



A Practical and Efficient Cellular Substrate for the Generation of Induced Pluripotent Stem Cells from Adults: Blood-Derived Endothelial Progenitor Cells

IMBISAAT GETI,^{a*} MARK L. ORMISTON,^{b*} FOAD ROUHANI,^{a*} MARK TOSHNER,^b
MEHREGAN MOVASSAGH,^b JENNIFER NICHOLS,^c WILLIAM MANSFIELD,^c MARK SOUTHWOOD,^b
ALLAN BRADLEY,^d AMER AHMED RANA,^{a,b*†} LUDOVIC VALLIER,^{a†} NICHOLAS W. MORRELL^{b†}

Key Words. Induced pluripotent stem cells • Endothelial cell • Progenitor cells • Reprogramming • Direct cell conversion

ABSTRACT

Induced pluripotent stem cells (iPSCs) have the potential to generate patient-specific tissues for disease modeling and regenerative medicine applications. However, before iPSC technology can progress to the translational phase, several obstacles must be overcome. These include uncertainty regarding the ideal somatic cell type for reprogramming, the low kinetics and efficiency of reprogramming, and karyotype discrepancies between iPSCs and their somatic precursors. Here we describe the use of late-outgrowth endothelial progenitor cells (L-EPCs), which possess several favorable characteristics, as a cellular substrate for the generation of iPSCs. We have developed a protocol that allows the reliable isolation of L-EPCs from peripheral blood mononuclear cell preparations, including frozen samples. As a proof-of-principle for clinical applications we generated EPC-iPSCs from both healthy individuals and patients with heritable and idiopathic forms of pulmonary arterial hypertension. L-EPCs grew clonally; were highly proliferative, passageable, and bankable; and displayed higher reprogramming kinetics and efficiencies compared with dermal fibroblasts. Unlike fibroblasts, the high efficiency of L-EPC reprogramming allowed for the reliable generation of iPSCs in a 96-well format, which is compatible with high-throughput platforms. Array comparative genome hybridization analysis of L-EPCs versus donor-matched circulating monocytes demonstrated that L-EPCs have normal karyotypes compared with their subject's reference genome. In addition, >80% of EPC-iPSC lines tested did not acquire any copy number variations during reprogramming compared with their parent L-EPC line. This work identifies L-EPCs as a practical and efficient cellular substrate for iPSC generation, with the potential to address many of the factors currently limiting the translation of this technology. *STEM CELLS TRANSLATIONAL MEDICINE 2012;1:000-000*

INTRODUCTION

Induced pluripotent stem cells (iPSCs) provide the potential to generate patient-specific tissues for disease modeling, drug and toxicology screening, tissue replacement, and delivery of gene therapy [1, 2]. However, for iPSC technology to progress to the translational phase, several obstacles remain. These include uncertainty regarding the ideal somatic cell type for reprogramming, the low kinetics and efficiency of reprogramming, and genomic differences between iPSCs and their somatic progenitors [3–5], which might preclude them from clinical applications.

Many advances have been made in the technology of transgene delivery used in nuclear reprogramming protocols. However, the choice of starting cell type is also critical. The characteristics of an ideal cellular substrate for simple, efficient, and large-scale iPSC generation would include cells that are (a) clearly defined; (b)

reproducibly isolated and easily cultured from subjects of all ages and across a spectrum of normal and disease states; (c) expandable in culture and suitable for banking (long-term storage); (d) in possession of a “normal,” unaltered genome; and (e) capable of highly efficient reprogramming.

Recent studies have shown that the genomes of many fibroblast-derived iPSC lines are altered compared with the cells from which they were derived [3–5]. In these reports, the genomes of the lines analyzed exhibited copy number variations (CNVs). The frequency of CNV correlated with the reprogramming technique, with retroviral methods resulting in the highest number of CNVs. It has been suggested that these defects are a consequence of the reprogramming process [3]. These findings have implications for the use of iPSCs in disease modeling, since genomic anomalies could alter phenotypic behavior.

^aWellcome Trust-Medical Research Council Cambridge Stem Cell Institute, Anne McLaren Laboratory for Regenerative Medicine, University of Cambridge, Cambridge, United Kingdom;

^bDivision of Respiratory Medicine, Department of Medicine, University of Cambridge, Addenbrooke's Hospital, Cambridge, United Kingdom; ^cThe Wellcome Trust-Medical Research Council Cambridge Stem Cell Institute, University of Cambridge, Cambridge, United Kingdom; ^dWellcome Trust Sanger Institute, Hinxton, Cambridge, United Kingdom

*Contributed equally.

†Joint senior authors.

Correspondence: Amer Ahmed Rana, Ph.D., Division of Respiratory Medicine, Department of Medicine, University of Cambridge, Box 157, Level 5, Addenbrooke's Hospital, Hills Road, Cambridge CB2 0QQ, United Kingdom. Telephone: 44-1223-336862; Fax: 44-1223-336846; E-Mail: ar332@cam.ac.uk

Received July 27, 2012; accepted for publication September 27, 2012; first published online in *SCTM EXPRESS* November 29, 2012.

©AlphaMed Press
1066-5099/2012/\$20.00/0

[http://dx.doi.org/
10.5966/sctm.2012-0093](http://dx.doi.org/10.5966/sctm.2012-0093)

Genomic instability also is a concern for cell-based therapies as this raises the potential for malignant transformation. An alternative cause for the observed increase in mutational load is that at least some of the CNVs are due to the genomes of clonally derived iPSCs being compared with a reference genome generated from a polyclonal population of fibroblasts, only some of which have a genome representative of the parent fibroblast cell. Therefore an additional characteristic of an ideal cellular substrate for the generation of iPSCs might be that the cellular substrate is clonally derived or capable of clonal expansion prior to reprogramming. This would allow the generation of a clonal reference genome, which is essential for subsequent high resolution genetic testing such as array comparative genome hybridization and comparison of genomic sequence.

Although skin fibroblasts are the most common cell type used for generating iPSCs, these cells may not be ideal for several reasons. The isolation of these cells requires a surgical procedure, which is undesirable in children, patients with skin disorders, and patients with abnormal coagulation or wound healing. In addition, fibroblasts reprogram with relatively low efficiency [1, 2]. A universally acceptable alternative is venous blood sampling, which is a relatively minor, well-tolerated procedure. Although T cells and myeloid cells derived from peripheral blood can be used to generate patient-specific iPSCs [6–11], the use of these cells is limited by several factors. These include the low capacity of these cells to expand in culture, their low reprogramming efficiency, and the presence of permanent genomic rearrangements in T cells. Hematopoietic stem cells derived from the bone marrow compartment can be reprogrammed with higher efficiencies, but obtaining these cells is not trivial, requiring mobilization of the bone marrow compartment or bone marrow aspiration [12, 13]. Therefore, we considered other blood-derived cell types that might have broad application. Previous reports have shown the utility of other mononuclear cells, such as CD34⁺ cells derived from cord and peripheral blood as substrates for iPSC generation [14–19]. However, as yet, no cell type, blood-derived or otherwise, has been described that demonstrates all the desired attributes necessary to move iPSCs to the translational phase.

Late-outgrowth endothelial progenitor cells (L-EPCs; also known as blood-outgrowth endothelial cells) arise from the mononuclear cell fraction of peripheral blood under endothelial-selective conditions [20, 21]. Although early reports suggested that these cells arise from the bone marrow [20], L-EPCs isolated from patients with chronic myeloid leukemia or polycythemia vera do not carry the somatic hematopoietic stem cell mutations associated with these conditions [22, 23], supporting the suggestion that L-EPCs are not of myeloid lineage. Here we report the use of adult blood-derived late-outgrowth endothelial progenitor cells (L-EPCs) [20, 21] as a novel cellular reprogramming substrate, with several potential advantages for the translation of iPSC technology.

