Human Induced Pluripotent Stem Cells Derived Under Feeder-Free Conditions Display Unique Cell Cycle and DNA Replication Gene Profiles

Henry C.Y. Chung,¹ Ruby C.Y. Lin,² Grant J. Logan,³ Ian E. Alexander,^{3,4}
Perminder S. Sachdev,¹ and Kuldip S. Sidhu¹

Use of animal feeder layers and serum containing media in the derivation and propagation of induced pluripotent stem cells (iPSCs) can hinder clinical translation, because of the presence of xeno-material/pathogens. A defined and standardized system would be ideal for generating a homogenous population of iPSCs, which closely resembles human embryonic stem cells (hESCs). This article presents a novel and extensive comparison between in-house produced iPSCs and hESCs under “feeder” and “feeder-free” conditions, using transcriptomic genome-wide microarray analysis. We generated a list of pluripotency-associated and bivalent domain-containing genes by meta-analysis to measure qualitatively the degree of reprogramming in feeder-free derived iPSCs, in which both profiles displayed similar levels of gene expression as in hESCs. Gene ontology analysis showed that feeder-free iPSCs have enriched terms belonging to DNA repair/replication and cell cycle, which are signature to pluripotent cells. Transcriptomic data combined with directed differentiation assays, indicated that variability among iPSC lines is minimized when using a feeder-free cultural system, which may serve as a platform for further developing regenerative medicine compliant human iPSCs.

Introduction

SINCE THE FIRST DERIVATION of mouse induced pluripotent stem cells (iPSCs) [1], several studies have derived iPSCs from human somatic cells [2,3] and by lentiviral transductions [4]. Increased efficiency of iPSC production with the aid of small molecules [5–8] and by using polycistronic vectors [9,10] has also been achieved; as has the elimination of transgene expression [11–13], and the utilization of protein transduction methods to avoid the use of genome-integrating viruses [14,15].

Several safety obstacles need to be addressed before iPSCs transits to the clinic; in particular, the risk of insertional mutagenesis when using integrative viruses and the transmission of pathogens when in media and/or feeder layers containing animal products. While transgene/viral-free methods have been developed, a majority of iPSCs are still derived on animal feeder layers, which offsets the benefits of a xeno-free autologous transplantation. The use of animal feeder layers also introduces inconsistent and variable reprogramming outcomes, making the screening of suitable,

fully reprogrammed iPSC colonies labor intensive [16]. This is reflected in iPSC lines generated under different experimental conditions and show varying degrees of differentiation into hematopoietic and neural lineages [17,18], which also limits its therapeutic potential. The use of defined media however, minimizes variable and/or inhibitory components present in serum and growth-factor secreting feeder cells.

Consequently, some laboratories have generated iPSCs under feeder-free conditions by using extracellular matrices and serum-free media [19,20]. While the phenotypic outcomes have been tested, transcriptomic characterization of feeder-free derived iPSCs has not been fully explored. The advent of whole transcript gene expression microarrays was capitalized on to identify underlying molecular events that may underpin the differences between feeder derived, feeder-free derived iPSCs and human embryonic stem cells (hESCs) propagated under identical conditions.

For the first time it is reported that feeder-free iPSCs (ff-iPSCs) resemble hESCs more than feeder derived iPSCs (f-iPSCs) in terms of overall gene expression patterns governing pluripotency and other biological functions. The

¹Stem Cell Laboratory, School of Psychiatry, University of New South Wales, Sydney, Australia.

²Ramaciotti Centre for Gene Function Analysis, School of Biotechnology and Biomolecular Sciences, University of New South Wales, Sydney, Australia.

³Gene Therapy Research Unit, Children's Medical Research Institute and the Children's Hospital, Sydney, Australia.

⁴Discipline of Pediatrics and Child Health, University of Sydney, Sydney, Australia.

processes related to pluripotent signature in hESCs (ie, DNA replication and cell cycle) were substantially enriched in ff-iPSCs and expression of bivalent genes was lower. The advantages of a feeder-free defined system are such that homogenous populations of patient-specific pluripotent stem cells can be generated, batch-to-batch differences created by serum and feeder-cells can be eliminated and scale-up cultures can be easily carried out. This is of particular interest in regenerative medicine.

Materials and Methods

Ethics statement

This study was conducted in accordance with the University of New South Wales (UNSW), Human Research Ethics Committee (HREC) approvals, that is, the use of human fetal fibroblasts (hFFs) (HREC 08021), hESCs and the derivation and use of human iPSCs (HREC 08037).

Generation and propagation of feeder/serum and feeder/serum-free derived iPSCs

hFFs were maintained in Dulbecco's modified Eagle's medium supplemented with 10% FBS, 2 mM Glutamax, 25 U/mL Penicillin/streptomycin at 37°C, 5% CO₂ in a humidified incubator. hESC lines—Endeavour-1 and 2 (E1 and E2 respectively) [21,22] were maintained in either Knock-Out™ Dulbecco's modified Eagle's medium supplemented with 20% KnockOut™ Serum replacement, 0.1 mM nonessential amino acids, 0.1 mM β-mercaptoethanol, 2 mM Glutamax, 25 U/mL Penicillin/streptomycin, 1×Insulin-transferrin-selenium (later referred to as KO-SR) or under feeder-free conditions on Matrigel™ (BD Biosciences) coated vessels in mTeSR™1 media (Stemcell Technologies).

