑	2.79 ↑	2.45 ↑	2.44 ↑
909	gi|119617360	Prostaglandin E synthase 3 (cytosolic), isoform CRA_b	1.39 ↑	2.98 ↑	2.31 ↑	1.62 ↑	1.86 ↑

(*continued*)

Table 2. (Continued)

Spot No.	Accession No.	Protein name	Fold change				
			H9 ESC/hFF	3 F hiPSC/hFF	4 F hiPSC/hFF	5 F hiPSC/hFF	6 F hiPSC/hFF
920	gi|4506217	Proteasome 26S non-ATPase subunit 10 isoform 1	2.95 ↑	1.86 ↑	1.98 ↑	1.74 ↑	1.64 ↑
922	gi|35068	Non-metastatic cells 1, protein (Nm23 protein)	2.72 ↑	1.83 ↑	1.08 ↓	1.66 ↑	1.52 ↑
926	gi|17402893	Phosphoserine aminotransferase isoform 1	1.26 ↑	2.93 ↑	2.29 ↑	2.81 ↑	2.45 ↑
930	gi|5031851	Stathmin 1 isoform a	1.66 ↑	2.92 ↑	2.66 ↑	2.19 ↑	1.94 ↑
972	gi|1160963	Transmembrane protein	2.38 ↓	2.17 ↓	1.66 ↓	1.89 ↓	1.19 ↑
981	gi|178277	*S*-Adenosylhomocysteine hydrolase	1.39 ↓	2.02 ↑	1.01 ↑	1.61 ↓	1.87 ↑
994	gi|178277	*S*-Adenosylhomocysteine hydrolase	2.48 ↑	1.11 ↓	1.41 ↑	1.21 ↑	1.39 ↑
1021	gi|976227	26S proteasome subunit p45	2.71 ↑	2.24 ↑	2.03 ↑	1.97 ↑	1.80 ↑
1037	gi|14790115	Caspase 3 preproprotein	1.65 ↑	2.04 ↑	2.68 ↑	2.25 ↑	1.86 ↑
1040	gi|2781202	Chain A, 3-Dimensional Structure Of Human Electron Transfer Flavoprotein To 2.1 A Resolution	1.71 ↑	1.60 ↑	2.68 ↑	2.22 ↑	1.68 ↑
1057	gi|16878077	FUBP1 protein	1.24 ↑	1.39 ↑	2.33 ↑	1.95 ↑	2.65 ↑
1072	gi|4929769	CGI-150 protein	2.07 ↓	1.23 ↓	1.88 ↓	1.60 ↓	2.61 ↓
1095	gi|16878077	FUBP1 protein	1.50 ↑	1.41 ↑	2.57 ↑	1.95 ↑	2.22 ↑
1112	gi|16924265	Enoyl Coenzyme A hydratase 1, peroxisomal	2.55 ↑	1.51 ↑	1.69 ↑	1.66 ↑	1.35 ↑
1124	gi|13027378	Glucosamine-6-phosphate deaminase 1	1.34 ↑	1.63 ↑	2.52 ↑	1.55 ↑	1.63 ↑
1181	gi|48255905	Transgelin	2.43 ↓	2.27 ↓	1.73 ↓	1.71 ↓	2.13 ↓
1194	gi|4504035	Guanine monophosphate synthetase	1.21 ↑	1.60 ↑	2.41 ↑	2.00 ↑	1.77 ↑
1209	gi|32129199	Cytokine induced protein 29 kDa	1.37 ↑	1.75 ↑	2.18 ↑	1.84 ↑	2.39 ↑
1240	gi|5031981	26S proteasome-associated pad1 homolog	2.19 ↑	1.28 ↑	1.04 ↓	1.04 ↑	1.07 ↓
1246	gi|54648253	KHSRP protein (FUBP2)	1.31 ↓	2.34 ↓	1.19 ↓	1.65 ↓	1.20 ↓
1257	gi|1575609	FUSE binding protein 3	2.32 ↓	1.59 ↓	1.12 ↓	1.85 ↓	2.32 ↓
1281	gi|5453595	Adenylyl cyclase-associated protein	1.61 ↓	2.26 ↓	2.01 ↓	2.29 ↓	1.92 ↓
1288	gi|184433	Histone-binding protein	1.70 ↑	2.29 ↑	1.86 ↑	1.11 ↑	1.79 ↑
1290	gi|546901	Cytosolic acetoacetyl-coenzyme A thiolase; CT	2.06 ↑	2.05 ↑	1.91 ↑	2.28 ↑	1.69 ↑
1296	gi|17402893	Phosphoserine aminotransferase isoform 1	1.53 ↑	2.28 ↑	1.84 ↑	1.45 ↑	1.50 ↑
1312	gi|4505763	Phosphoglycerate kinase 1	2.26 ↑	1.96 ↑	2.16 ↑	2.02 ↑	1.88 ↑
1320	gi|4758484	Glutathione-*S*-transferase omega 1	1.30 ↓	1.64 ↓	2.26 ↓	1.37 ↓	1.21 ↓
1332	g i|847724	Methylthioadenosine phosphorylase	1.65 ↑	2.25 ↑	2.19 ↑	2.17 ↑	1.76 ↑
1358	gi|10863927	Peptidylprolyl isomerase A	1.16 ↑	1.50 ↑	2.22 ↑	1.28 ↑	1.72 ↑
1376	gi|4504035	Guanine monophosphate synthetase	1.13 ↑	1.33 ↑	2.20 ↑	1.90 ↑	1.35 ↑
1392	gi|4557032	L-lactate dehydrogenase B	2.17 ↑	1.60 ↑	1.16 ↑	1.19 ↑	1.32 ↑
1411	gi|19743875	Fumarate hydratase precursor	2.15 ↑	2.12 ↑	1.49 ↑	1.26 ↑	1.56 ↑
1450	gi|4506181	Proteasome α2 subunit	1.33 ↑	2.11 ↑	1.65 ↑	2.03 ↑	1.38 ↑
1471	gi|4507645	Triosephosphate isomerase 1	2.08 ↓	1.60 ↓	1.52 ↓	1.73 ↓	1.82 ↓
1488	gi|4506195	Proteasome β2 subunit	1.59 ↑	1.85 ↑	2.06 ↑	1.67 ↑	2.05 ↑
1507	gi|6912586	6-phosphogluconolactonase	1.34 ↑	1.38 ↑	2.04 ↑	1.03 ↑	1.44 ↑
1530	gi|181969	elongation factor 2	1.94 ↑	2.02 ↑	1.98 ↑	2.01 ↑	1.71 ↑