MATERIALS AND METHODS

Isolation and Characterization of L-EPCs

Human mononuclear cells were obtained from 40–80 ml of peripheral blood by density gradient centrifugation using Ficoll Paque Plus (GE Healthcare, Little Chalfont, U.K., <http://www.gehealthcare.com>). Washed samples were cultured at 1.5×10^6 cells per cm² on collagen (BD Biosciences, Franklin Lakes, NJ, <http://www.bdbiosciences.com>)-coated T75 or T25 flasks in

EGM-2MV medium (Lonza, Basel, Switzerland, <http://www.lonza.com>) containing 20% embryonic stem cell-grade fetal bovine serum (HyClone, Thermo Scientific, Basingstoke, U.K., <http://www.hyclone.com>). Culture medium was changed every 2 days. Late-outgrowth EPCs appeared between 10 and 14 days in culture. Following generation, EPCs were passaged onto tissue culture plastic and maintained in EGM-2MV containing 10% standard fetal bovine serum (FBS) (Gibco, Invitrogen, Paisley, U.K., <http://www.invitrogen.com>).

Flow Cytometry

Peripheral blood mononuclear cells (PBMCs) were isolated from whole blood samples by Ficoll density gradient centrifugation and monocytes were isolated by positive magnetic selection using CD14-microbeads (Miltenyi Biotec, Bergisch Gladbach, Germany, <http://www.miltenyibiotec.com>) as per the manufacturer's instructions. L-EPCs and human pulmonary artery endothelial cells (PAECs) were trypsinized prior to resuspension in staining buffer containing 2% FBS and 2 mM EDTA. Monocytes, L-EPCs, and PAECs were stained with allophycocyanin (APC)-conjugated mouse- α -human vascular endothelial growth factor receptor 2 (clone 89106; R&D Systems Inc., Minneapolis, MN, <http://www.rndsystems.com>), fluorescein isothiocyanate (FITC)-conjugated mouse- α -human CD31 (clone WM59), APC-conjugated mouse- α -human CD34 (clone 581), FITC-conjugated mouse- α -human CD14 (clone M5E2), and FITC-conjugated mouse- α -human CD45 (clone HI30, all from BD Biosciences) or the appropriate isotype controls prior to analysis.

Generation and Culturing of Fibro-iPSCs and L-EPC-iPSCs

Generation and culturing of fibro-iPSCs was performed as previously described [5, 24] using 100,000 starting cells and with the following modifications. Fibroblasts were transduced at 32°C. The next day cells were washed three times with phosphate-buffered saline (PBS) and switched to mouse embryonic fibroblast (MEF) medium at 37°C. On day 5 cells were split and added to an MEF feeder plate. From day 7, cells were cultured in KSR + fibroblast growth factor 2 (FGF2) medium. L-EPC-iPSC generation and culturing was performed as for fibro-iPSCs except that L-EPCs were maintained in EGM-2MV + 10% serum until day 5 when they were transferred to MEF-coated feeder plates and MEF medium. Percentage of reprogramming efficiency was calculated as follows: iPSC colony number at day 20/33,333 \times 100.

Activin and FGF2 were provided by Marco Hyvonen (Department of Biochemistry, University of Cambridge).

Assessment of Oct4 Promoter Methylation, Bisulfite-Polymerase Chain Reaction, and Pyrosequencing

Bisulfite modification of genomic DNA (1 μ g) was performed using the EpiTect DNA methylation kit (Qiagen, Crawley, U.K., <http://www.qiagen.com>) as recommended by the manufacturer. Bisulfite-polymerase chain reaction (BS-PCR) was performed to amplify the promoter region of the *Oct3/4* gene (GRCh37, Chr6: 31,140,564–31,140,784), using previously reported primers [10]. The reverse primer was biotinylated for the template strand and the streptavidin-captured single-strand DNA was pyrosequenced using pyrosequencing primers 2 and 3 to cover all the CpGs sites within this region with the exception of

the first CpG. The first CpG was pyrosequenced using the biotinylated forward primer for the BS-PCR instead, and pyrosequencing was performed using the pyrosequencing primer-4. Pyrosequencing runs were performed using PyroGold Q96 SQA reagents on the PyroMark ID pyrosequencer (Qiagen) as per the manufacturer's recommendation. The pyrosequencing data were analyzed using Pyro Q-CpG software (Qiagen), and results are presented as percentage of methylation for each of the CpG sites. All primer sequences for BS-PCR and pyrosequencing can be found in supplemental online Table 1. BS-PCR was performed at a final magnesium chloride concentration of 3 mM with the following program: 95°C for 10 minutes; 50 cycles of 95°C for 20 seconds, 55°C for 20 seconds, and 72°C for 1 minutes; and 1 cycle of 72°C for 10 minutes.

Comparative Genomic Hybridization Analysis

Comparative genomic hybridization (CGH) analysis and the identification of CNVs were performed as previously described [5]. In summary, genomic DNA was extracted using the DNeasy kit (Qiagen). Agilent 244k human genome arrays (Agilent Technologies, Palo Alto, CA, <http://www.agilent.com>) were used following the manufacturer's protocol. The arrays were scanned using an Agilent microarray scanner, and the data were generated by Agilent Feature Extraction software. The analysis was performed using Agilent Genomic Workbench software, and CGH calls were made using the ADM-2 algorithm (6.0 threshold) with a minimum of three consecutive probes detecting a region of abnormality.

Directed Differentiation Indicates Chemically Defined Medium

Serum directed differentiation of extraembryonic and neuroectoderm was performed as previously described [24]. Mesendoderm differentiation was performed in a 3-day differentiation protocol in the following way: day 1 cells were cultured in chemically defined medium + polyvinyl alcohol (as previously described) + 100 ng/ml Activin + 100 ng/ml FGF2 + 10 ng/ml bone morphogenetic protein 4 (BMP4) + 10 μ M Ly + 3 μ M chir. On day 2 cells were switched to 100 ng/ml Activin + 100 ng/ml FGF2 + 10 ng/ml BMP4 + 10 μ M Ly excluding chir. On day 3 cells were switched to RPMI medium + 100 ng/ml Activin + 100 ng/ml FGF2.

Generation and Histological Analysis of Teratomas

EPC-iPSCs were injected into SCID or SCID Beige mice either intraperitoneally, intramuscularly, or under the kidney capsule. Mice were maintained for at least 14 weeks postinjection of iPSCs, and every care was taken in following strict local ethical policies and Home Office rules concerning animal uses and regulated procedures.

Formalin-fixed, paraffin-embedded EPC-iPSC-derived teratoma specimens were sectioned (4 μ m), and antigen retrieval was performed using target retrieval solution (Dako, Ely, U.K., <http://www.dako.com>). Monoclonal mouse anti-human smooth muscle actin- α , CD31, p63, cytokeratin-7, cytokeratin-20, and HMB45 (all Dako) were immunolabeled using a peroxidase coupled dextran polymer (Envision; Dako). Antibodies were visualized using 3,3'-diaminobenzidine tetrahydrochloride to create a brown reaction product, counterstained with hematoxylin (Dako), and examined by light microscopy.

Immunofluorescence Staining

Immunostaining was performed as previously described [24], with donkey and goat serum C07SA from AbD Serotec (Raleigh, NC, <http://www.ab-direct.com>). A list of primary antibodies can be found in supplemental online Table 2.

Alkaline Phosphatase Staining of iPSC Colonies

Cells were fixed in 4% paraformaldehyde for 20 minutes at 4°C and then rinsed three times in PBS followed by alkaline phosphatase (AP) solution (0.1 M Tris, pH 9.5, 0.1 M NaCl). They were then incubated for 24 hours at 4°C in 10 ml of AP solution supplemented with 200 μ l of Nitro Blue Tetrazolium + 20 μ l of BCIP (5-bromo-4-chloro-3-indolyl-phosphatase; catalog no. S3771; Promega, Madison, WI, <http://www.promega.com>). Finally, cells were washed with PBS once staining was complete.

Viral Insertion Copy Number Analysis

Viral insertion copy number analysis was performed as previously described [24] using the Quant-iT DNA Assay Kit, Broad Range (Molecular Probes Q33135) using the viral copy number and the transgene-specific and endogenous gene expression primers listed in supplemental online Tables 3–5.