Cells were reprogrammed using Lenti-EF1α-STEMCCA [10], a polycistronic vector encoding *OCT4*, *KLF4*, *SOX2* and *CMYC* (kind gift from Gustavo Mostoslavsky, Boston University), which was packaged and titred using established methodologies [23]. For f-iPSC generation, 100,000 hFFs were seeded overnight onto a single well of a 6-well plate (Griener). Cells were transduced with the virus at a multiplicity of infection of 45, in fresh fibroblast media supplemented with 5 μg/mL polybrene (Sigma-Aldrich). Alternate day media changes were performed until day 5 post-infection. All cells were harvested using TrypLE™ Select and reseeded onto 0.1% gelatin (Sigma-Aldrich) coated, irradiated hFF feeder layers in KO-SR supplemented with 10 ng/mL fibroblast growth factor 2 (FGF2). For ff-iPSC generation, 100,000 hFFs were seeded onto Matrigel™ coated 60 mm dishes and transduced as above. Media was changed to mTeSR™1 on day 6 post-infection. Colonies resembling hESCs were physically dissected under a microscope, into feeder or feeder-free conditions respectively and propagated as per standard hESC procedures.

Both feeder hESC and iPSC lines were continuously passaged using 1:4 diluted TrypLE™ Select on a 5–7 day basis and cultured in KO-SR media supplemented with 4 ng/mL FGF2. Feeder-free cultures were passaged using 1 mg/mL dispase on a 5–7 day basis and cultured on Matrigel™ coated surfaces in mTeSR™1 media.

All reagents were purchased from Life Technologies unless otherwise stated.

In vitro differentiation

Embryoid bodies (EBs) from hESCs and iPSCs were generated from feeder and feeder-free cultures using 1 mg/mL Collagenase IV or 0.5 mg/mL dispase solutions respectively, in non-tissue culture adherent plates (BD Biosciences) for 7–9 days before attachment to gelatin or Matrigel™ coated plates, respectively, for an additional 6–7 days in KO-SR media without FGF2 supplement.

In vivo differentiation—teratoma assays

Both f-iPSCs and ff-iPSCs were injected intramuscularly into hind legs of SCID mice. Teratomas were removed 6–8 weeks after injection and stained with Hematoxylin and Eosin as per standard procedures.

In vitro differentiation—direct differentiation to dopaminergic neurons

EBs were generated as above and cultured for 5–7 days in KO-SR and mTeSR™1 media for f-iPSCs and ff-iPSCs respectively. EBs were subsequently seeded onto PA6 cells grown as monolayers in 0.1% gelatin-coated plates for co-culture in GMEM + 20% SR, supplemented with 2 mM Glutamax, 0.1 mM nonessential amino acids, 1 mM sodium pyruvate for a period of 28 days with partial media changes performed every other day [24].

Karyotyping

Metaphase cells were prepared by colcemid treatment for 3–4 h and a single cell suspension was prepared. Cell suspension was incubated with 37.5 mM KCl for 30 min at 37°C and fixed in cold methanol and acetic acid (3:1 vol/vol). Karyotyping analysis was performed at the Cytogenetics lab (Prince of Wales Hospital, Sydney, Australia).

Gene expression

Total RNA was extracted using TRIzol® and 1 μg was reverse transcribed to cDNA using SuperScript® III kit with random hexamers. Genes of interest were quantitatively detected using SensiMix Plus SYBR master mix (Quantace) and relative expression calculated using Pfaffl's methods [25]. A list of primers can be found in Supplementary Table S1 (Supplementary Data are available online at www.liebertonline.com/scd). All reagents were purchased from Life Technologies following manufacturer's instructions unless stated otherwise. Biological triplicates were used.

Immunocytochemistry staining

Cell samples were fixed in 4% paraformaldehyde for 20 min, permeabilized with 0.1% Triton X-100/phosphate-buffered saline for 10 min and blocked with 2% bovine serum albumin/phosphate-buffered saline for 1 h. Cell samples were stained with appropriate primary antibodies in blocking solution overnight at 4°C and AlexaFluor® conjugated secondary antibodies (Life Technologies) for 30 min at room temperature. Cells were counterstained with DAPI (Life Technologies). Primary antibodies included AFP, OCT4, NANOG and SOX17 (R&D Systems), DESMIN, SSEA4, TRA160 and SMA (Abcam), βIII tubulin (Covance), NESTIN (Millipore).

Whole genome transcriptomic analysis

One hundred nanogram of total RNA from hESCs (passage 20–30) and iPSCs (passage 10–15) were labeled and hybridized onto Affymetrix Human GeneChip® Gene 1.0 ST arrays according to the manufacturer's instructions at the Ramaciotti Centre (UNSW, Australia). Two to three biological replicates were used for each cell line. RMA normalized data [26] was subjected to 1-way analysis of variance to look for differential gene expression between cell lines (Partek Genome Suite v6.5). A False Discovery Rate of 0.05 was used. Candidate genes were functionally annotated according to Gene Ontology (GO) terms [27] and were visualized in pathways using KEGG [28] and DAVID [29]. Gene lists were generated based on a cutoff of greater than 1.5-fold change.

The transcriptomic distance ratio was calculated by dividing the number of differentially expressed genes between 2 cell types over the total amount of differentially expressed genes between hFF/iPSC and iPSC/hESC [30].