hiPSC, human iPSC.

necessary to elucidate the involvement mechanism of Nups in reprogramming.

The protein SET (also known as protein phosphatase 2A inhibitor) was significantly increased in both hESCs and hiPSCs (Fig. 4 and Table 2). SET binds to histones as a subunit of the INHAT (inhibitor of acetyltransferase) complex [37]. The binding of the INHAT complex to histones suppresses their acetylation and thereby induces transcriptional repression. The overexpression of SET inhibits DNA demethylation, resulting in gene silencing [38]. Additionally, SET binds and blocks the DNase activity of NM23-H1, a tumor metastasis suppressor [38]. SET also plays a role in hematopoietic differentiation [39]. The dramatically increased levels of SET in hESCs and hiPSCs may reflect its role in reprogramming, although the details of role are unknown.

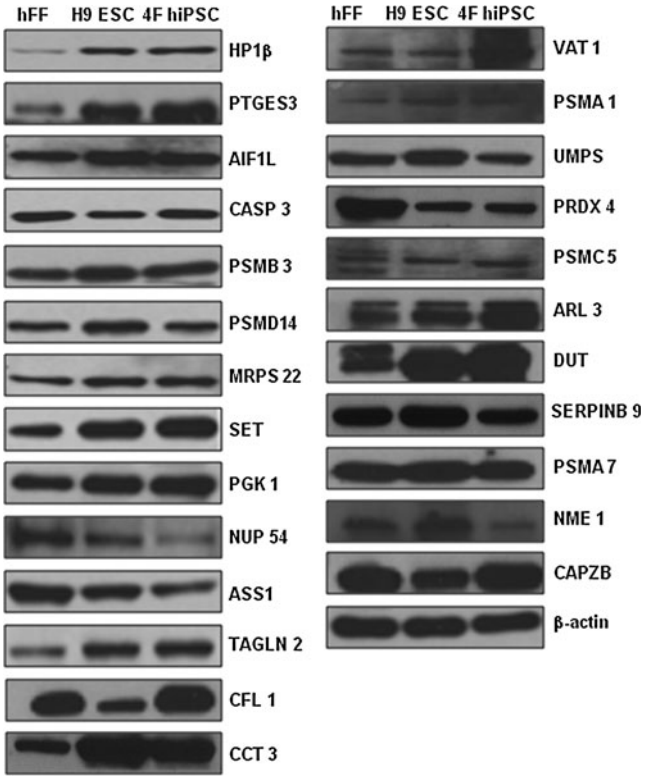


FIG. 4. Validation of 2-DE results by western blot analysis. Western blot confirmation of the expression patterns of heterochromatin protein 1β (HP1β), prostaglandin E synthase 3 (PTGES3), ionized calcium binding adapter molecule 2 isoform 1 (AIF1L), caspase 3 preproprotein (CASP3), proteasome beta 3 subunit (PSMB3), 26S proteasome-associated pad1 homolog (PSMD14), mitochondrial ribosomal protein