



SUBJECT AREAS:

REGENERATIVE
MEDICINEINDUCED PLURIPOTENT STEM
CELLS

REPROGRAMMING

STEM-CELL BIOTECHNOLOGY

Optimized Surface Markers for the Prospective Isolation of High-Quality hiPSCs using Flow Cytometry Selection

Ramzey Abujarour*, Bahram Valamehr*, Megan Robinson, Betsy Rezner, Florin Vranceanu & Peter Flynn

Fate Therapeutics Inc., 3535 General Atomics Court Suite 200, San Diego, CA 92121, USA.

Received
26 October 2012Accepted
10 January 2013Published
31 January 2013

Correspondence and requests for materials should be addressed to P.F. (peter.flynn@fatetherapeutics.com)

* These authors contributed equally to this work.

hiPSC derivation and selection remains inefficient; with selection of high quality clones dependent on extensive characterization which is not amenable to high-throughput (HTP) approaches. We recently described the use of a cocktail of small molecules to enhance hiPSC survival and stability in single cell culture and the use of flow cytometry cell sorting in the HTP-derivation of hiPSCs. Here we report an enhanced protocol for the isolation of bona fide hiPSCs in FACS-based selection using an optimized combination of cell surface markers including CD30. Depletion of CD30⁺ cells from reprogramming cultures almost completely abolished the NANOG and OCT4 positive sub-population, suggesting it is a pivotal marker of pluripotent cells. Combining CD30 to SSEA4 and TRA-1-81 in FACS greatly enhanced specificity and efficiency of hiPSC selection and derivation. The current method allows for the efficient and automated, prospective isolation of high-quality hiPSC from the reprogramming cell milieu.

The recent development of human induced pluripotent stem cells (hiPSCs) from somatic cells through the exogenous expression of transcription factors such as OCT4, SOX2, KLF4, and MYC^{1,2}, enables the generation of pluripotent cells from any genetic background. hiPSCs represent an attractive platform for disease modeling, drug discovery and ultimately cell therapy³, however several improvements to the reprogramming process are required prior to effective industrial and clinical applications. Efficiencies for the generation of hiPSCs vary greatly depending on the reprogramming methodology applied (i.e. delivery vehicle, number of reprogramming factors, and the use of feeder cells), but the reprogramming process usually results in the generation of a small sub-population of successfully reprogrammed hiPSCs (generally < 1% input cell number) in a largely heterogeneous population of non-iPSCs⁴.

Manual picking of hiPSC colonies based on morphology (3D growth of colonies with sharp edges, and high nucleus to cytoplasm ratio) is the current method of choice for the isolation of the rare hiPSC sub-population during the reprogramming process. Staining live reprogramming cell cultures with surface markers can be used as an additional selection criterion^{5,6} and the analysis of selected clones using multiple markers has been shown to be essential to distinguish the fully reprogrammed pluripotent state from the partially reprogrammed intermediates⁶. However, manual picking as a tool for the selection of bona fide hiPSCs remains time consuming, requires specifically trained operators and is not amenable to high-throughput manipulation⁷. We have previously described a high-throughput platform for the selection of clonal hiPSCs and maintenance in feeder-free cultures based on fluorescence activated cell sorting (FACS) and a small chemical media additive (SMC4)⁸. The use of FACS permits the efficient and automated derivation of a large number of hiPSC clones and subclones from any given reprogramming experiment and lowers the workload and technical barriers for multiplex derivation of hiPSC from multiple donors. Such a system represents a meaningful step towards large-scale studies of disease-specific phenotypes and hiPSC banking. The use of FACS as the method for hiPSC selection eliminates morphology as a criterion, instead relying on the specificity of the surface markers used. For the hiPSC selection process to be efficient, it is critical that the surface marker(s) selected are extremely specific to human pluripotent stem cells (hPSCs) and can discriminate successfully reprogrammed hiPSCs from partially reprogrammed or differentiated cells.

The most common surface markers used to distinguish hPSCs are SSEA3, SSEA4, TRA-1-60, and TRA-1-81⁹. The expression of SSEA3 and SSEA4 by reprogramming cells usually precedes the expression of TRA-1-60 and TRA-1-81, which are detected only at later stages of reprogramming⁶. It has been proposed that the antibodies specific for the TRA-1-60 and TRA-1-81 antigens recognize distinct and unique epitopes on the same large glycoprotein Podocalyxin (also called podocalyxin-like, PODXL)¹⁰. Other surface modifications including the



presence of specific lectins have also been shown to distinguish hiPSCs from non-hiPSCs¹¹. Cluster of differentiation (CD) molecules are proteins expressed on the surface of many cell types, and function as integrins, adhesion molecules, glycoproteins, and receptors⁹. CD molecules have been used extensively as markers of several types of adult stem cells such as CD34 for hematopoietic stem cells and progenitors¹². However, CD molecules are less commonly used to label hPSCs, even though several CD molecules have been associated with pluripotency such as CD30 (tumor necrosis factor receptor superfamily, member 8, TNFRSF8), CD9 (leukocyte antigen, MIC3), CD50 (intercellular adhesion molecule-3, ICAM3), CD200 (MRC OX-2 antigen, MOX2) and CD90 (Thy-1 cell surface antigen, THY1)^{13–18}. The true utility of these cell surface markers in the isolation of hiPSCs has yet to be investigated. Most, if not all of the cell surface markers currently used for the characterization of hiPSCs are also expressed in normal and malignant tissues. For example, SSEA4 is expressed by dorsal root ganglion cells¹⁹ and in fetal forebrain and spinal cord^{20,21}. Podocalyxin is expressed in multiple tissues, with highest expression in kidney, pancreas, and heart²² and multiple cancers²³. CD30 is seen to be expressed on activated T and B cells and in certain hematopoietic malignancies^{24,25}. It is therefore likely that combinations of markers will provide the most specificity in hiPSC selection.