Hierarchical clustering was carried out to look for possible co-regulation of significant differentially expressed candidate genes based on Euclidean and complete linkage distances using MeV 4.6.2 [31]. Further analysis to highlight specific pathways was carried out by obtaining gene lists from SABiosciences-curated pathways. Gene lists were then annotated against the array data and depicted as box-whisker plots.

Statistical analysis

Statistical analysis was performed in GraphPad Prism 5.0. Relative gene expression was analyzed using Student's *t*-test. Comparison of array data with SABiosciences-curated gene lists was carried out using 1-way analysis of variance.

Accession number

Array data have been deposited at the Gene Expression Omnibus database (www.ncbi.nlm.nih.gov/geo/), with the accession number, GSE21655.

Results

Initial characterization of human iPSCs derived under feeder and feeder-free conditions

Human iPSCs were generated from fetal fibroblasts under feeder and feeder-free conditions via transduction with a polycistronic virus encoding *OCT4*, *KLF4*, *SOX2* and *CMYC*. Colonies were visible as early as 9 days post-infection and physically dissected by 17–25 days and expanded under standard hESC protocols. iPSCs produced under both conditions displayed similar morphologies to hESCs. Feeder derived iPSCs and feeder-free derived iPSCs are subsequently denoted as f-iPSCs and ff-iPSCs, respectively. Both f-iPSCs and ff-iPSCs expressed markers for pluripotent genes and proteins (Fig. 1A–C). Interestingly, *KLF4* and *CMYC* genes were significantly downregulated in ff-iPSCs compared to hESCs ($P < 0.0005$ and 0.05 respectively).

To establish their capacity to differentiate, EBs were generated allowing them to spontaneously differentiate. Representative immunocytochemistry staining for early differentiation markers was used to confirm in vitro pluripotency abilities (Fig. 2A, B). Representative in vivo teratoma assays also showed tissues derived from all 3 germ lineages (Fig. 2C). In

addition, DNA fingerprinting using short tandem repeat analysis revealed that both f-iPSCs and ff-iPSCs were derived from parental hFFs (data not shown). Further, karyotype analysis for ff-iPSCs at passage 20 was conducted, which showed no chromosomal abnormalities (Supplementary Fig. S1).

Basic pluripotency characterization revealed no distinctive morphological differences between f-iPSCs and ff-iPSCs. Underlying global gene expression patterns between these cell types were then examined.

ff-iPSCs are more similar to hESCs based on their global gene expression signature

Microarray comparisons were performed with a total of 4 iPSC lines (iPS52, iPS60, iPS62, iPS65) that were either derived/maintained under feeder or feeder-free conditions and 2 hESC lines (E1, E2), see also Table 1. To determine the degree of reprogramming in these iPSCs, the gene expression signatures were compared to those of hESCs grown under similar conditions. Scatterplots of global gene expression data of 33,297 transcripts indicated that the strongest correlation with hESCs was ff-iPSCs, followed by f-iPSCs, whereas hFFs showed the least correlation with pluripotent cells (Fig. 3A).

Next the differentially expressed genes were examined between all cell lines using a criterion of having a false discovery rate-corrected *P* value of < 0.05 and acquiring a > 1.5 -fold change difference (Fig. 3B). As expected, fibroblasts exhibited the most number of differentially expressed genes to pluripotent cell types, whereas amongst iPSC and hESC lines, this number was particularly low, indicating a global similarity in the transcriptome of pluripotent cells. To quantitatively set a measure of similarity, a previously determined distance metric [30] was calculated based on the proportion of differentially expressed genes between fibroblasts/iPSCs or iPSCs/hESCs. The lower the ratio, the closer the 2 cell types and vice versa. The ratios for ff-iPSCs: hESCs were 0.01 and 0.02 and f-iPSCs: hESCs were 0.47 and 0.13 (Fig. 3C).

Pluripotency-associated genes used to measure the degree of reprogramming

Two gene lists from meta-analysis studies were combined to examine the expression of pluripotency-associated genes among iPSCs derived and propagated under feeder and feeder-free conditions. Briefly, the PluriNet gene list was collated from a range of multipotent and pluripotent cells that shared the same pluripotent protein-protein interactions [32]. The consensus hESC gene list comprised of genes that were overexpressed in a large cohort of ~ 50 hESCs lines [33]. The combined list included well-documented pluripotency-related genes; *DNMT3B*, *CMYC*, *NANOG*, *OCT4*, *TDGF1*, *ZFP42* (mouse homologue of *REX1*, a full list of genes can be found in Supplementary Table S2).

The expression profiles of this pluripotency list was represented as a heatmap (Fig. 4A), however only 276/299 genes of the PluriNet and 782/1,076 genes of the consensus hESC list were mapped for which expression values were known. Hierarchical clustering indicates that ff-iPSCs (ie, iPS60, iPS65) clustered closely with hESCs whereas f-iPSCs (ie, iPS52, iPS62) were clustered separately from hFFs, hESCs and ff-iPSCs.