RESULTS

Generation and Characterization of Late-Outgrowth Endothelial Progenitor Cells

Using our modified protocol, we were able to isolate L-EPCs from all individuals tested ($n = 25$; patient data and method of isolation are given in supplemental online Table 6). Late-outgrowth EPCs emerged from within cultures of early endothelial progenitor cells (Fig. 1A). Early-EPCs formed adherent cultures with a monocyte-derived macrophage morphology and predominated in the culture flask up to around day 15, after which they senesced. However, L-EPCs emerged around day 10, forming between two and seven highly proliferative colonies per flask. These become the predominant cell type exhibiting a cobblestone morphology resembling endothelial cell cultures and were able to form endothelial cell-like networks in vitro. Although the circulating cell type that gives rise to L-EPCs is unclear, once apparent in culture they were unequivocally defined as CD31^{high}:CD34⁺:CD146⁺:kinase insert domain receptor⁺:von Willebrand factor⁺:CD14⁻:CD45⁻ (Fig. 1B, 1C). L-EPCs were highly proliferative, doubling every 24–28 hours, could be passaged (at least 10 times), and could be frozen and thawed without impacting on viability or cell phenotype (data not shown). Importantly, we also found that EPCs are readily generated from frozen PBMNC preparations ($n = 3$; supplemental online Table 6), which makes storage and transportation of samples straightforward. In this case, following density gradient centrifugation, freshly isolated PBMNCs were frozen in 90% ES-cell grade FBS and 10% dimethyl sulfoxide. To confirm the robustness of L-EPCs as a cellular substrate, we derived iPSCs using cells from healthy control subjects ($n = 3$) and from patients with heritable pulmonary arterial hypertension (PAH), a condition characterized by endothelial dysfunction ($n = 5$) [25].

Generation and Characterization of EPC-iPSCs

Figure 2A provides an overview of the approach used to generate iPSCs from L-EPCs. iPSC generation was performed by standard techniques using the exogenous expression of OCT4, SOX2, KLF4,

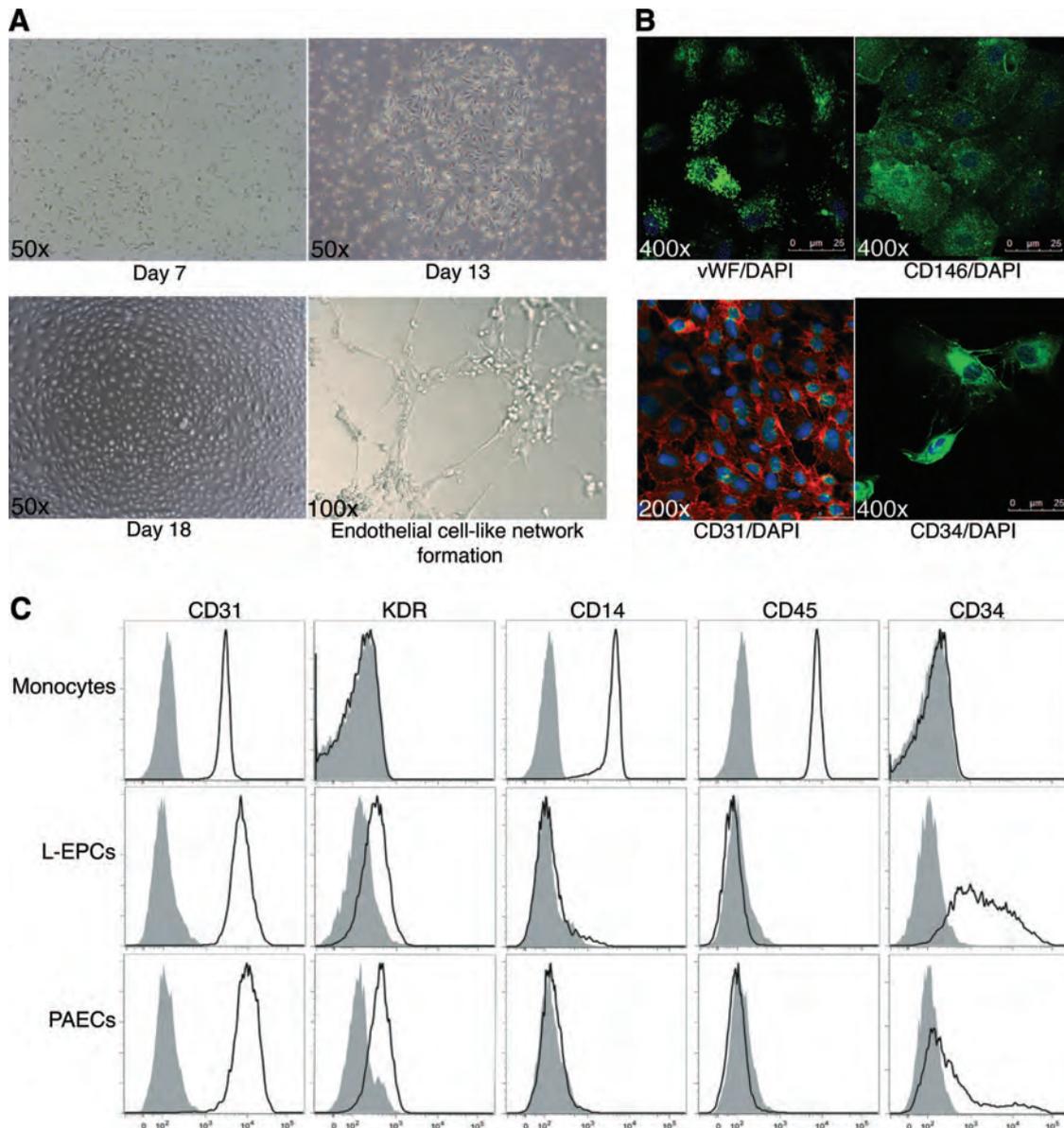


Figure 1. Generation and characterization of L-EPCs. **(A):** Generation of L-EPCs. Late-outgrowth endothelial progenitor cells emerged from within cultures of early endothelial progenitor cells. Early-EPCs (E-EPCs) formed adherent cultures with a monocyte-derived macrophage type morphology (day 7) and predominated in the culture flask up to around day 15; however, L-EPCs emerged by day 10, forming highly proliferative colonies (day 13) of cells resembling endothelial cells, and became the predominant cell type in the flask exhibiting the cobblestone morphology shared with endothelial cell cultures (day 18). L-EPCs were able to form endothelial cell-like networks in vitro. **(B):** Characterization of established L-EPCs. Immunostaining of L-EPCs at passage 4 revealed expression of endothelial cell-specific markers vWF (green; DAPI, blue), CD146 (green; DAPI, blue), and CD31 (red; DAPI, blue) and also the progenitor cell marker CD34 (green; DAPI, blue). **(C):** Flow cytometric analysis demonstrated that L-EPCs (passage 4) expressed surface marker expression similar to that of pulmonary artery endothelial cells when compared with freshly isolated monocytes (which predominated in the E-EPC population). Abbreviations: DAPI, 4',6-diamidino-2-phenylindole; L-EPC, late-outgrowth endothelial progenitor cell; KDR, kinase insert domain receptor; PAEC, pulmonary artery endothelial cell; vWF, von Willebrand factor.

and c-MYC as previously described by our laboratories to generate fibroblast-iPSCs [5, 24]. We chose to use retroviral methods since this is the approach used by most laboratories to reprogram fibroblasts and allows some comparison with previous results obtained using fibroblasts as the starting cell substrate [3–5]. As stated above, L-EPCs were derived from three healthy control individuals (C3-EPC, C4-EPC, and C7-EPC) and five patients with PAH (supplemental online Table 6). Three patients carried a mutation in the gene encoding the bone morphogenetic protein type II receptor (*BMPR2*), the gene most commonly mutated in heritable PAH [25].

Of these, patient 1-endothelial progenitor cells (P1-EPC) and P2-EPC carried a W9X mutation, and P5-EPC carried a C347R mutation. In two patients (P3-EPC and P4-EPC) no *BMPR2* mutation was identified. As comparators of reprogramming dynamics we used the fibroblast lines CRL (Fibro1, from neonatal foreskin) and PedB (Fibro2, from dermal skin taken from a male patient in his 50s), previously shown to reprogram to iPSCs at the higher end of the range of fibroblast reprogramming efficiency [24] (a summary of all iPSC lines generated and characterization performed can be found in supplemental online Table 7).