In this study, we have identified specific combinations of pluripotency surface markers for the isolation of bona fide hiPSCs using FACS. A survey of a panel of CD molecules revealed CD30 as a specific pluripotency marker that distinguishes fully reprogrammed hiPSCs from other reprogramming derivatives. Depleting the CD30-positive cells from a reprogramming human cell culture almost completely eliminates the hiPSCs in that culture, suggesting that CD30 is a pivotal marker of pluripotent cells. Incorporating CD30 into our previously established, SSEA4 and TRA-1-81 based FACS workflow produced a highly enhanced enrichment of hiPSCs, compared to using SSEA4 and TRA-1-81 alone. We show that hiPSCs isolated using FACS and the surface markers SSEA4, TRA-1-81 and CD30 are pluripotent as determined by extensive expression analyses and tri-lineage differentiation. Furthermore, the hiPSC lines generated using this protocol, are seen to maintain a stable genome over prolonged culture in feeder-free conditions.

Results

Identifying cell surface markers for the isolation of fully reprogrammed hiPSCs. hPSCs grow and are routinely passaged as clumps of cells, a characteristic that confers survival and maintenance of pluripotency⁷. Manipulation of hPSCs via FACS, which requires cells to be completely dissociated, is therefore challenging. We have previously described a chemical media additive (SMC4) for the enhanced survival and increased clonogenicity of hiPSCs following single cell dissociation, allowing the use of FACS for the derivation of clonal hiPSCs under feeder-free conditions⁸. This system combined the conventional pluripotency cell surface markers SSEA4 and TRA-1-81 in the selection of hiPSC from the reprogramming cellular milieu. Clonal and subclonal selection at single cell per well into a 96-well plate directly from FACS results in a panel of potential hiPSC lines which can then be further characterized for their pluripotent state. We noticed however that not every cell pre-selected in this way resulted in a robust, fully reprogrammed iPSC line on full characterization, suggesting that additional selection criteria would improve the platform. For example, in a reprogramming experiment where human neonatal fibroblasts (FTc1) were infected with a lentivirus containing a polycistronic cassette expressing OCT4 (POU5F1), SOX2 and KLF4, derivation of clonal hiPSC lines was completed by sorting SSEA4 and TRA-1-81 double positive cells into Matrigel-coated 96-well plates at clonal densities (Supplementary Fig. 1). Many cell lines were carried forward and confirmed to be SSEA4 and TRA-1-81 double positive (Fig. 1a). However, gene

expression analysis showed that some, but not all cells selected as SSEA4 and TRA-1-81 double positive, expressed the pluripotency markers *NANOG*, *OCT4* and *SOX2* (Fig. 1b). Further, immunofluorescence analysis of *NANOG* expression confirmed the poor correlation between *NANOG* and TRA-1-81 in some hiPSC clones (Fig. 1c). Therefore, while conventional pluripotency surface markers SSEA4 and TRA-1-81 can aid in the isolation of bona fide hiPSCs using FACS, additional characterization is required to distinguish between fully reprogrammed and partially reprogrammed hiPSC lines.

To enhance the efficiency and specificity of true hiPSCs isolation by FACS, we screened additional cell surface proteins (CD30, CD9, CD50, CD200, and CD90) that have been suggested to be markers of pluripotency or expressed in hPSCs^{13–18,26}. The expression profiles of these cell surface proteins along with the traditional pluripotency markers: *Podocalyxin* (the carrier of TRA-1-60 and TRA-1-81 antigens), *OCT4*, *NANOG* and the fibroblast marker *CD13* were examined in a wide array of cell types. The cell types tested included hiPSCs cultured on feeder-free conditions (Matrigel and SMC4 media), hPSCs (hESCs and hiPSCs) cultured in conventional hESC media on growth-arrested mouse embryonic fibroblasts (MEFs), partially reprogrammed cell lines that have failed to activate endogenous *NANOG* and *OCT4*, hPSC-derived differentiated cells and adult primary cells (fibroblasts, adipocytes and myoblasts). As expected *Podocalyxin* was expressed in hPSCs, and was undetected in fibroblasts (Fig. 2a). However, *Podocalyxin* was expressed at high levels in the hPSC-derived differentiated cells and primary myoblasts. *Podocalyxin* was also expressed in the partially reprogrammed cell lines, validating the results obtained by flow cytometry and immunofluorescence analyses (Figs. 1 and 2a). While undetected in somatic cells, *CD30* was expressed at higher levels in hPSCs compared to partially reprogrammed cell lines and hPSC-derived differentiated cells. Relative to *Podocalyxin*, *CD30* showed a more dynamic expression window that largely mirrored that of *NANOG* (Fig. 2a). *CD200* also showed an expression profile that was relatively specific to hPSCs and similar to the expression profiles of *CD30* and *Podocalyxin*. The remaining surface markers were less specific in labeling hPSCs. Although high in fibroblasts, *CD90* appeared to be suppressed in the partially reprogrammed cell lines (Fig. 2a). *CD30* and *CD200* were undetected in the 3 different primary somatic cell types examined (fibroblasts, adipocytes, and myoblasts). To further investigate their specificity, a panel of 48 different normal human tissues was assayed for the expression of *CD30*, *CD200*, *CD9* and *Podocalyxin* by qRT-PCR (Fig. 2b). *CD30* showed the most restrictive expression pattern, limited to blood, stomach, and the mammary gland. *CD200* and *Podocalyxin* were expressed in the majority of tissues examined and *CD9* displayed the widest range of expression. Intriguingly, *CD30* expression did not overlap with that of *Podocalyxin* (e.g. *Podocalyxin* is not expressed in lymphocytes), suggesting that these two markers could provide better specificity when used together in the selection of hiPSCs.