The pluripotency-associated genes were significantly enriched for GO terms; DNA replication, DNA repair, cell cycle

FIG. 1. Pluripotent characterization of iPSCs derived under feeder and feeder-free conditions. **(A)** Pluripotency-associated gene expression profiles of f-iPSCs and ff-iPSCs in comparison to control hESCs. Student's *t*-test was used to determine statistical significance in comparison to hESCs, where $n=3$ and data presented as mean $\pm$ standard error of the mean. * $P < 0.05$, *** $P < 0.0005$. Immunocytochemistry staining of pluripotent nuclear and surface markers in **(B)** f-iPSCs with spindle-like shaped colonies and **(C)** ff-iPSCs with circular shaped colonies. Scale bars = 100 μ m. iPSCs, induced pluripotent stem cells; hESCs, human embryonic stem cells; f-iPSCs, feeder derived iPSCs; ff-iPSCs, feeder-free iPSCs. Color images available online at www.liebertonline.com/scd

(Fig. 4B). The expression profiles of these gene lists were quantitatively expressed as box-whisker graphs to demonstrate similarity across all iPSC lines relatively to hESCs (Fig. 4C). Interestingly, the expression profiles of iPS60, iPS62 and iPS65 were visually similar, with noticeable variability in iPS62

whereas iPS52 had a similar profile to hFF3 (data not shown). With respect to pluripotent gene expression, all iPSC lines were shown to be statistically different, except for iPS60 and iPS65.

Furthermore, 630 genes associated with lineage commitment (ie, Consensus differentiation gene list [33]) were also

FIG. 2. Differentiation of f-iPSCs and ff-iPSCs. Representative immunocytochemistry results for early differentiation markers; β III tubulin, NESTIN—ectoderm, Smooth Muscle Actin (SMA), DESMIN—mesoderm and α -feto protein (AFP), SOX17—endoderm at (A) low magnification (Scale bars=100 μ m) and (B) high magnification (Scale bars=25 μ m). (C) Representative immunohistochemistry Hematoxylin and Eosin staining after teratoma formation in SCID mice. Tissue consisted of neural rosettes (ectoderm), cartilage (mesoderm) and gut-like structures (endoderm). Images were obtained from the ff-iPSC line. Scale bars=25 μ m. Color images available online at www.liebertonline.com/scd

analyzed amongst all iPSC lines. f-iPSCs exhibited a higher and more significant expression of lineage commitment compared to ff-iPSCs, as indicated by higher mean and median values (Supplementary Fig. S2).

TABLE 1. CLASSIFICATION OF CELL LINES

Cell line	Cell type	Derivation/Maintenance type
E1	hESC	Feeder-free
E2/E2-f	hESC	Feeder-free/Feeder
iPS52	iPSC	Feeder
iPS60	iPSC	Feeder-free
iPS62	iPSC	Feeder
iPS65	iPSC	Feeder-free

Each cell line was derived and maintained under feeder or feeder-free conditions.

hESC, human embryonic stem cell; iPSC, induced pluripotent stem cell.

A distinctive set of biomarkers for partially reprogrammed cells had been previously characterized [34] and this gene set was utilized to characterize our iPSCs. Genes upregulated in partially reprogrammed iPSCs including *CDH13*, *CDKN1A*, *CDKN2A*, *FGF4* and *FOXD1* were all expressed at higher levels in f-iPSCs (Supplementary Fig. S3). Interestingly, f-iPSCs also stained positive for OCT4, NA-NOG and TRA160 (Fig. 1B), which have been suggested to be *bone fide* iPSC markers [16].

Comparing differential transcriptomic profiles of iPSCs to hESCs

There were a total of ~1000–3000 and ~120–160 genes differentially expressed in f-iPSCs and ff-iPSCs when compared to hESCs, respectively (Fig. 3B). On analyzing the functions of these genes, it was found that GO terms,

FIG. 3. Global gene profiling of iPSCs derived under feeder and feeder-free conditions. **(A)** Expression profiling of ~33,000 transcripts, represented as mean signal intensity values of corresponding cells. The global correlation coefficients (r) are also listed. **(B)** We obtained the numbers of differentially expressed genes by 1-way analysis of variance and filtering through expression profiles that had a false discovery rate-corrected P value < 0.05 and acquiring a fold change > 1.5 . **(C)** Relative genetic distances between parental fibroblasts, resulting iPSC lines (iPS52, iPS62 were feeder derived and iPS60, iPS65 were feeder-free derived) and hESCs (E1, E2). The lower the ratio, the closer the genetic distance and vice versa. hFF, human fetal fibroblast.

biological adhesion and developmental process were highly enriched in the (f-iPSC vs. hESC) category (represented by an enrichment score > 3 , which is equivalent to a significant P value < 0.05) (Table 2). In particular, the observation was that different biological process functions even among f-iPSCs - iPS52 had differential genes highly enriched for cellular processes, such as apoptosis and cell proliferation while iPS62 differential genes were enriched for system development. In contrast, biological adhesion and developmental process terms were either not significant or not highly enriched in ff-iPSCs as compared to f-iPSCs (Table 2).

Pluripotency-associated differences are present between f-iPSCs and ff-iPSCs

On comparing (f-iPSC vs. ff-iPSC), cell cycle and DNA replication GO terms were significantly enriched (Fig. 5A). Interestingly, these GO terms were also highly enriched in our established pluripotency-associated gene list. This suggests

that there may be a partial difference between f-iPSCs and ff-iPSCs in terms of pluripotency transcriptome. In contrast, the same terms were not significantly different in (f-hESC vs. ff-hESC), suggesting that hESCs cultured under feeder or feeder-free conditions do not distinguish its pluripotency (Fig. 5B).