S22 (MRPS22), SET, phosphoglycerate kinase 1 (PGK1), nucleoporin 54 kDa (NUP54), argininosuccinate synthetase 1 (ASS1), transgelin 2 (TAGLN2), cofilin 1 (CFL1), γ subunit of CCT chaperonin (CCT3), vesicle amine transport protein 1 (VAT1), proteasome α1 subunit isoform 2 (PSMA1), UMP synthase (UMPS), thioredoxin peroxidase (PRDX4), 26S proteasome subunit p45 (PSMC5), ADP-ribosylation factor-like 3 (ARL3), deoxyuridine 5′-triphosphate nucleotidohydrolase (DUT), serpin peptidase inhibitor, clade B, member 9 (SERPINB9), proteasome subunit HSPC (PSMA7), nonmetastatic cells 1 protein (NME1), and F-actin capping protein β subunit (CAPZB). β-Actin was employed as a loading control.

Protein network analysis of identified proteins

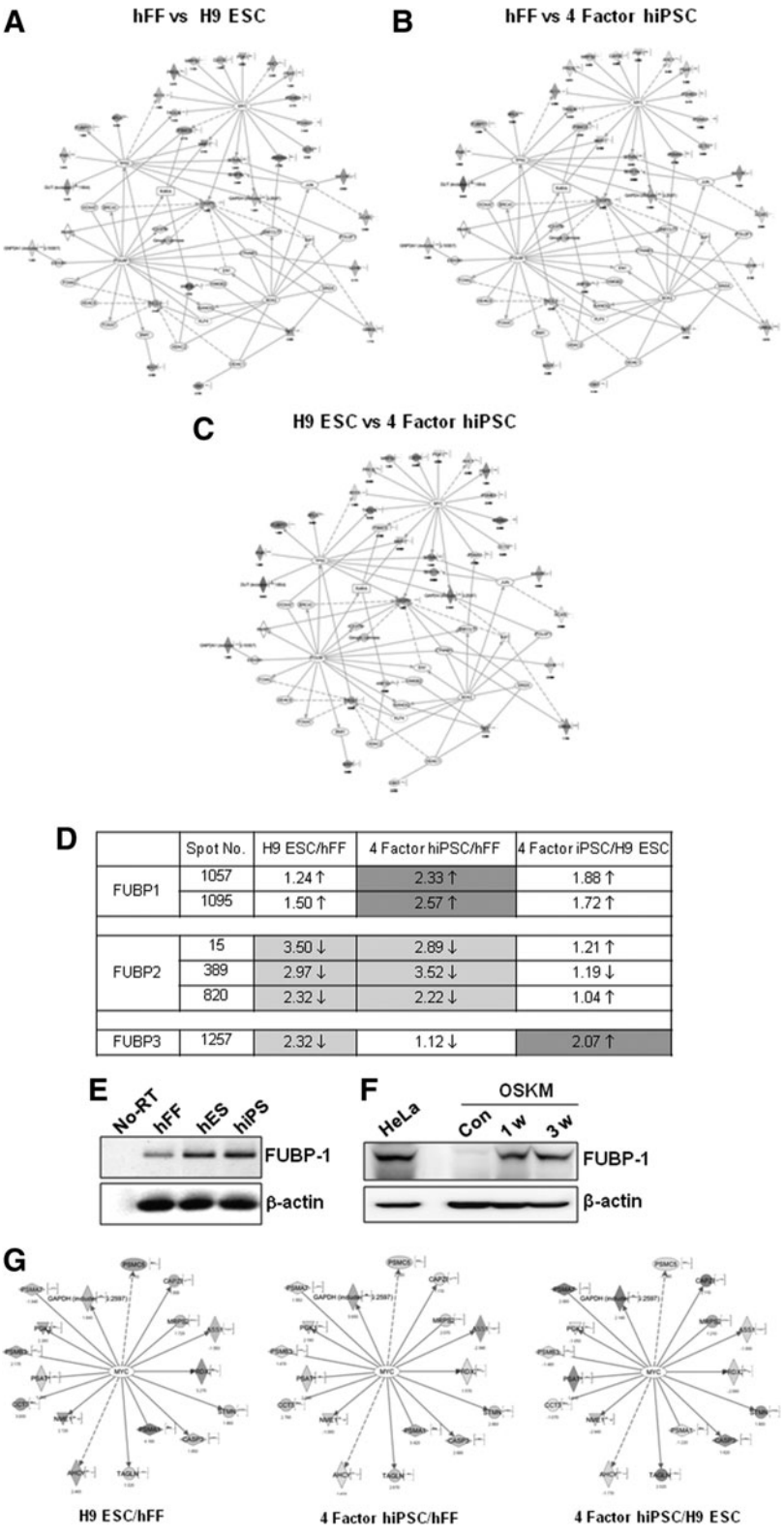
The proteins displaying patterns of differential expression among hFFS, hESCs, and OSKM-hiPSCs were analyzed in silico using the IPA software [40,41]. Data from the comparison of donor cells and hESCs are represented in Fig. 5A, and data from the comparison of donor cells and OSKM-hiPSCs are depicted in Fig. 5B. There are similarities between hESCs and OSKM-hiPSCs, but subtle differences were also detected (Fig. 5C), providing insight into the reprogramming process and differences between hESCs and OSKM-hiPSCs.