In addition, we tested the panel of surface markers by flow cytometry. A previously selected hiPSC line seen to be partially differentiated during routine culture was stained with combinations of three surface markers: SSEA4 and TRA-1-81, and either TRA-1-60, CD200, CD90, CD9, CD50, or CD30. We investigated the ability of the different surface markers to discriminate between the SSEA4⁺ TRA-1-81⁺ cell pool (the undifferentiated pool) and the SSEA4⁻ TRA-1-81⁻ cell pool (the differentiated pool). The TRA-1-60 antibody labeled the double positive pool almost exclusively, confirming that TRA-1-60 and TRA-1-81 are redundant *Podocalyxin* markers¹⁰. Among the markers tested, only CD30 clearly separated the two different cell pools (Fig. 3). Taken together, these results indicate that CD30 is a specific marker of pluripotency, with minimal expression on differentiated cell types.

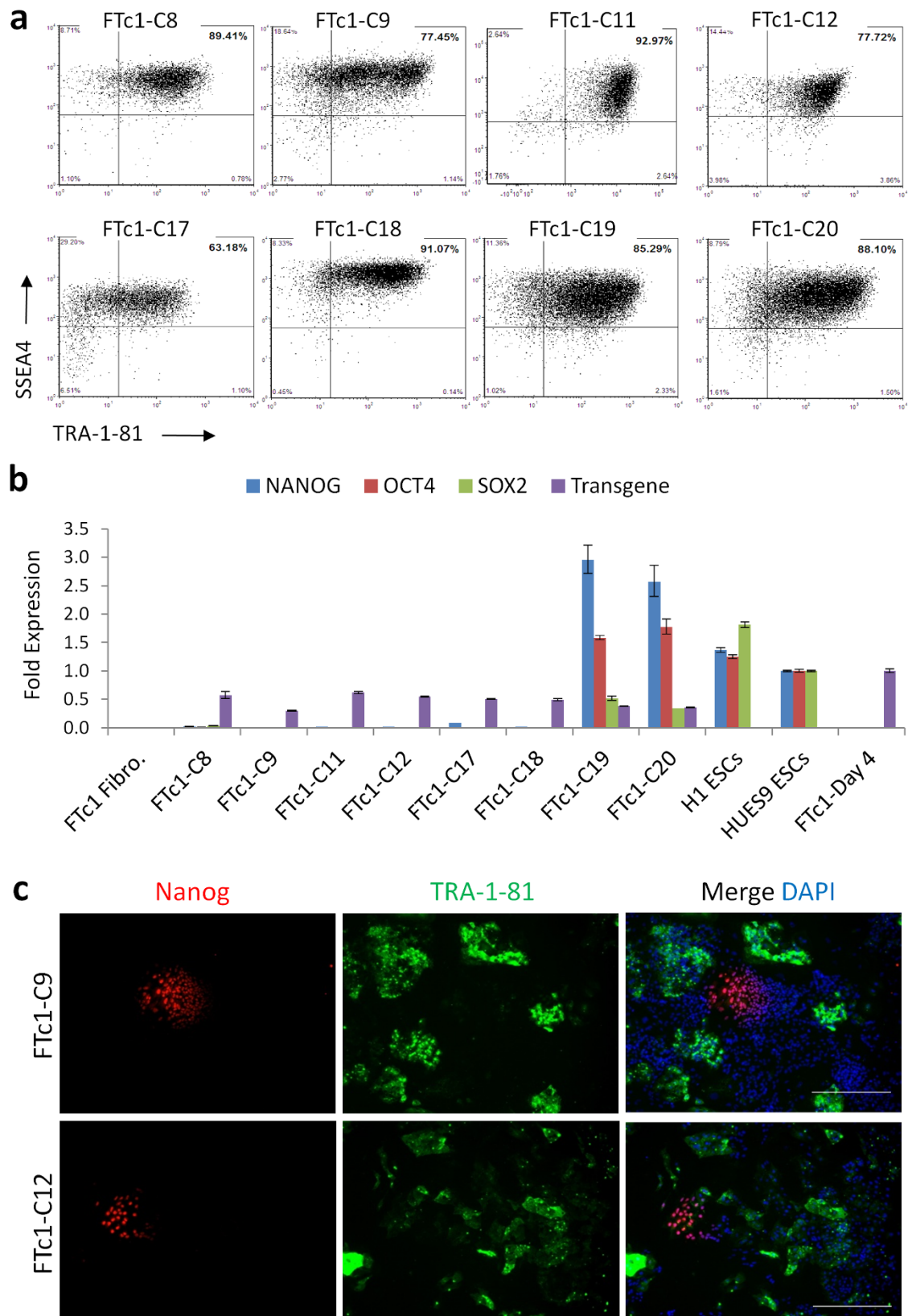


Figure 1 | Conventional pluripotency markers SSEA4 and TRA-1-81 do not discriminate between partially and fully reprogrammed hiPSCs. (a) Flow cytometry analysis of hiPSC lines generated by sorting human FTc1 fibroblasts directly into 96-well plate in SMC4 media 6 weeks post infection with polyclonic reprogramming lentivirus (expressing OCT4, SOX2, and KLF4). (b) SSEA4⁺ TRA-1-81⁺ cell lines were analyzed by qRT-PCR for the endogenous expression of pluripotency markers (*NANOG*, *OCT4* and *SOX2*) and transgene. Error bars represent standard deviation of duplicates. Differences in gene expression of pluripotency markers in all FTc1 cell lines are significant (p -values < 0.05) relative to the H1 ESCs sample. (c) SSEA4⁺ TRA-1-81⁺ cell lines FTc1-C9 and FTc1-C12 were stained for expression of TRA-1-81 (green) and NANOG (red) and examined under fluorescence microscopy. Nuclei were stained with DAPI (blue). Scale bar is 400 μ m.