There were 1209 genes that were common between these 2 gene sets and they were highly enriched for GO terms; extracellular matrix organization, cell-substrate adhesion and cell adhesion (data not shown). This falls in line with the cultural system differences that would be expected from a feeder and feeder-free system, that is, (f-iPSC vs. ff-iPSC) and (f-hESC vs. ff-hESC).

DNA replication and cell cycle gene profiles distinguish f-iPSCs and ff-iPSCs

DNA metabolic and cell cycle processes GO terms were significantly enriched in the following gene lists: (hFF vs. f-iPSC), (hFF vs. ff-iPSC), (hFF vs. f-hESC), (hFF vs. ff-hESC),

FIG. 4. Expression of pluripotency-associated genes. **(A)** Heatmap representation of 951 pluripotency-associated genes across all cell lines. ff-iPSC lines (iPS60 and iPS65) were tightly clustered to hESCs (E1, E2), whereas f-iPSCs (iPS52 and iPS62) clustered separately (Euclidean and complete linkage parameters were used). **(B)** GO enrichment analysis of these 951 genes reveal they are highly related to DNA metabolic processes and mostly occur within the nucleus. **(C)** Quantitative fold changes of these genes in all iPSC lines relatively to hESCs. Only ff-iPSCs were not significantly (n/s) different. "+" depicts the mean value. GO, gene ontology. Color images available online at www.liebertonline.com/scd

FIG. 5. Differentially expressed genes of iPSCs and hESCs under different culture conditions. **(A)** A total of 854 genes were differentially expressed in (f-iPSC vs. ff-iPSC) and **(B)** 1196 genes in (f-hESC vs. ff-hESC). GO enrichment analysis was performed in Partek GS and revealed that cell cycle process, DNA replication and DNA repair were significantly different in f-iPSCs and ff-iPSCs but not between f-hESCs and ff-hESCs. Significance is indicated by an enrichment score >3, which is equivalent to a P value <0.05. Color images available online at www.liebertonline.com/scd

TABLE 2. GENE ONTOLOGY ENRICHMENT ANALYSIS OF FEEDER DERIVED INDUCED PLURIPOTENT STEM CELLS AND FEEDER-FREE-INDUCED PLURIPOTENT STEM CELLS AGAINST HUMAN EMBRYONIC STEM CELLS

Cell line	Function	Enrichment score
iPS52	Biological adhesion	11
	Developmental process	7.8
	Cellular component organization	3.1
	Metabolic process	1.5
iPS62	Developmental process	14
	Biological adhesion	9.9
	Metabolic process	3.7
	Cellular component organization	0.9
iPS60	Developmental process	2.8
	Biological adhesion	2.7
	Metabolic process	0.5
	Cellular component organization	0.2
iPS65	Biological adhesion	4.1
	Cellular component organization	3.3
	Metabolic process	2
	Developmental process	0.38

Broad gene ontology biological process terms. An enrichment score of 3 is equivalent to a significant P -value <0.05.

(f-iPSC vs. ff-iPSC) and also in the pluripotency-associated genes (Figs. 4B and 5A, Supplementary Fig. S4). Therefore, the expression profiles of cell cycle, DNA damage and DNA repair in f-iPSCs and ff-iPSCs were examined. ff-iPSCs were not significantly different to each other (ie, iPS60 and iPS65), whereas f-iPSCs were significantly different to ff-iPSCs and to each other (ie, iPS52 and iPS62) across the 3 profiles (Fig. 6).

Expression profile of bivalent gene also distinguishes f-iPSCs and ff-iPSCs

Recently, a study demonstrated that the expression profile of bivalent domain containing developmental genes can be a good indicator of iPSC quality [35]. Similarly a list was generated consisting of 1,028 bivalent genes from 3 independent studies [36–38] to characterize our f-iPSCs and ff-iPSCs (Fig. 7A, Supplementary Table S2). These genes are highly enriched for developmental processes coinciding with bivalent developmental genes (Fig. 7B). Interestingly, hierarchical clustering closely groups ff-iPSCs with hESCs, whereas f-iPSCs were more closely clustered with hFFs, indicating that some developmental genes are over-expressed in f-iPSCs (Fig. 7C, D).

FIG. 6. Expression profiles of pluripotency-associated biological functions. Box-whisker plots of the expression profile of f-iPSCs and ff-iPSCs (expression is relative to hESCs). Significance was calculated by not assuming Gaussian distributions using the Friedman test. * $P < 0.05$, ** P value < 0.005 , *** P value < 0.0005 .

Directed differentiation of iPSCs to dopaminergic neurons

Spontaneously differentiating f-iPSCs and ff-iPSCs did not show any dissimilarity, so hESCs, f-iPSCs and ff-iPSCs were

directly differentiated to dopaminergic neurons by co-culturing with stromal-cell derived inducing activity PA6 cells [24]. Interestingly, at the end of the differentiation, neurite-like structures were only observed in hESC and ff-iPSC cultures and in some instances (Fig. 8A), rosette-like structures were only observed in differentiating f-iPSC cultures (data not shown). Lastly, only differentiated hESCs and ff-iPSCs stained positive for TUJ1 (Fig. 8B).