FUBPs bind to an upstream element of the *c-myc* promoter and regulate the level of *c-myc* mRNA [42,43]. The c-Myc protein is part of the basic helix-loop-helix leucine zipper family of transcription factors and is involved in cell growth, proliferation, differentiation, and apoptosis [42]. Further, c-Myc participates in the dedifferentiation process and is widely used as a key reprogramming factor [1–5]. Omission of c-Myc during reprogramming significantly lengthens the time required for reprogramming and dramatically decreases reprogramming efficiency [44]. However, c-Myc induces tumor formation in half of the chimeric mice generated from iPSCs [1,45]. Therefore, the alternative gene of *c-myc* is highly needed to overcome these obstacles of c-Myc during reprogramming process. In our proteomic analysis, 3 FUBP family protein members were differentially expressed (Table 1). FUBP-1 and FUBP-3 were significantly upregulated in hESCs compared with somatic cells. On the other hand, increased FUBP-1 levels and decreased FUBP-2 levels were detected in hiPSCs. Moreover, only FUBP-3 expression was significantly enhanced in hiPSCs compared with hESCs (Fig. 5D). The target genes of FUBPs are influenced by the absolute and relative intranuclear stoichiometries of individual FUBPs [46]. Therefore, it is estimated that absolute and relative amount of FUBPs may also be important in the regulation of c-Myc expression during the reprogramming process. In consistent, expression of FUBP-1 mRNAs was upregulated in pluripotent hESCs and hiPSCs compared with hFFs (Fig. 5E). Importantly, it was confirmed that the expression of FUBP-1 proteins was notably increased in hFFs at 1 and 3 weeks after OSKM transduction (Fig. 5F). The result from protein network analysis clearly showed differences in FUBPs among hFFs, hESCs, and OSKM-hiPSCs (Fig. 5G). Therefore, we suggest the possibility that FUBPs or FUBP-upregulating molecule(s) may be used as reprogramming factors.

Concluding Remarks

We performed a comparative proteome analysis of hESCs, hiPSCs, and their corresponding donor cells (hFFs). Through this approach, we identified many proteins that may be directly or indirectly involved in reprogramming. The identified proteins are involved in various biological processes, including transcription cofactor activity, proteasome activator activity, lipid metabolic processes, cell redox homeostasis, and nucleoside metabolic processes, indicating that

significant physiological changes occur during reprogramming. Further, we identified several proteins with differential expression patterns between hESCs and hiPSCs. In future studies, we will perform a detailed investigation of the roles of the identified proteins during reprogramming and examine whether they can be effectively utilized to induce or regulate reprogramming at will. In combination with our proteomic analyses, further characterization of these proteins should provide valuable new insights into the mechanism of reprogramming.