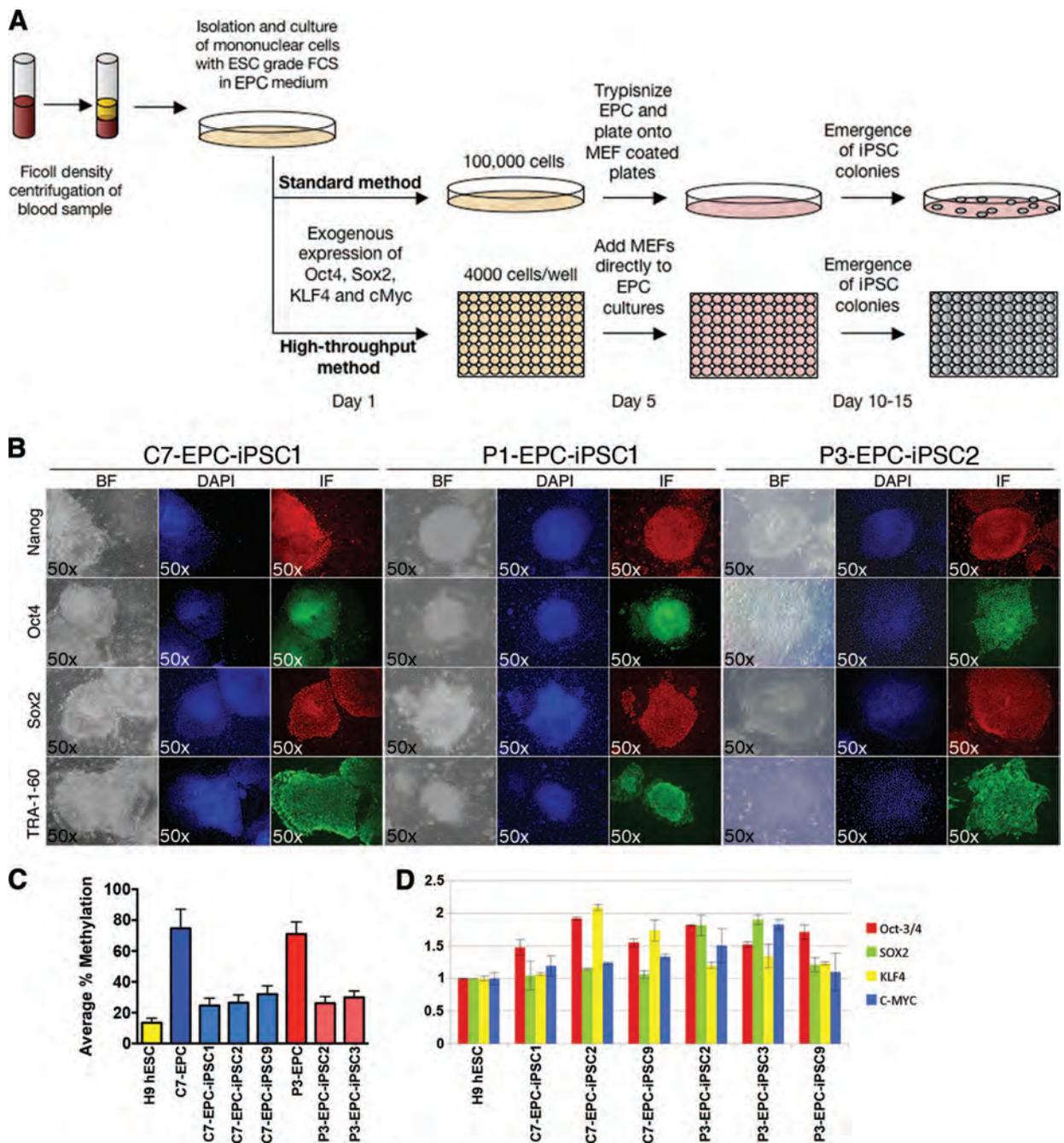


Figure 2. Summary of methodology and basic characterization of EPC-iPSCs. **(A):** Overview of EPC-iPSC derivation. iPSC derivation in standard 10-cm dishes and 96-well high-throughput method. Mononuclear cells are isolated from nonmobilized peripheral blood. These cells are transduced with retrovirus expressing Oct4, Sox2, c-Myc, and KLF4 and either added to MEFs or MEFs are added to infected late-outgrowth endothelial progenitor cell (L-EPC) cultures. Colonies of iPSCs emerge and are either picked or stained. **(B):** Pluripotency marker expression in three exemplar lines of iPSCs generated from EPCs. Panels in the first group of columns (C7-EPC-iPSC1) are subline 1 of iPSCs derived from L-EPC line C7-EPC (from normal control). Panels in the middle group of columns (P1-EPC-iPSC1) are subline 1 of iPSCs derived from L-EPC line P1-EPC (from a patient with pulmonary arterial hypertension [PAH] harboring a mutation in bone morphogenetic protein type II receptor (BMPR2)). Panels in the third group of columns (P3-EPC-iPSC2) are subline 2 of iPSCs derived from L-EPC line P3-EPC (from a patient with PAH without an identifiable mutation in BMPR2). All iPSC lines expressed the markers NANOG, OCT4, SOX2, and TRA-1-60, which is consistent with a pluripotent state. **(C):** Relative levels of DNA methylation on the *Oct4* promoter of EPC-iPSCs generated from L-EPC compared with H9 hESCs. L-EPC lines C7-EPC and P3-EPC3 were analyzed along with their derivative iPSC sublines 1, 2, and 9 and 2 and 3, respectively. **(D):** Viral transgene insertion rates of EPC-iPSCs generated from L-EPC lines C7-EPC (see Results section and supplemental online Table 6) and P3-EPC (see Results section and supplemental online Table 6). The number of viral integration was obtained by quantitative polymerase chain reaction genotyping analyses quantifying the number of copies of each gene relative to the endogenous levels of each gene in H9 hESCs. Thus, the H9 control has two copies represented as one on the y-axis. Abbreviations: BF, bright field; C, control; DAPI, 4',6-diamidino-2-phenylindole; EPC, endothelial progenitor cell; FCS, fetal calf serum; hESC, human embryonic stem cell; IF, immunofluorescence; iPSC, induced pluripotent stem cell; MEF, mouse embryonic fibroblast; P, patient.