Gene expression analysis between undifferentiated (U) and differentiated (D) cells displayed some interesting patterns (Fig. 8C). Firstly, *OCT4* and *NANOG* were significantly downregulated in hESCs and ff-iPSCs, consistent with differentiating pluripotent stem cells. Early neuroectoderm markers *SOX1* and *PAX6* demonstrated an upregulation trend among all differentiated cells with *PAX6* significantly different in f-iPSCs. Other late dopaminergic markers demonstrated an upward trend in gene expression (*PITX3* and *TH*) among all differentiated cells with the exception of *NURR1* being significantly downregulated in f-iPSCs.

FIG. 7. f-iPSC/ff-iPSC expression profiles of bivalent genes. (A) Merging of bivalent domain containing gene lists from 3 independent studies. (B) GO enrichment analysis of the 1,028 genes. (C) Box-whisker plots of the expression profile in parental fibroblasts, f-iPSCs and ff-iPSCs. Fold changes were relative to hESCs. (D) Hierarchical clustering of bivalent genes across all cell lines (Euclidean and complete linkage parameters were used). Color images available online at www.liebertonline.com/scd

Discussion

iPSCs have been generated from a diverse background of somatic cells using a number of transcription-factor delivery methods. Although these cell lines are coined iPSCs, they vary between laboratories and exhibit a persistent donor cell signature [30]. Thus, it is of great interest to generate iPSCs in a defined system to minimize variability. Here, we performed an extensive transcriptomic analysis of iPSCs derived and maintained under feeder and feeder-free conditions and subsequently identified distinctive gene patterns. The GO terms analyzed serve as distinguishing features for iPSCs and/or hESCs. this study demonstrated that ff-iPSCs display the closest correlation with hESCs.

The first piece of evidence displaying iPSC-hESC similarity was the distance metric calculated using proportionate differentially expressed genes. This only provided a global measurement of the degree of reprogramming. However, GO enrichment analysis revealed more insight into these genes. A novel "pluripotency-associated gene list" via meta analysis was established. Genes from this list had the most

FIG. 8. Directed differentiation into dopaminergic neurons. **(A)** Morphological analysis at the end of the differentiation period. **(B)** Immunocytochemistry of TUJ1 neuronal marker in hESCs, f-iPSCs and ff-iPSCs (DAPI inset, scale bars = 200 μ m). **(C)** Gene expression profiles of hESCs, f-iPSCs and ff-iPSCs from the undifferentiated (U) state to their differentiated (D) state. Genes include; pluripotent (*OCT4*, *NANOG*), neuroectoderm (*SOX1*, *PAX6*), early mid brain (*NURR1*), late mid brain (*PITX3*) and dopaminergic markers (*DAT*, *TH*). * $P < 0.05$, ** $P < 0.005$, *** $P < 0.0005$. Color images available online at www.liebertonline.com/scd

over-represented GO terms in biological processes of DNA replication, DNA repair and cell cycle process, which are processes involved in cellular reprogramming. Coincidentally, several of these terms were also differentially expressed in (f-iPSC vs. ff-iPSC). This was confirmed because the same comparison with hESCs did not reveal the same signature. Furthermore, expression profiles of cell cycle-, DNA damage- and DNA repair-related genes revealed significant differences between the profiles of f-iPSCs, ff-iPSCs and even amongst f-iPSCs, but not between ff-iPSCs. Additionally, the higher levels of cell cycle expression in ff-iPSCs may be a direct result of media used for propagation, as previously demonstrated that feeder-free culture in mTeSRTM1 media exhibited a shorter doubling time [39]. Overall, this demonstrated that iPSCs derived and maintained under feeder conditions exhibit variability, which may be the cause of the differentiation variability [17,18] or transcriptomic variability [40], reported in other studies.

Developmental genes are normally silenced (or expressed at low levels) in the undifferentiated state of hESCs but are

poised for activation [36,37]. A recent study claims that the expression profile of these bivalent genes can be a good indicator of reprogramming quality [35]. Here, the demonstration shows the bivalent gene profile of ff-iPSCs closely resembles that of hESCs, indicating superiority in a defined, feeder-free system. Interestingly, f-iPSCs had a higher expression level of partially reprogrammed genes, but simultaneously, stained positive for TRA160, a proposed true iPSC marker [16]. Though contradicting, this may indicate the existence of multiple intermediate/pluripotent states, different from previously published studies [1,34].

To further ascertain the differences in f-iPSCs and ff-iPSCs, directed differentiation of these cells to dopaminergic neurons was carried out. Interestingly, ff-iPSCs closely resembled hESC in specification of dopaminergic neurons as evident by morphological and immunocytochemistry analyses. The probable explanation for this is the improper downregulation of master pluripotent regulators, *OCT4* and *NANOG*, in differentiating f-iPSCs. This was also supported by the significant difference in gene expression of early

neural differentiation that is, *PAX6* and *NURR1*, suggesting that these cells may be withheld from progressing further to the early mid brain stage.

In conclusion, it was found that generating iPSCs under defined conditions minimizes degrees of variability in the resultant iPSC lines as evidenced by having similar transcriptomic profiles and differentiation capacity to hESCs. Key biological functions of pluripotent cells, in particular DNA replication and cell cycle, were distinguishing features of high quality iPSCs. Importantly, comparison between hESC under different culture conditions did not reveal the same signature. Additionally, distinctive signatures present in ff-iPSCs suggest that a defined system is favorable for inducing or at the minimum, maintaining the pluripotent state in hESCs and iPSCs. iPSCs created under different experimental conditions differ at both the transcriptomic and differentiation levels. It is important to highlight here that a standardized protocol can generate a homogenous population of iPSCs which closely resemble blastocyst-derived hESCs. The data presented here are of clinical relevance and the gene signatures have the potential to distinguish high quality human iPSCs for regenerative medicine.