	Spot No.	H9 ESC/hFF	4 Factor hiPSC/hFF	4 Factor iPSC/H9 ESC
FUBP1	1057	1.24 ↑	2.33 ↑	1.88 ↑
	1095	1.50 ↑	2.57 ↑	1.72 ↑
FUBP2	15	3.50 ↓	2.89 ↓	1.21 ↑
	389	2.97 ↓	3.52 ↓	1.19 ↓
	820	2.32 ↓	2.22 ↓	1.04 ↑
FUBP3	1257	2.32 ↓	1.12 ↓	2.07 ↑

FIG. 5. Protein network analysis using Ingenuity Pathway Analysis (IPA) software. Proteins are represented as nodes, and the biological relationship between 2 nodes is represented as an edge (*line*). All edges are supported by at least 1 reference from published sources or from information stored in the Ingenuity Pathways Knowledge Base (IPKB). The node indicates up or down regulation. Nodes are displayed using several shapes that represent the functional classes of proteins. **(A)** Comparison between somatic donor cells (hFFs) and hESCs. **(B)** Comparison between somatic donor cells and 4 Factor hiPSCs (OSKM-hiPSCs). **(C)** Comparison between hESCs and 4 Factor hiPSCs (OSKM-hiPSCs). **(D)** Fold changes in FUBPs among somatic cells, ESCs, and iPSCs. **(E)** Confirmation of FUBP-1 mRNA level in hFFs, hESCs, and hiPSCs. **(F)** mRNA level of FUBP-1 at 1 and 3 weeks after OSKM transduction. **(G)** Network representation of molecular interactions, focused on the c-Myc protein.

Acknowledgments

The authors are grateful to Drs. Sang J. Chung, Jae Ho Kim, Sayeon Cho, Jeong-Ki Min, and Doo Byung Oh for their careful reading of the manuscript. The authors also thank Eun-Young Kim and Eonryung Kim for technical assistance and helpful discussions. This work was supported by the KRIBB/KRCF research program (NAP), the National Research Foundation (NRF) of Korea (No. 2010-0020305), and the KRIBB.

Author Disclosure Statement

No competing financial interests exist.