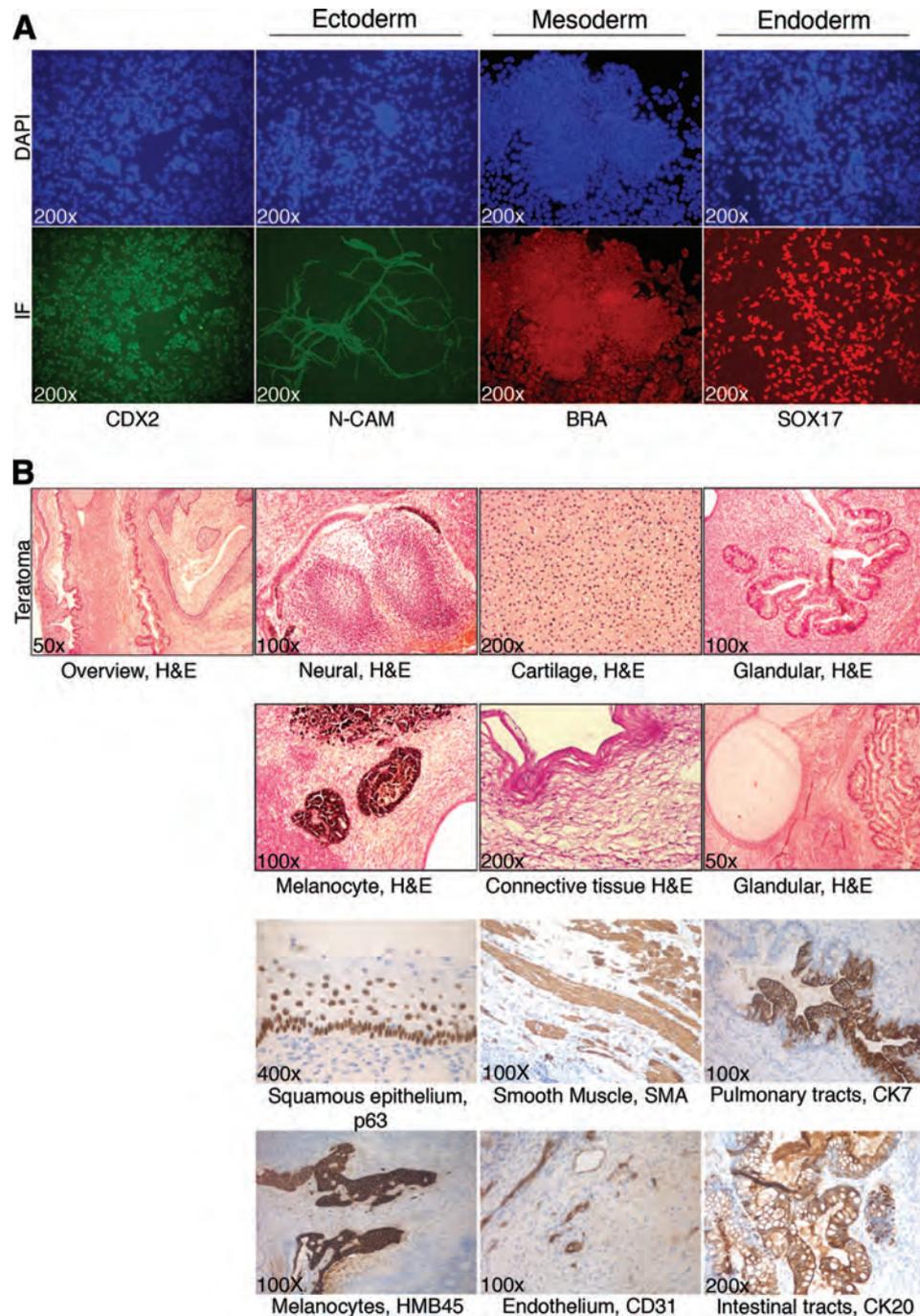


Figure 3. Differentiation of endothelial progenitor cell (EPC)-induced pluripotent stem cells (iPSCs). **(A):** A single iPSC exemplar line (C7-EPC-iPSC1). EPC-iPSCs could be differentiated using chemically defined differentiation protocols *in vitro* into derivatives of the three germ layers, ectoderm (N-CAM), mesoderm (BRA), and endoderm (SOX17), and tissues expressing posterior mesoderm/extraembryonic tissues (CDX2). **(B):** Two exemplar lines of EPC-iPSCs (C7-EPC-iPSC1 and P1-EPC-iPSC3). *In vivo* teratoma analyses EPC-iPSCs also differentiated into derivatives of the three germ layers identifiable by H&E staining (bottom two rows) or using specific antibodies (ectodermal derivatives, P63, HMB45; mesodermal derivatives, SMA, CD31; endoderm, CK7, CK20). Abbreviations: BRA, brachyury; CK, cytokeratin; DAPI, 4',6-diamidino-2-phenylindole; H&E, hematoxylin and eosin; IF, immunofluorescence; N-CAM, neural cell adhesion molecule; SMA, smooth muscle actin.

Pluripotency marker expression on all lines confirmed that L-EPCs reprogrammed to a state resembling human embryonic stem cells (Fig. 2B; supplemental online Fig. 1). L-EPC-derived iPSCs (EPC-iPSCs) showed demethylation of the OCT4 promoter (Fig. 2C) and appropriate silencing of the exogenous reprogramming factor expression (supplemental online Fig. 1), which in-

serted at 1–2 copies per factor (Fig. 2D). The differentiation capacity of EPC-iPSCs was tested *in vitro*. We confirmed expression of markers associated with each of the three germ layers (ectoderm, mesoderm, and endoderm) and extraembryonic tissues (Fig. 3A) in all EPC-iPSC lines. We further tested whether EPC-iPSCs could generate teratomas *in vivo* by injecting cells from

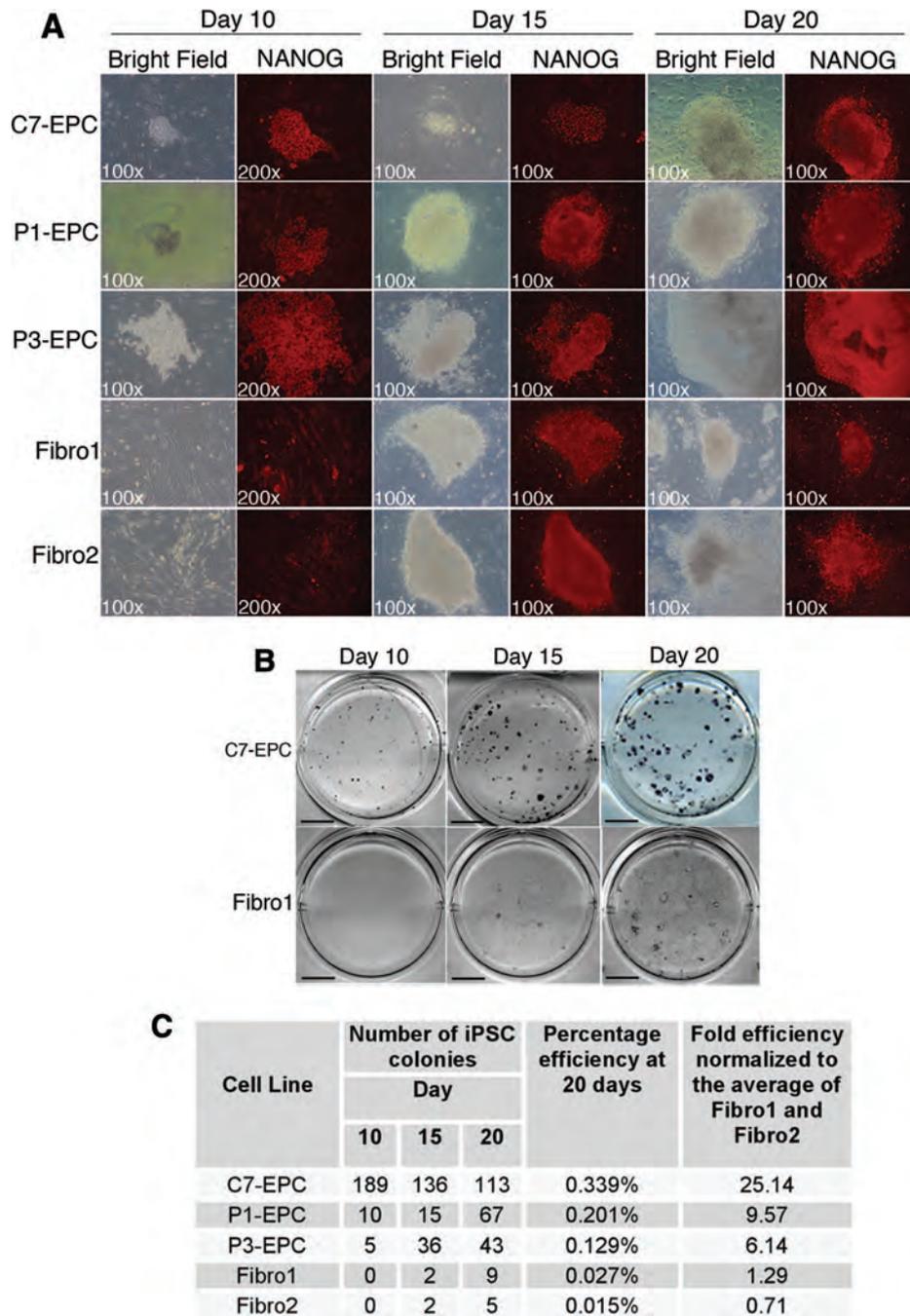


Figure 4. Kinetics of iPSC generation from late-outgrowth endothelial progenitor cells (L-EPCs). **(A):** NANOG-expressing colonies were detected from day 10 of reprogramming of L-EPCs and day 15 of fibroblasts (day 10 NANOG expression panels are shown at twice the magnification of all other panels). **(B):** Exemplar comparison of iPSC derivation efficiency between C7-EPC and Fibro1. Each well contained 33,333 infected cells. iPSC colonies stained with alkaline phosphatase (AP) positive at different time points. Scale bars = 8.75 mm. **(C):** Frequencies of AP-positive colonies derived from three L-EPC lines and two fibroblast lines. Experimental design was the same as described for **(B)**. Abbreviations: C, control; EPC, endothelial progenitor cell; Fibro, fibroblast; iPSC, induced pluripotent stem cell; P, patient.