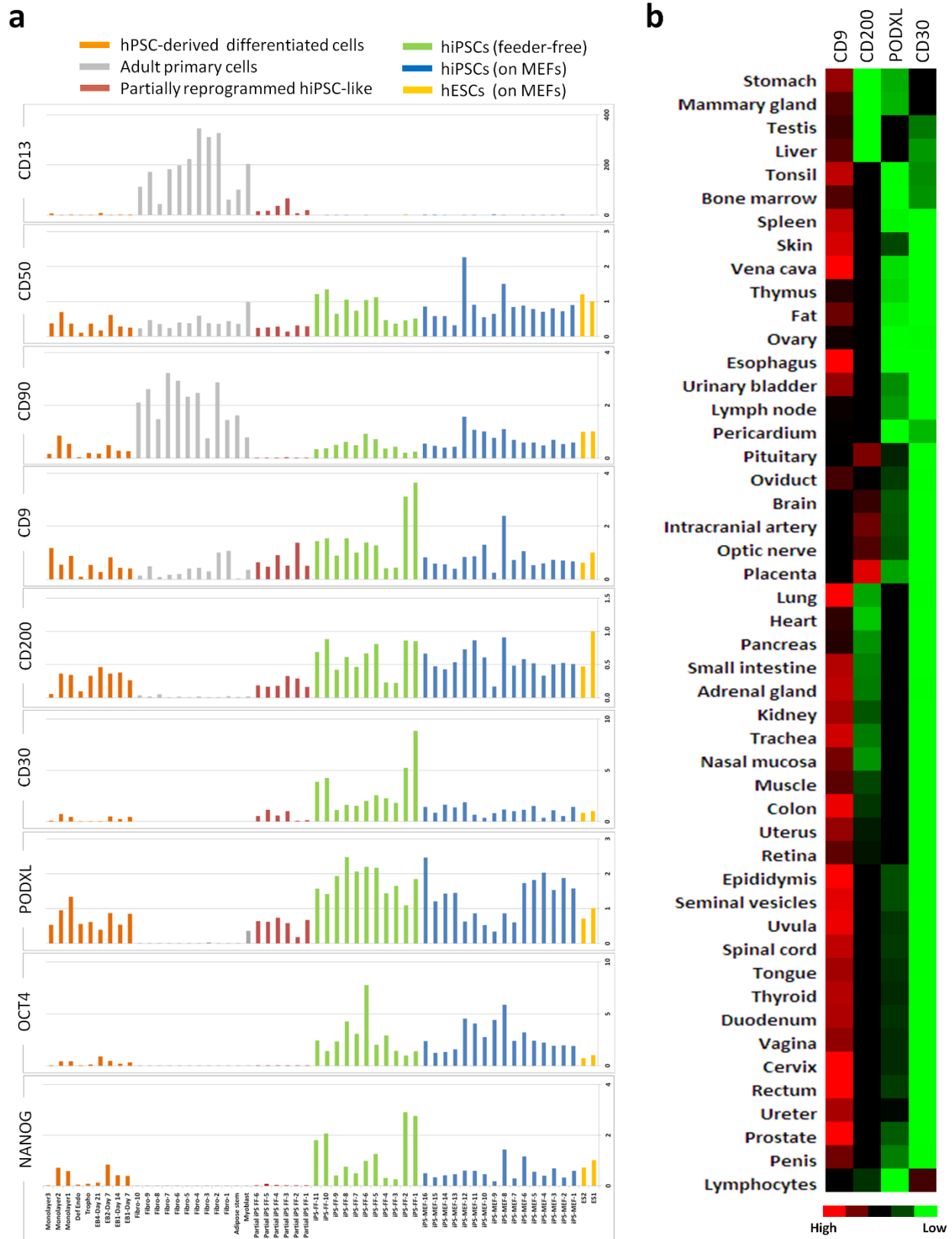


Figure 2 | Surface marker expression in hPSCs and somatic cells. (a) Expression of indicated surface markers was compared to that of NANOG and OCT4 in an array of cell types including hPSCs (hESCs; $n = 2$, hiPSCs; $n = 16$) cultured in conventional media and on MEFs, hiPSCs cultured in SMC4 media on Matrigel ($n = 11$), partially reprogrammed cells cultured in SMC4 media on Matrigel ($n = 6$), primary adult cells (fibroblasts, $n = 10$; adipose stem cells, $n = 1$; myoblasts, $n = 1$), and hPSC-derived differentiated cells (embryoid bodies, $n = 4$; definitive endoderm, $n = 1$; trophoblast, $n = 1$; monolayer differentiation, $n = 3$). Expression of indicated genes was examined by qRT-PCR, and values normalized within each set to that of hESC. (b) Expression of indicated surface markers was examined by qRT-PCR in normal tissue samples ($n = 48$) and results are depicted as a heat map. Distances between samples and assays are calculated for hierarchical clustering based on the Δ CT values using the Pearson's Correlation. The Δ CT range values are between -3 (High) and 15 (Low).



CD30 distinguishes fully reprogrammed hiPSCs from reprogramming intermediates and derivatives. To further explore the application of CD30 as a tool for the selection of hiPSCs, we used flow cytometry to examine its expression in the SSEA4 and TRA-1-81 double positive cell lines that failed to initiate the endogenous pluripotency program, as indicated by lack of robust, homogeneous NANOG and OCT4 expression. In four of these partially reprogrammed cell lines, 52–84% of the cells were double positive for SSEA4 and TRA-1-81, whereas only 0.6–20% of the cells were positive for CD30 (Supplementary Fig. 2a). In contrast, a fully reprogrammed hiPSC line (FTc1-C19) was seen to consist of over 90% CD30 positive cells (Fig. 4a and Supplementary Fig. 2a). In addition, immunofluorescence analysis validated the more restrictive expression of CD30

compared to that of TRA-1-81 in these partially reprogrammed cell lines, where most TRA-1-81⁺ cells were not CD30⁺ but most CD30⁺ cells were TRA-1-81⁺ (Supplementary Fig. 2b). To better understand the significance of the restrictive expression of CD30 in these cells, we examined CD30 and TRA-1-81 as they relate to the expression of NANOG. The partially reprogrammed cells were co-stained with antibodies specific to CD30, TRA-1-81 (or its equivalent TRA-1-60) and NANOG and examined using immunofluorescence. Whereas TRA-1-81 and TRA-1-60 labeled both NANOG⁺ and NANOG⁻ cells, there was an almost complete correlation between CD30 expression and that of NANOG (Fig. 4a and Supplementary Fig. 3a). These observations were also confirmed using two additional partially reprogrammed cell lines derived from a