References

1. Takahashi K and S Yamanaka. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 126:663–676.
2. Winkler J, I Sotiriadou, S Chen, J Hescheler and A Sachinidis. (2009). The potential of embryonic stem cells combined with–omics technologies as model systems for toxicity. Curr Med Chem 16:4814–4827.
3. Jalving M and H Schepers. (2009). Induced pluripotent stem cells: will they be safe? Curr Opin Mol Ther 11:383–393.
4. Smith KP, MX Luong and GS Stein. (2009). Pluripotency: toward a gold standard for human ES and iPS cells. J Cell Physiol 220:21–29.
5. Yamanaka S. (2007). Strategies and new developments in the generation of patient-specific pluripotent stem cells. Cell Stem Cell 1:39–49.
6. Hochedlinger K and K Plath. (2009). Epigenetic reprogramming and induced pluripotency. Development 136:509–523.
7. Daley GQ. (2010). Stem cells: roadmap to the clinic. J Clin Invest 120:8–10.
8. Banito A and J Gil. (2010). Induced pluripotent stem cells and senescence: learning the biology to improve the technology. EMBO Rep 11:353–359.
9. Huangfu D, K Osafune, R Maehr, W Guo, A Eijkelenboom, S Chen, W Muhlestein and DA Melton. (2008). Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. Nature Biotechnol 26:1269–1275.
10. Park I-H, R Zhao, JA West, A Yabuuchi, H Huo, TA Ince, PH Lerou, MW Lensch and GQ Daley. (2008). Reprogramming of human somatic cells to pluripotency with defined factors. Nature 451:141–146.
11. Kim JB, B Greber, MJ Arauzo-Bravo, J Meyer, K-I Park, H Zaehres and HR Scholer. Direct reprogramming of human neural stem cells by OCT4. Nature 461:649–654.
12. Brambrink T, R Foreman, GG Welstead, CJ Lengner, M Wernig, H Suh and R Jaenisch. (2008). Sequential expression of pluripotent markers during direct reprogramming of mouse somatic cells. Cell Stem Cell 2:151–159.
13. Dihazi H, GH Dihazi, J Nolte, S Meyer, O Jahn, GA Muller and W Engel. (2009). Multipotent adult germline stem cells and embryonic stem cells: comparative proteomic approach. J Proteome Res 8:5497–5510.
14. Rigbolt KT, TA Prokhorova, V Akimov, J Henningsen, PT Johansen, I Kratchmarova, M Kassem, M Mann, JV Olsen and B Blaqoev. (2011). System-wide temporal characterization of the proteome and phosphoproteome of human embryonic stem cell differentiation. Sci Signal 4:rs3.
15. Wang S, R Tian, D Figeys and L Wang. (2011). An enhanced chemically defined SILAC culture system for quantitative proteomics study of human embryonic stem cells. Proteomics doi:10.1002/pmic.201100052.
16. Graumann J, NC Hubner, JB Kim, K Ko, M Moser, C Kumar, J Cox, H Scholer and M Mann. (2008). Stable isotope labeling by amino acids in cell culture (SILAC) and proteome quantitation of mouse embryonic stem cells to a depth of 5,111 proteins. Mol Cell Proteomics 7:672–683.
17. Kim MS, J Kim, HW Han, YS Cho, YM Han and JK Park. (2007). Microfabricated embryonic stem cell divider for large-scale propagation of human embryonic stem cells. Lab Chip 7:513–515.
18. Kang TH, K-H Bae, Mj Yu, WK Kim, H-R Hwang, H Jung, PY Lee, S Kang, T-S Yoon, et al. (2007). Phosphoproteomic analysis of neuronal cell death by glutamate-induced oxidative stress. Proteomics 7:2624–2635.
19. Kim S-Y, PY Lee, H-J Shin, DH Kim, S Kang, H-B Moon, SW Kang, J-M Kim, SG Park, et al. (2009). Proteomic analysis of liver tissue from *HBx*-transgenic mice at early stages of hepatocarcinogenesis. Proteomics 9:5056–5066.
20. Kang S, EY Kim, JW Chung, T-S Yoon, DH Lee, SG Park, BC Park and K-H Bae. (2007). Proteomic analysis of effect of MAPK pathway activation on L-glutamate-induced neuronal cell death. Cell Mol Biol Lett 12:139–147.
21. Collins GA and WP Tansey. (2006). The proteasome: a utility tool for transcription? Curr Opin Genet Dev 16:197–202.
22. Szutorisz H, A Georgiou, L Tora and N Dillon. (2006). The proteasome restricts permissive transcription at tissue-specific gene loci in embryonic stem cells. Cell 127:1357–1388.
23. Ayoub N, AD Jeyasekharan, JA Bernal and AR Venkitaraman. (2008). HP1-β mobilization promotes chromatin changes that initiate the DNA damage response. Nature 453: 682–686.
24. Rowe HM, J Jakobsson, D Mesnard, J Rougemont, S Reynard, T Aktas, PV Maillard, H Layard-Liesching, S Verp, et al. KAP1 controls endogenous retroviruses in embryonic stem cells. Nature 463:237–240.
25. Djuric U and J Ellis. (2010). Epigenetics of induced pluripotency, the seven-headed dragon. Stem Cell Res 1:3.
26. Hochedlinger K and K Plath. (2009). Epigenetic reprogramming and induced pluripotency. Development 136:509–523.
27. Yosida Y, K Takahashi, K Okita, T Ichisaka and S Yamanaka. (2009). Hypoxia enhances the generation of induced pluripotent stem cells. Cell Stem Cell 5:237–241.
28. Cho YM, S Kwon, YK Park, HW Seol, YM Choi, DJ Park, KS Park and HK Lee. (2006). Dynamic changes in mitochondrial biogenesis and antioxidant enzymes during the spontaneous differentiation of human embryonic stem cells. Biochem Biophys Res Comm 348:1472–1478.
29. Prigione A, B Fauler, R Lurz, H Lehrach and J Adjaye. (2010). The senescence-related mitochondrial/oxidative stress pathway is repressed in human induced pluripotent stem cells. Stem Cells 28:721–733.
30. Rehman J. (2010). Empowering self-renewal and differentiation: the role of mitochondria in stem cells. J Mol Med 88:981–986.
31. Hara MR, N Agrawal, SF Kim, MB Cascio, M Fujimuro, Y Ozeki, M Takahashi, JH Cheah, SK Tankou, et al. (2005). S-nitrosylated GAPDH initiates apoptotic cell death by nuclear translocation following Siah1 binding. Nat Cell Biol 7:665–674.
32. Tristan C, N Shahani, TW Sedlak and A Sawa. (2011). The diverse functions of GAPDH: Views from different subcellular compartments. Cell Signal 23:317–323.
33. Jang M, HJ Kang, SY Lee, SJ Chung, S Kang, S-W Chi, S Cho, SC Lee, C-K Lee, et al. (2009). Glycerladehyde-3-phosphate,