two lines (C7-EPC-iPSC line 1 and P1-EPC-iPSC line 1) into immunocompromised mice. Analysis of the resulting tumors clearly demonstrated structures and cell types specific to each of the three germ layers, confirming the pluripotent nature of the reprogrammed EPCs (Fig. 3B). We also confirmed that the *BMPR2* mutations described in P1-EPC and P5-EPC were also present in EPC-iPSC lines derived from them (a *BMPR2* ± W9X mutation in P1-EPC-iPSC line 1 and a *BMPR2* ± C347R mutation in P5-EPC-iPSC line 1; supplemental online Fig. 2).

Efficiency and Kinetics of L-EPC Reprogramming

We next assessed the kinetics and efficiency of reprogramming of L-EPCs, controlling our method with the reprogramming of two fibroblast cell lines previous used to generate iPSC [24]. Using L-EPCs, NANOG, and alkaline phosphatase-positive iPSC colonies appeared as early as 10 days after expression of the reprogramming factors, compared with 15 days for fibroblasts, suggesting that the kinetics of L-EPC reprogramming was significantly faster (Fig. 4A, 4C). Sec-

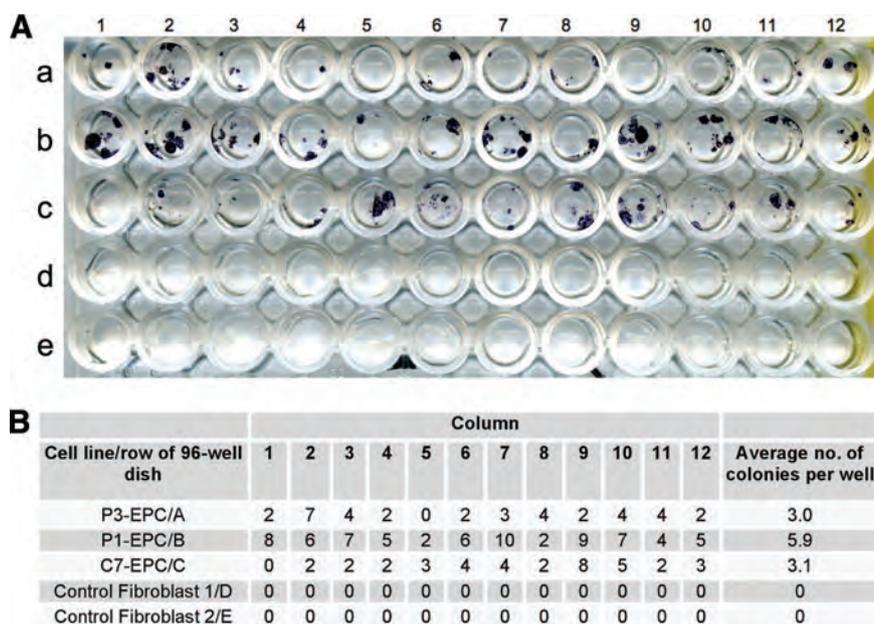


Figure 5. Ninety-six-well format reprogramming of late-outgrowth endothelial progenitor cells (L-EPCs). **(A):** 4,000 L-EPCs or fibroblasts were added to one well of a 96-well dish. This was replicated 11 times, so that every well in a row had 4,000 cells of the same line. These were infected with the reprogramming viruses, and on day 5 mouse embryonic fibroblasts were added to the wells. Alkaline phosphatase (AP) staining was carried out on day 15, and induced pluripotent stem cell colonies appeared blue/purple. **(B):** Frequencies of AP-positive iPSC colonies per well. Abbreviations: C, control; EPC, endothelial progenitor cell; P, patient.

ond, iPSC colonies formed from L-EPCs with an average efficiency of 0.22%, around 10-fold more frequently than for the two fibroblast lines tested (Fig. 4B, 4C); however, direct comparisons between reprogramming kinetics and efficiencies were not tested here, as the fibroblasts and EPCs were not isogenic and were derived differently. Nonetheless the kinetics and efficiency of L-EPC reprogramming compares favorably with other cell types used in the field. We therefore decided to take advantage of these features of L-EPC reprogramming and designed a method that could be used to generate iPSCs from a limited number of cells in parallel. We added either 4,000 EPCs or fibroblasts to single wells of a 96-well tissue culture dish and infected them with the retrovirally encoded reprogramming factors. On day 5 postinfection MEF feeder cells were added to each well. The cells were then left for a further 10 days and stained for alkaline phosphatase (Fig. 5A). iPSC colonies were not observed in any of the 12 wells for either fibroblast cell line used, whereas between three and six colonies were observed in 34 of 36 wells for the three L-EPC lines tested (Fig. 5B), providing further evidence for the favorable efficiency of L-EPCs as a cellular substrate and demonstrating the potential for these cells to be used in applications that would benefit from higher throughput generation of iPSCs.

Karyotype Analysis of EPC-iPSCs

One potential source of CNV in iPSCs is the acquisition of abnormalities during somatic cell aging *in vivo* or prolonged culture *ex vivo* [26–28]. Although the karyotype of L-EPCs has been shown to be normal [21], higher resolution analyses have not been undertaken. To test the genomic stability of L-EPCs in culture we performed array comparative genomic hybridization (aCGH) analysis at passages 3–4 (Table 1) compared with DNA from freshly isolated CD14⁺ monocytes from the same individual. This analysis revealed that the genomes of five of six L-EPC lines were either identical to their corresponding matched monocyte reference genome or demonstrated a single 13.1-kb copy gain at 1p21.3. This region contains the LCE3B/LCE3C genes, previously linked with susceptibility to psoriasis [29]. Deletion of this region as assessed by aCGH occurs in 55%–71% of European populations using whole blood-derived DNA as the reference genome

[29]. It is possible that the occurrence of independent copy gains of this specific region in three of six L-EPC genomes is explained by the common loss of this region in the blood-derived CD14⁺ monocytes, used as our reference genome. In one L-EPC line (EPC-C6) a proportion (20%–30%) of the L-EPC population appeared to have monosomy chromosome 18, although the majority of cells possessed a matched genome. Therefore, L-EPCs appear to demonstrate a reassuringly high level of genetic stability in culture when compared with the reference genome obtained from freshly isolated circulating monocytes.

To determine the impact of reprogramming on copy number variation we conducted aCGH analysis on iPSCs grown for 3–11 passages compared with their corresponding parental EPC lines. Compared with parental L-EPC DNA, 9 of 11 EPC-iPSC lines (derived from three different EPC lines) showed no detectable genomic abnormalities (Table 2). The remaining two EPC-iPSC lines demonstrated single copy gains of regions of 36 or 632 kb compared with the corresponding parental L-EPC genome (Table 2; supplemental online Fig. 3). Interestingly we have not observed CNVs on chromosomes 8, 12, 17, or 20, reported to be common in human embryonic stem cell and human iPSC (hiPSC) lines [3, 4, 26, 28]. The data presented here demonstrate that it is possible to generate iPSCs whose genomes are unaltered compared with their parental cell type in a significant proportion of isolates.

To examine this further, DNA from four EPC-iPSC lines from two subjects were compared with matched DNA derived from their CD14⁺ monocytes. One line from subject C7 (passage 10) showed a single copy gain and loss relative to the corresponding matched monocyte genome covering 230 kb. The three lines from individual C4 (at early passages, 3–4) showed between 10 and 17 copy number gains and losses compared with the matched monocyte genome (supplemental online Table 8), whereas fibroblast-iPSC at the same passage have been described to have >100 CNVs [3]. In this case we did observe CNVs on chromosomes 8, 12, 17, and 20.