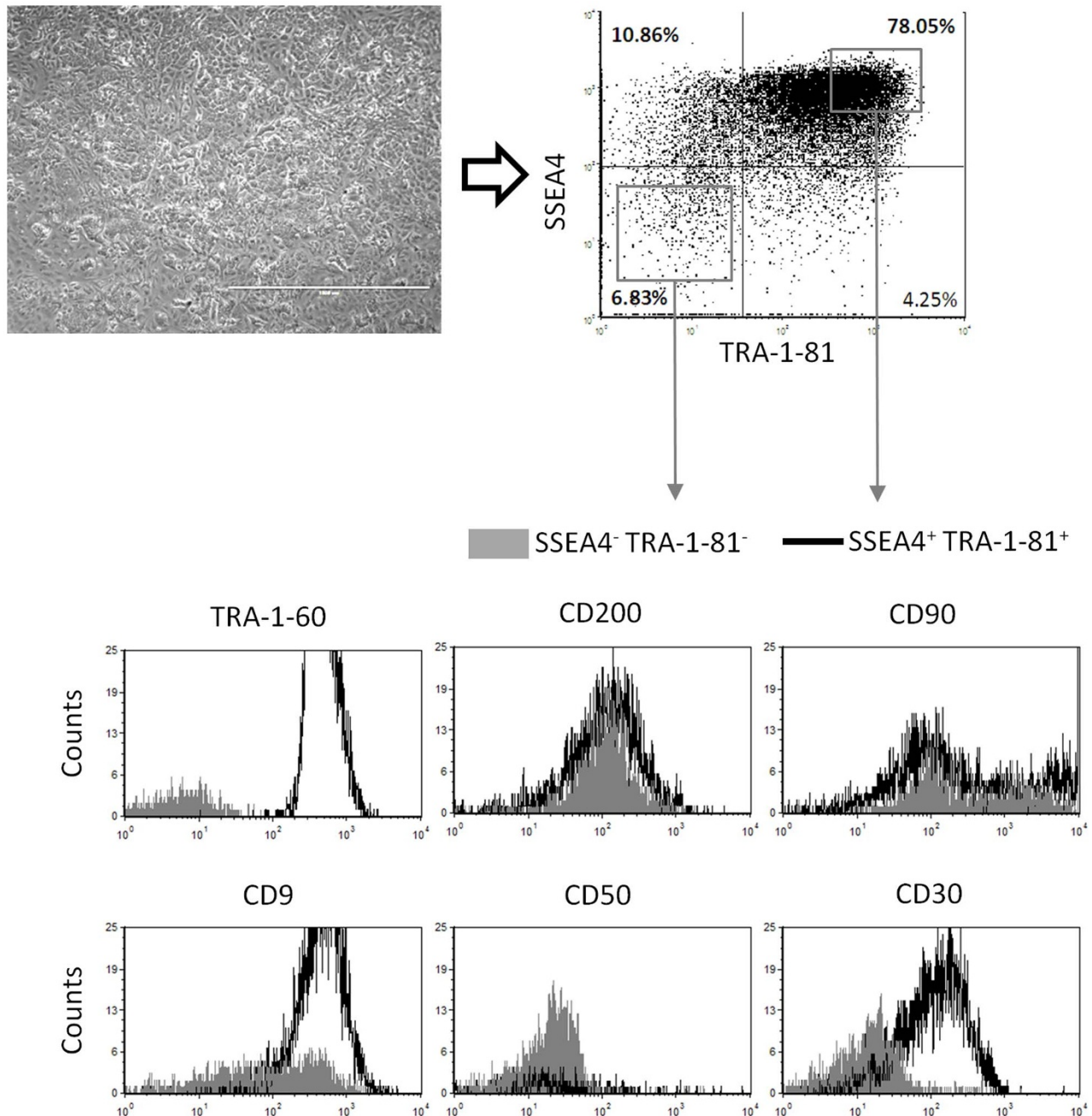


Figure 3 | Screening CD molecules for ability to identify iPSCs from non-iPSC contaminants reveals CD30 as a specific pluripotency marker. A partially differentiated hiPSC line (upper left) was stained with a combination of three surface markers: SSEA4 and TRA-1-81, and either TRA-1-60, CD200, CD90, CD9, CD50, or CD30. Two sub-populations were gated as shown (upper right): SSEA4⁻ TRA-1-81⁻ (non-iPSCs; grey area) and SSEA4⁺ TRA-1-81⁺ (iPSCs; black line). Histograms show expression of the indicated markers within each gate.