a glycolytic intermediate, plays a key role in controlling cell fate via inhibition of caspase activity. Mol Cells 28:559–563.
34. Sabri N, P Roth, F Xylourgidis, F Sadeghifar, J Adler and C Samakovlis. (2007). Distinct functions of the *Drosophila* Nup153 and Nup214 FG domains in nuclear protein transport. J Cell Biol 178:557–565.
35. Terry L and S Wente. (2007). Nuclear mRNA export requires specific FG nucleoporins for translocation through the nuclear pore complex. J Cell Biol 178:1121–1132.
36. Lupu F, A Alves, K Anderson, V Doye and E Lacy. (2008). Nuclear pore composition regulates neural stem/progenitor cell differentiation in the mouse embryo. Dev Cell 14: 831–842.
37. Karetsou Z, G Martic, G Sflomos and T Papamarcaki. (2005). The histone chaperone SET/TAF-Iβ interacts functionally with the CREB-binding protein. Biochem Biophys Res Comm 335:322–327.
38. Muto S, M Senda, Y Akai, L Sato, T Suzuki, R Nagai, T Senda and M Horikoshi. (2007). Relationship between the structure of SET/TAF-Iβ/INHAT and its histone chaperone activity. Proc Natl Acad Sci 104:4285–4290.
39. Kandilci A, E Mientjes and G Grosveld. (2004). Effects of SET and SET-CAN on the differentiation of the human promonocytic cell line U937. Leukemia 18:337–340.
40. Katagiri T, N Hatano, M Aihara, H Kawano, M Okamoto, Y Liu, T Izumi, T Maekawa, S Nakamura, et al. (2010). Proteomic analysis of proteins expressing in regions of rat brain by a combination of SDS-PAGE with nano-liquid chromatography-quadrupole-time of flight tandem mass spectrometry. Proteome Sci 8:41.
41. Ortiz PA, ME Bruno, T Moore, S Nesnow, W Winnik and Y Ge. (2010). Proteomic analysis of propiconazole responses in mouse liver: comparison of genomic and proteomic profiles. J Proteome Res 9:1268–1278.
42. Wierstra I and J Alves. (2008). The c-myc promoter: still MysterY and challenge. Adv. Cancer Res 99:113–333.
43. Jang M, BC Park, S Kang, S-W Chi, Cho S, SJ Chung, SC Lee, K-H Bae and SG Park. (2009). Far upstream element-binding protein-1, a novel caspase substrate, acts as a cross-talker between apoptosis and the *c-myc* oncogene. Oncogene 28: 1529–1536.
44. Wernig M, A Meissner, JP Cassady and R Jaenisch. (2008). c-Myc is dispensable for direct reprogramming of mouse fibroblasts. Cell Stem Cell 2:10–12.
45. Okita K, M Nakagawa, H Hyenjong, T Ichisaka and S Yamanaka. (2008). Generation of mouse induced pluripotent stem cells without viral vectors. Science 322: 949–953.
46. Chung H-J, J Liu, M Dundr, Z Nie, S Sanford and D Levens. (2006). FBPs are calibrated molecular tools to adjust gene expression. Mol Cell Biol 26:6584–6597.

Address correspondence to:
Dr. Kwang-Hee Bae
Medical Proteomics Research Center
KRIBB
50 Eon-dong
Yusung-gu
Daejeon 305-806
South Korea

E-mail: khbae@kribb.re.kr

Dr. Yee Sook Cho
Development and Differentiation Research Center
KRIBB
50 Eon-dong, Yusung-gu
Daejeon 305-806
South Korea

E-mail: june@kribb.re.kr

Received for publication May 12, 2011
Accepted after revision July 25, 2011
Prepublished on Liebert Instant Online July 25, 2011