DISCUSSION

The present study demonstrates that L-EPCs are highly proliferative, passageable, and bankable and have normal karyotypes.

Table 1. Array comparative genomic hybridization analysis of EPC genomes compared with matched monocyte genomes

| Subject | Reference genome (from fresh blood sample) | EPC genome (passages 3–4) | Copy number variation | Chromosome and band | Size (kb) | Gene involved |
|---------|--|---------------------------|---|---------------------|-----------|---------------|
| C1 | C1 monocyte | C1-EPC | None | | | |
| C2 | C2 monocyte | C2-EPC | None | | | |
| C3 | C3 monocyte | C3-EPC | 1 copy gain | 1q21.3 | 13.1 | <i>LCE3C</i> |
| C4 | C4 monocyte | C4-EPC | 1 copy gain | 1q21.3 | 13.1 | <i>LCE3C</i> |
| C5 | C5 monocyte | C5-EPC | 1 copy gain | 1q21.3 | 13.1 | <i>LCE3C</i> |
| C6 | C6 monocyte | C6-EPC | Partial monosomy in 20%–30% of population | Chromosome 18 | | |

Abbreviations: C, healthy control; EPC, endothelial progenitor cell.

Table 2. Array comparative genomic hybridization analysis of EPC-iPSC genomes compared with parental late-outgrowth EPC genomes

| Subject | Reference genome | EPC-iPSC line genome (passages 3–9) | Copy number variation | Chromosome and band | Size (kb) | Genes |
|---------|------------------|-------------------------------------|-----------------------|---------------------|-----------|-----------------------------------|
| P1 | P1-EPC | P1-EPC-iPSC1 | None | | | |
| | P1-EPC | P1-EPC-iPSC2 | None | | | |
| | P1-EPC | P1-EPC-iPSC3 | 1 copy gain | 15q14 | 36.6 | <i>LOC723972</i> |
| P2 | P2-EPC | P2-EPC-iPSC1 | None | | | |
| | P2-EPC | P2-EPC-iPSC2 | None | | | |
| | P2-EPC | P2-EPC-iPSC3 | None | | | |
| | P2-EPC | P2-EPC-iPSC4 | 1 copy gain | 16q23.1 | 632.7 | <i>NUDT7, VAT1L, CLEC3A, WWOX</i> |
| | P2-EPC | P2-EPC-iPSC5 | None | | | |
| P3 | P3-EPC | P3-EPC-iPSC1 | None | | | |
| | P3-EPC | P3-EPC-iPSC2 | None | | | |
| | P3-EPC | P3-EPC-iPSC3 | None | | | |

Abbreviations: EPC, endothelial progenitor cell; iPSC, induced pluripotent stem cell; P, patient with pulmonary arterial hypertension.

L-EPCs display high reprogramming kinetics and efficiencies compatible with high-throughput platforms, taking just 10 days to emerge in culture and forming iPSC colonies around 10-fold more readily than the two fibroblast lines we used as comparators. Taking the clonal expansion rate of L-EPCs (doubling rate of 24–28 days) together with the observation that as few as 4,000 EPCs can be reprogrammed to generate three to six iPSC colonies, the process of generating EPC-iPSC directly from a peripheral blood sample could take 24 days. Other methods have been described that have reported generating hiPSCs using polyclonal populations of starting cells in less time. However, when we consider that (a) both the L-EPC parent and iPSC lines are clonally derived, which allows meaningful genetic testing, (b) L-EPCs have a genome representative of a subject's normal genome, and (c) the reassuringly high level of genetic similarity between the EPC-iPSCs and their L-EPC parent line, then using L-EPCs as a starting cell for the refinement of iPSC technology for translational medicine applications has particular advantages.

A question that arises from our study is why L-EPCs are so amenable to reprogramming. One reason might be that L-EPCs possess progenitor-like characteristics, for example their high level of expression of CD34, when compared with endothelial cells (Fig. 1). However, in previous reports this was not associated with higher reprogramming efficiency [14, 17]. Another possibility is that cells of the endothelial lineage are more plastic than is widely thought. Recent work has shown that L-EPCs can trans-differentiate to a smooth muscle-like phenotype when exposed to hypoxia or transforming growth factor- β 1 [30]. In addition, endothelial plasticity has been demonstrated in a number of reports that have shown that these cells can transdifferentiate into a variety of mesenchymal fates, including osteogenic, chondrogenic, and adipogenic identities [31–34]. Additional studies will be necessary to uncover the molecular mechanisms by which the biology of L-EPC can influence the efficiency of direct reprogramming.

A further question is why EPC-iPSCs appear more karyotypically similar to their parent L-EPC lines than previous studies

have reported for fibroblast-iPSCs. Although we did not make a direct comparison between fibroblast- and L-EPC-derived iPSCs in the present study, previous reports have shown that the vast majority of fibroblast-iPSCs accumulate significant CNVs compared with their parent fibroblast line [3–5]. We generated EPC-iPSCs and performed aCGH analyses on these lines using methods comparable to those used in one of these reports [5]. In contrast to these analyses of fibroblast-iPSCs, EPC-iPSCs generally have fewer CNVs compared with their parent L-EPC line. Interestingly in comparison with fibroblasts, which represent a polyclonal population, individual L-EPC lines arise from a single clone or a small number of clonal colonies (fewer than seven per line); thus, the normality of iPSCs derived from L-EPC lines may reflect the relatively close lineage history of the iPSCs and their progenitor cells, rather than unique properties of L-EPCs. Either way, L-EPCs provide a potentially powerful tool for refining the reprogramming process for the production of genetically healthy iPSCs for clinical applications. The lack of consistency in the genomic analysis of L-EPCs versus iPSCs and iPSCs versus monocytes demonstrates the importance of the selected reference genome in reaching conclusions about the genome under assessment. The L-EPC versus iPSC comparisons were likely to be mostly normal because of the oligoclonal nature of EPCs and the fact that the resultant iPSCs were separated from their parental cells by relatively few cell divisions. The monocyte versus iPSC comparisons reflect the difference between clonal iPSCs and an average polyclonal somatic genome. Underlying somatic variation may make the derivation of "genetically pure" iPSCs impossible [26–28], since such a state might not exist in vivo [27]. Additional studies will be necessary to define the natural occurrence of genomic changes in somatic cells and to understand the significance and consequences of these anomalies on derived iPSC lines. In addition, further genetic and epigenetic profiling, such as exome sequencing, will be necessary to fully assess the usefulness of EPC-iPSCs in regenerative applications. The possibility of deriving clonal L-EPC lines

could greatly facilitate these studies since it will simplify the evaluation of the resulting variation in iPSCs by avoiding the use of potentially heterogeneous populations of somatic cells.

In the current study retrovirus vectors were used to generate EPC-iPSCs. This was necessary to allow us to compare the utility of L-EPCs as a reprogramming substrate relative to previous studies, both in terms of reprogramming efficiencies and genetic analysis. iPSCs generated using retroviral vectors can be used effectively in drug/toxicology screens and in in vitro disease modeling, but transgene-free generation of EPC-iPSC will need to be optimized for clinical applications. A variety of new methods have been developed, such as Sendai virus [5], modified RNA [35], and episomes [36], each with varying degrees of success in different cell types and in different laboratories. Thus, a major future goal will be to systematically test these and other methods with L-EPCs to generate more clinically relevant EPC-iPSCs.

CONCLUSION

L-EPCs are readily obtainable from peripheral blood, with minimal manipulation. Their use as a nuclear reprogramming substrate allows the routine, efficient, and potentially high-throughput generation of iPSCs, with the majority of iPSC lines having a karyotype matching that of their parent somatic line.

ACKNOWLEDGMENTS

This work was funded by the British Heart Foundation, Medical Research Council, Wellcome Trust, and Fondation Leducq and

supported by the Cambridge Cardiovascular Consortium, Cambridge NIHR Biomedical Research Centre, and Cambridge Stem Cell Institute. A.A.R. was also supported by a Pfizer European Young Investigator Award. The authors thank Dr. J. Trotman and Dr. R.J.L. Treacy for assistance with sequencing.