Acknowledgments

Thanks goes to the Ramaciotti Centre (UNSW, Australia) for their service in Affymetrix microarray processing. This study was supported in part by NHMRC Program Grant No. 568969 to P.S.S. and UNSW Faculty of Medicine initiative on stem cells.

Author Disclosure Statement

The authors indicate no conflict of interests.

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. Yu J, MA Vodyanik, K Smuga-Otto, J Antosiewicz-Bourget, JL Frane, S Tian, J Nie, GA Jonsdottir, V Ruotti, et al. (2007). Induced pluripotent stem cell lines derived from human somatic cells. *Science* 318:1917–1920.
3. Takahashi K, K Tanabe, M Ohnuki, M Narita, T Ichisaka, K Tomoda and S Yamanaka. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 131:861–872.
4. Hockemeyer D, F Soldner, EG Cook, Q Gao, M Mitalipova and R Jaenisch. (2008). A drug-inducible system for direct reprogramming of human somatic cells to pluripotency. *Cell Stem Cell* 3:346–353.
5. Huangfu D, R Maehr, W Guo, A Eijkelenboom, M Snitow, AE Chen and DA Melton. (2008). Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. *Nat Biotech* 26:795–797.
6. Marson A, R Foreman, B Chevalier, S Bilodeau, M Kahn, RA Young and R Jaenisch. (2008). Wnt signaling promotes reprogramming of somatic cells to pluripotency. *Cell Stem Cell* 3:132–135.
7. Shi Y, C Desponts, JT Do, HS Hahm, HR Schöler and S Ding. (2008). Induction of pluripotent stem cells from mouse embryonic fibroblasts by Oct4 and Klf4 with small-molecule compounds. *Cell Stem Cell* 3:568–574.
8. Silva J, O Barrandon, J Nichols, J Kawaguchi, TW Theunissen and A Smith. (2008). Promotion of reprogramming to ground state pluripotency by signal inhibition. *PLoS Biol* 6:e253.
9. Carey BW, S Markoulaki, J Hanna, K Saha, Q Gao, M Mitalipova and R Jaenisch. (2009). Reprogramming of murine and human somatic cells using a single polycistronic vector. *Proc Natl Acad Sci U S A* 106:157–162.
10. Sommer CA, M Stadtfeld, GJ Murphy, K Hochedlinger, DN Kotton and G Mostoslavsky. (2008). iPS cell generation using a single lentiviral stem cell cassette. *Stem Cells* 27:543–549.
11. Soldner F, D Hockemeyer, C Beard, Q Gao, GW Bell, EG Cook, G Hargus, A Blak, O Cooper, et al. (2009). Parkinson's disease patient-derived induced pluripotent stem cells free of viral reprogramming factors. *Cell* 136:964–977.
12. Woltjen K, IP Michael, P Mohseni, R Desai, M Mileikovsky, R Hämäläinen, R Cowling, W Wang, P Liu, et al. (2009). piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. *Nature* 458:766–770.
13. Yu J, K Hu, K Smuga-Otto, S Tian, R Stewart, II Slukvin, and JA Thomson. (2009). Human induced pluripotent stem cells free of vector and transgene sequences. *Science* 324:797–801.
14. Kim D, CH Kim, JI Moon, YG Chung, MY Chang, BS Han, S Ko, E Yang, KY Cha, R Lanza and KS Kim. (2009). Generation of human induced pluripotent stem cells by direct delivery of reprogramming proteins. *Cell Stem Cell* 4:472–476.
15. Zhou H, S Wu, JY Joo, S Zhu, DW Han, T Lin, S Trauger, G Bien, S Yao, et al. (2009). Generation of induced pluripotent stem cells using recombinant proteins. *Cell Stem Cell* 4:381–384.
16. Chan EM, S Ratanasirintrao, IH Park, PD Manos, YH Loh, H Huo, JD Miller, O Hartung, J Rho, et al. (2009). Live cell imaging distinguishes bona fide human iPS cells from partially reprogrammed cells. *Nat Biotechnol* 27:1033–1037.
17. Feng Q, SJ Lu, I Klimanskaya, I Gomes, D Kim, Y Chung, GR Honig, KS Kim and R Lanza. (2010). Hemangioblastic derivatives from human induced pluripotent stem cells exhibit limited expansion and early senescence. *Stem Cells* 28:704–712.
18. Hu BY, JP Weick, J Yu, LX Ma, XQ Zhang, JA Thomson and SC Zhang. (2010). Neural differentiation of human induced pluripotent stem cells follows developmental principles but with variable potency. *Proc Natl Acad Sci U S A* 107:4335–4340.
19. Sun N, NJ Panetta, DM Gupta, KD Wilson, A Lee, F Jia, S Hu, AM Cherry, RC, Robbins, MT Longaker and JC Wu. (2009). Feeder-free derivation of induced pluripotent stem cells from adult human adipose stem cells. *Proc Natl Acad Sci U S A* 106:15720–15725.
20. Vallier L, T Touboul, S Brown, C Cho, B Bilican, M Alexander, J Cedervall, S Chandran, L Ahrlund-Richter, A Weber and RA Pedersen. (2009). Signaling pathways controlling pluripotency and early cell fate decisions of human induced pluripotent stem cells. *Stem Cells* 27:2655–2666.
21. Sidhu KS, JP Ryan, JG Lees and BE Tuch. (2010). Derivation of a new human embryonic stem cell line, Endeavour-2, and its characterization. *In Vitro Cell Dev Biol Anim* 46:269–275.
22. Sidhu KS, JP Ryan and BE Tuch. (2008). Derivation of a new hESC line, Endeavour-1 and its clonal propagation. *Stem Cells Dev* 17:41–52.
23. Ginn SL, J Fleming, PB Rowe and IE Alexander. (2003). Promoter interference mediated by the U3 region in early-generation HIV-1-derived lentivirus vectors can influence