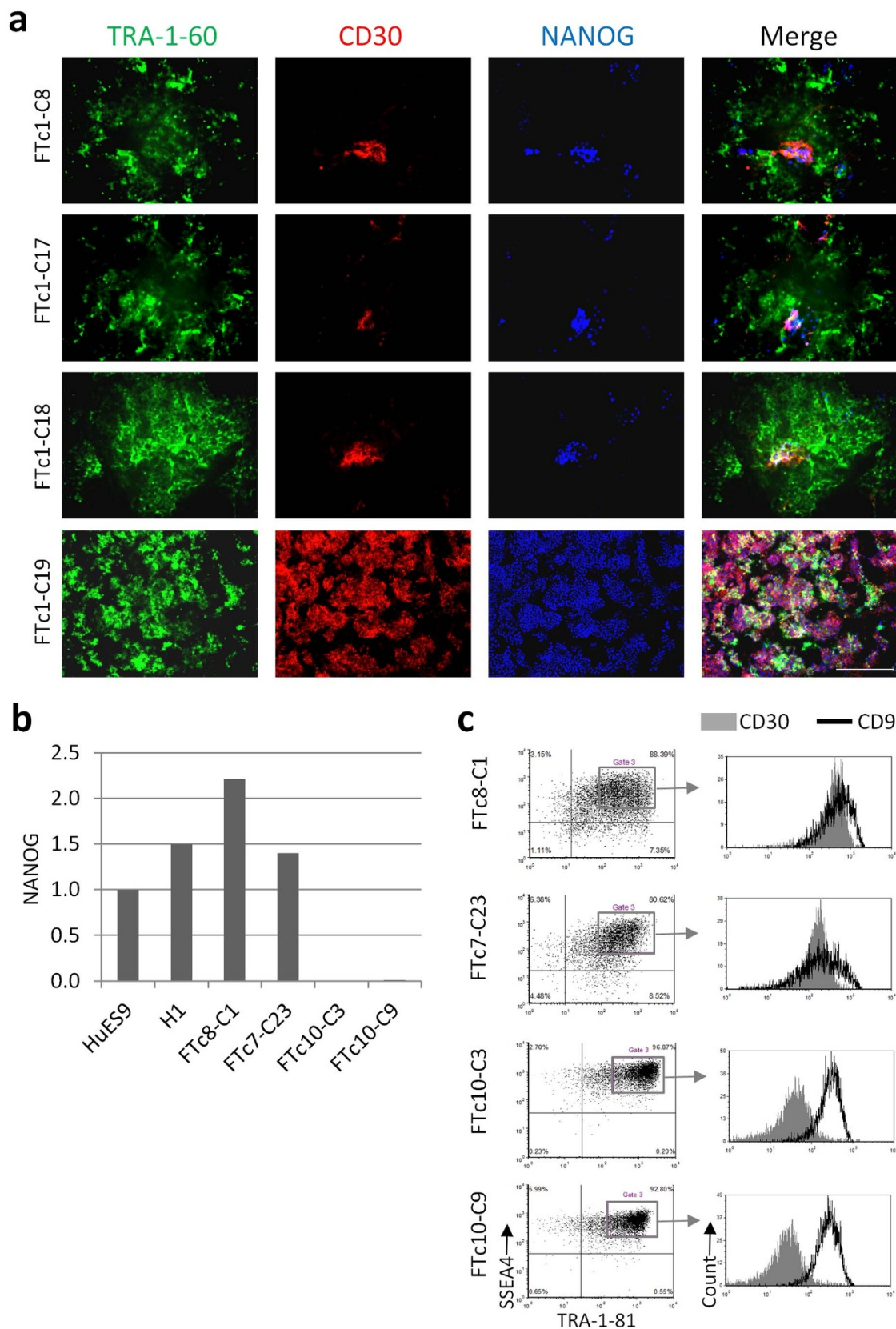


Figure 4 | CD30 surface marker correlates with NANOG expression. (a) Three partially reprogrammed cell lines FTc1-C8, FTc1-C17 and FTc1-C18, and one fully reprogrammed hiPSC line FTc1-c19 were stained for the surface markers TRA-1-60 (green), CD30 (red) and NANOG (blue), and examined under fluorescence microscopy. Scale bar is 200 μ m. (b) Four cell lines generated by sorting SSEA4⁺ TRA-1-81⁺ cells were analyzed by qRT-PCR for expression of NANOG. (c) The two NANOG⁺ cell lines FTc8-C1 and FTc7-C23 and the two NANOG⁻ cell lines FTc10-C3 and FTc10-C9 were analyzed by flow cytometry for expression of the surface markers: SSEA4, TRA-1-81, CD30, and CD9. The histograms (right) depict the gated population of cells in the dot plots (left).



different human fibroblast line (FTc10). The two cell lines (FTc10-C3 and FTc10-C9) were positive for SSEA4 and TRA-1-81, but again failed to initiate endogenous *NANOG* expression as assessed by qRT-PCR (Fig. 4b). Flow cytometry analysis revealed that the two partially reprogrammed cell lines were largely SSEA4⁺ and TRA-1-81⁺ but CD30⁻, whereas two fully reprogrammed hiPSC lines (FTc7-C23 and FTc8-C1), with high levels of *NANOG*, were largely positive for all 3 surface markers (Figs. 4b,c and Supplementary Fig. 3b). The ability of CD30 to distinguish between the partially and fully reprogrammed cell lines was in contrast to that of CD9, which stained all 4 cell lines equally well. These results suggest that CD30 is a potentially powerful marker for the selection of pluripotent cells,

essentially acting as a cell surface surrogate for *NANOG* expression when combined with SSEA4 and TRA-1-81.

Enhanced hiPSC isolation using the CD30 cell surface marker in FACS-Based selection. Based on our observations that CD30 markedly improves identification of cells with endogenous *NANOG* expression, we explored its use in FACS-based hiPSC selection. We first tested the effects, if any, of depleting or enriching CD30⁺ cells. We sorted two sub-populations: SSEA4⁺ TRA-1-81⁺ CD30⁺ cells and SSEA4⁺ TRA-1-81⁺ CD30⁻ cells from a partially differentiated hiPSC line (FTc8-C9) and surveyed the expression of a panel of pluripotency markers. Before sorting, the FTc8-C9 line