AUTHOR CONTRIBUTIONS

I.G., M.T., M.M., J.N., W.M., and M.S.: experiment design, collection of data, data analysis and interpretation; M.L.O.: experiment design, collection of data, data analysis and interpretation, manuscript writing; F.R.: financial support, experiment design, collection of data, data analysis and interpretation; A.B.: financial support, experiment design, data analysis and interpretation, manuscript writing, final approval of manuscript; A.A.R.: conception and design, financial support, collection and assembly of data, data analysis and interpretation, manuscript writing, final approval of manuscript; L.V.: conception and design, financial support, data analysis and interpretation, manuscript writing, final approval of manuscript; N.W.M.: conception and design, financial support, provision of study material and patients, data analysis and interpretation, manuscript writing, final approval of manuscript.

DISCLOSURE OF POTENTIAL CONFLICTS OF INTEREST

The authors indicate no potential conflicts of interest.

REFERENCES

- 1 Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 2006;126:663–676.
- 2 Takahashi K, Tanabe K, Ohnuki M et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 2007;131:861–872.
- 3 Hussein SM, Batada NN, Vuoristo S et al. Copy number variation and selection during reprogramming to pluripotency. *Nature* 2011;471:58–62.
- 4 Martins-Taylor K, Nisler BS, Taapken SM et al. Recurrent copy number variations in human induced pluripotent stem cells. *Nat Biotechnol* 2011;29:488–491.
- 5 Yusa K, Rashid ST, Strick-Marchand H et al. Targeted gene correction of α 1-antitrypsin deficiency in induced pluripotent stem cells. *Nature* 2011;478:391–394.
- 6 Loh YH, Agarwal S, Park IH et al. Generation of induced pluripotent stem cells from human blood. *Blood* 2009;113:5476–5479.
- 7 Ye Z, Zhan H, Mali P et al. Human-induced pluripotent stem cells from blood cells of healthy donors and patients with acquired blood disorders. *Blood* 2009;114:5473–5480.
- 8 Seki T, Yuasa S, Oda M et al. Generation of induced pluripotent stem cells from human terminally differentiated circulating T cells. *Cell Stem Cell* 2010;7:11–14.
- 9 Loh YH, Hartung O, Li H et al. Reprogramming of T cells from human peripheral blood. *Cell Stem Cell* 2010;7:15–19.
- 10 Staerk J, Dawlaty MM, Gao Q et al. Reprogramming of human peripheral blood cells to induced pluripotent stem cells. *Cell Stem Cell* 2010;7:20–24.
- 11 Chou BK, Mali P, Huang X et al. Efficient human iPSC cell derivation by a non-integrating plasmid from blood cells with unique epigenetic and gene expression signatures. *Cell Res* 2011;21:518–512.
- 12 Okabe M, Otsu M, Ahn DH et al. Definitive proof for direct reprogramming of hematopoietic cells to pluripotency. *Blood* 2009;114:1764–1767.
- 13 Eminli S, Foudi A, Stadtfeld M et al. Differentiation stage determines potential of hematopoietic cells for reprogramming into induced pluripotent stem cells. *Nat Genet* 2009;41:968–976.
- 14 Haase A, Olmer R, Schwanke K et al. Generation of induced pluripotent stem cells from human cord blood. *Cell Stem Cell* 2009;5:434–441.
- 15 Ho PJ, Yen ML, Lin JD et al. Endogenous KLF4 expression in human fetal endothelial cells allows for reprogramming to pluripotency with just OCT3/4 and SOX2: Brief report. *Arterioscler Thromb Vasc Biol* 2010;30:1905–1907.
- 16 Panopoulos AD, Ruiz S, Yi F et al. Rapid and highly efficient generation of induced pluripotent stem cells from human umbilical vein endothelial cells. *PLoS One* 2011;6:e19743.
- 17 Mack AA, Kroboth S, Rajesh D et al. Generation of induced pluripotent stem cells from CD34+ cells across blood drawn from multiple donors with non-integrating episomal vectors. *PLoS One* 2011;6:e27956.
- 18 Chou BK, Mali P, Huang X et al. Efficient human iPSC cell derivation by a non-integrating plasmid from blood cells with unique epigenetic and gene expression signatures. *Cell Res* 2011;21:518–529.
- 19 Hu K, Yu J, Suknutha K et al. Efficient generation of transgene-free induced pluripotent stem cells from normal and neoplastic bone marrow and cord blood mononuclear cells. *Blood* 2011;117:e109–e119.
- 20 Lin Y, Weisdorf DJ, Solovey A et al. Origins of circulating endothelial cells and endothelial outgrowth from blood. *J Clin Invest* 2000;105:71–77.
- 21 Medina R, O'Neill CL, Humphreys MW et al. Outgrowth endothelial cells: Characterization and their potential for reversing ischemic retinopathy. *Invest Ophthalmol Vis Sci* 2010;51:5906–5913.
- 22 Yoder M, Mead LE, Prater D et al. Redefining endothelial progenitor cells via clonal analysis and hematopoietic stem/progenitor cell principals. *Blood* 2007;109:1801–1809.
- 23 Piaggio G, Rosti V, Corselli M et al. Endothelial colony forming cells (ECFCs) from patients with chronic myeloproliferative disorders lack the disease-specific molecular clonality marker. *Blood* 2009;114:3127–3130.
- 24 Vallier L, Touboul T, Brown S et al. Signaling pathways controlling pluripotency and early cell fate decisions of human induced pluripotent stem cells. *STEM CELLS* 2009;27:2655–2666.
- 25 International PPH Consortium, Lane KB, Machado RD et al. Heterozygous germline mutations in BMPR2, encoding a TGF-beta receptor, cause familial primary pulmonary hypertension. *Nat Genet* 2000;26:81–84.

26 Laurent LC, Ulitsky I, Slavin I et al. Dynamic changes in the copy number of pluripotency and cell proliferation genes in human ESCs and iPSCs during reprogramming and time in culture. *Cell Stem Cell* 2011;8:106–118.

27 Liang Q, Conte N, Skarnes WC et al. Extensive genomic copy number variation in embryonic stem cells *Proc Natl Acad Sci USA* 2008; 105:17453–17456.

28 International Stem Cell Initiative, Amps K, Andrews PW et al. Screening ethnically diverse human embryonic stem cells identifies a chromosome 20 minimal amplicon conferring growth advantage. *Nat Biotechnol* 2011;29:1132–1144.

29 de Cid R, Riveira-Munoz E, Zeeuwen PL et al. Deletion of the late cornified envelope

LCE3B and LCE3C genes as a susceptibility factor for psoriasis. *Nat Genet* 2009;41:211–215.

30 Imamura H, Ohta T, Tsunetoshi K et al. Transdifferentiation of bone marrow-derived endothelial progenitor cells into the smooth muscle cell lineage mediated by transforming growth factor-beta1. *Atherosclerosis* 2010; 211:114–121.

31 Medici D, Shore EM, Lounev VY et al. Conversion of vascular endothelial cells into multipotent stem-like cells. *Nat Med* 2010;16: 1400–1406.

32 Wylie-Sears J, Aikawa E, Levine RA et al. Mitral valve endothelial cells with osteogenic differentiation potential. *Arterioscler Thromb Vasc Biol* 2011;31:598–607.

33 Tran KV, Gealekman O, Frontini A et al. The vascular endothelium of the adipose tissue gives rise to both white and brown fat cells. *Cell Metab* 2012;15:222–229.

34 Gupta RK, Mepani RJ, Kleiner S et al. Zfp423 expression identifies committed preadipocytes and localizes to adipose endothelial and perivascular cells. *Cell Metab* 2012;15:230–239.

35 Warren L, Manos PD, Ahfeldt T. Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. *Cell Stem Cell* 2010;7: 618–630.

36 Okita K, Matsumura Y, Sato Y et al. A more efficient method to generate integration-free human iPSCs. *Nat Methods* 2011;8:409–412.



See www.StemCellsTM.com for supporting information available online.