- detection of transgene expression in a cell-type and species-specific manner. *Hum Gene Ther* 14:1127–1137.
24. Kawasaki H, K Mizuseki, S Nishikawa, S Kaneko, Y Kuwana, S Nakanishi, SI Nishikawa and Y Sasai. (2000). Induction of midbrain dopaminergic neurons from ES cells by stromal cell-derived inducing activity. *Neuron* 28:31–40.
 25. Pfaffl MW. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 29:e45.
 26. Lin RC, KL Weeks, XM Gao, RB Williams, BC Bernardo, H Kiriazis, VB Matthews, EA Woodcock, RD Bouwman, et al. (2010). PI3K(p110 alpha) protects against myocardial infarction-induced heart failure: identification of PI3K-regulated miRNA and mRNA. *Arterioscler Thromb Vasc Biol* 30:724–732.
 27. Ashburner M, CA Ball, JA Blake, D Botstein, H Butler, JM Cherry, AP Davis, K Dolinski, SS Dwight, et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 25:25–29.
 28. Kanehisa M and S Goto. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 28:27–30.
 29. Dennis G, Jr., BT Sherman, DA Hosack, J Yang, W Gao, HC Lane and RA Lempicki. (2003). DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol* 4:P3.
 30. Ghosh Z, KD Wilson, Y Wu, S Hu, T Quertermous and JC Wu. (2010). Persistent donor cell gene expression among human induced pluripotent stem cells contributes to differences with human embryonic stem cells. *PLoS ONE* 5:e8975.
 31. Saeed AI, NK Bhagabati, JC Braisted, W Liang, V Sharov, EA Howe, J Li, M Thiagarajan, JA White and J Quackenbush. (2006). TM4 microarray software suite. *Methods Enzymol* 411:134–193.
 32. Müller FJ, LC Laurent, D Kostka, I Ulitsky, R Williams, C Lu, IH Park, MS Rao, R Shamir, et al. (2008). Regulatory networks define phenotypic classes of human stem cell lines. *Nature* 455:401–405.
 33. Assou S, T Le Carrour, S Tondeur, S Strom, A Gabelle, S Marty, L Nadal, V Pantesco, T Reme, et al. (2007). A meta-analysis of human embryonic stem cells transcriptome integrated into a web-based expression atlas. *Stem Cells* 25:961–973.
 34. Mikkelsen TS, J Hanna, X Zhang, M Ku, M Wernig, P Schorderet, BE Bernstein, R Jaenisch, ES Lander and A Meissner. (2008). Dissecting direct reprogramming through integrative genomic analysis. *Nature* 454:49–55.
 35. Boue S, I Paramonov, MJ Barrero and JC Izpisua Belmonte. (2010). Analysis of human and mouse reprogramming of somatic cells to induced pluripotent stem cells. What is in the plate? *PLoS ONE* 5:e12664.
 36. Pan G, S Tian, J Nie, C Yang, V Ruotti, H Wei, GA Jonsdottir, R Stewart and JA Thomson. (2007). Whole-genome analysis of histone H3 lysine 4 and lysine 27 methylation in human embryonic stem cells. *Cell Stem Cell* 1:299–312.
 37. Zhao XD, X Han, JL Chew, J Liu, KP Chiu, A Choo, YL Orlov, WK Sung, A Shahab, et al. (2007). Whole-genome mapping of histone H3 Lys4 and 27 trimethylations reveals distinct genomic compartments in human embryonic stem cells. *Cell Stem Cell* 1:286–298.
 38. Ku M, RP Koche, E Rheinbay, EM Mendenhall, M Endoh, TS Mikkelsen, A Presser, C Nusbaum, X Xie, et al. (2008). Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. *PLoS Genet* 4:e1000242.
 39. Chin AC, J Padmanabhan, SK Oh and AB Choo. (2010). Defined and serum-free media support undifferentiated human embryonic stem cell growth. *Stem Cells Dev* 19:753–761.
 40. Chin MH, MJ Mason, W Xie, S Volinia, M Singer, C Peterson, G Ambartsumyan, O Aimiwu, L Richter, et al. (2009). Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. *Cell Stem Cell* 5:111–123.

Address correspondence to:

*Dr. Kuldip S. Sidhu
Stem Cell Laboratory
School of Psychiatry
University of New South Wales
G65 Clinical Sciences Building
Prince of Wales Hospital
Randwick, NSW 2031
Australia*

E-mail: k.sidhu@unsw.edu.au

Received for publication October 1, 2010

Accepted after revision April 20, 2011

Prepublished on Liebert Instant Online April 20, 2011