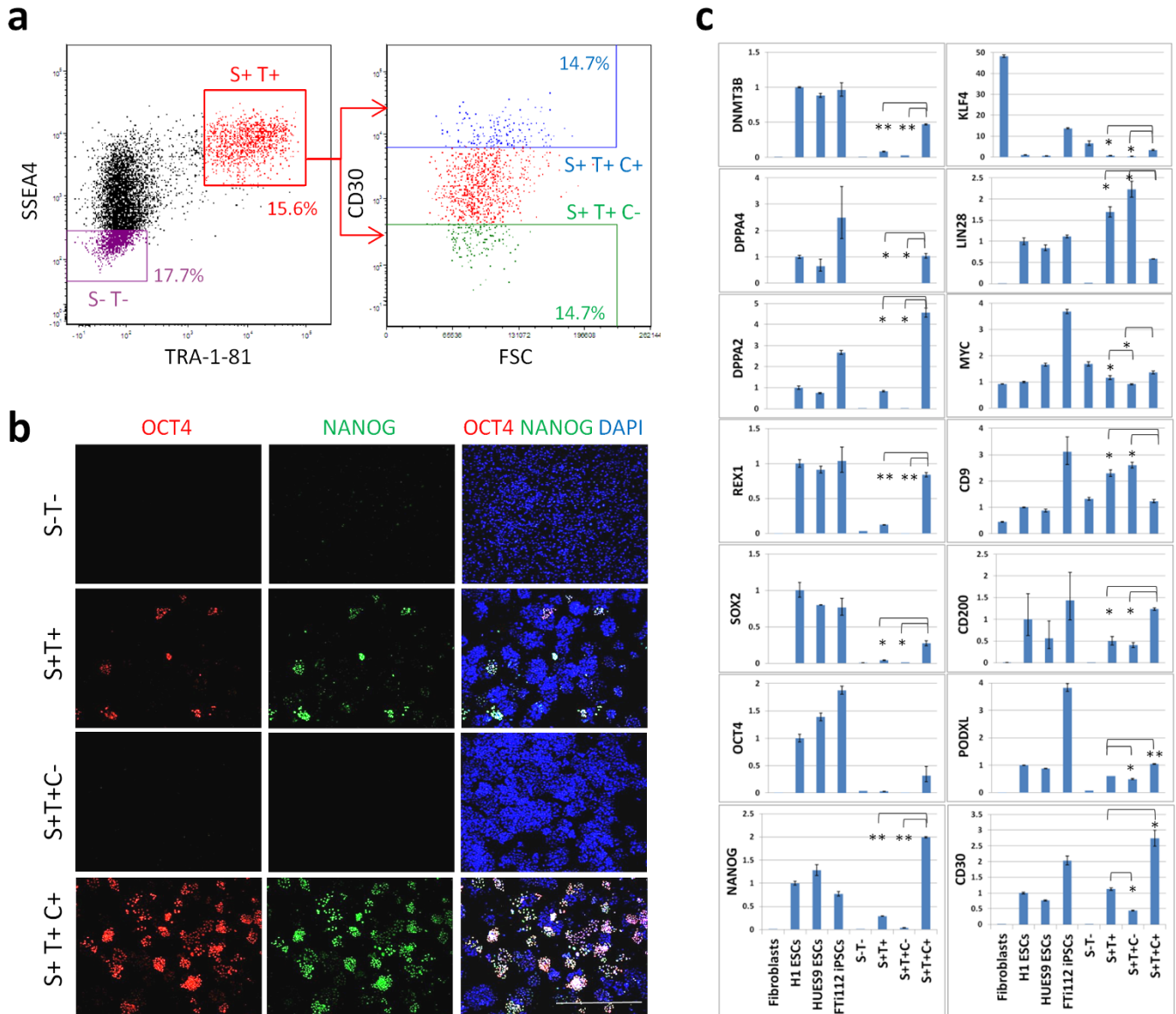


Figure 5 | Selection of hiPSCs from a reprogramming pool using FACS is highly enhanced by combining CD30 with the conventional pluripotency markers SSEA4 and TRA-1-81. (a) 37 days post initiation of reprogramming of FTc63 human fibroblasts, cells were stained for the surface markers: SSEA4, TRA-1-81, and CD30. Four different sub-populations of the reprogramming pool were separated by FACS as indicated: SSEA4⁻ TRA-1-81⁻ (S-T-), SSEA4⁺ TRA-1-81⁺ (S+T+), SSEA4⁺ TRA-1-81⁺ CD30⁻ (S+T+C-), SSEA4⁺ TRA-1-81⁺ CD30⁺ (S+T+C+). Sorted cells were collected and seeded on Matrigel-coated plate in SMC4 medium. (b) The 4 different sorted sub-populations were analyzed for the expression of OCT4 (red) and NANOG (green) by fluorescence microscopy. Nuclei were stained with DAPI (blue). Scale bar in merged color image is 1000 μ m. (c) The 4 different sorted sub-populations were analyzed by qRT-PCR for the endogenous expression of pluripotency markers, reprogramming factors, and indicated surface markers. Established hiPSC line FT1112 cultured in SMC4 on Matrigel, hESC lines H1 and HuES9 cultured in conventional media and on feeder cells were used as references. Podocalyxin (PODXL) is the carrier for TRA-1-60 and TRA-1-81. Error bars represent standard deviation of duplicates. The asterisks denote corresponding p-values (* < 0.05, ** < 0.001).



showed markedly reduced *NANOG* expression compared to hESCs (Supplementary Fig. 4a). Following the sort, the triple positive sub-population displayed levels of expression for *NANOG* and other pluripotency markers that are equivalent to hESCs (Supplementary Fig. 4b). In contrast, the CD30 depleted sub-population showed a significant reduction in the expression of pluripotency markers, and displayed a predominantly differentiated morphology (Supplementary Fig. 4b,c).

We next compared the ability of the two conventional pluripotency markers SSEA4 and TRA-1-81 either alone or when combined with CD30 to separate true hiPSCs from other reprogramming intermediates and derivatives during the reprogramming process. Five weeks post infection with the 3-factor reprogramming lentivirus, human neonatal fibroblasts (FTc63), were separated by FACS into four different sub-populations: SSEA4⁻ TRA-1-81⁻ (S⁻ T⁻), SSEA4⁺ TRA-1-81⁺ (S⁺ T⁺), SSEA4⁺ TRA-1-81⁺ CD30⁻ (S⁺ T⁺ C⁻), SSEA4⁺ TRA-1-81⁺ CD30⁺ (S⁺ T⁺ C⁺) (Fig. 5a). The 4 pools of cells were collected, seeded on Matrigel-coated plates and analyzed 4 days later by immunofluorescence and qRT-PCR for expression of a panel of pluripotency markers. As expected, the double negative sub-population was notably devoid of pluripotent cells when examined by microscopy and qRT-PCR (Fig. 5b, c). The double positive sub-population (S⁺ T⁺) contained a number of pluripotent colonies, whereas the triple positive sub-population (S⁺ T⁺ C⁺) contained significantly more *NANOG*⁺ and *OCT4*⁺ hiPSC colonies (Fig. 5b). Gene expression analysis of *CD30* and *Podocalyxin*, confirmed the effective FACS separation of different cell sub-populations (Fig. 5c). Additionally, the triple positive sub-population (S⁺ T⁺ C⁺) expressed *NANOG*, *OCT4*, *SOX2*, *REX1*, *DNMT3B*, *DPPA2*, and *KLF4* at 5–12 fold higher levels than in the double positive sub-population (S⁺ T⁺) (Fig. 5c). *LIN28*, which has been shown recently to promote transformation²⁷, was expressed at lower levels in the triple positive cell population (S⁺ T⁺ C⁺) compared to the double positive (S⁺ T⁺) and CD30-depleted (S⁺ T⁺ C⁻) cell pools (Fig. 5c). Notably, *MYC* and *KLF4*, known oncogenes, were expressed at comparable levels or slightly higher in the double negative sub-population relative to the triple positive population. Therefore isolating hiPSCs using a combination of SSEA4, TRA-1-81, and CD30 (on comparison with SSEA4 and Tra-1-81 alone) selectively and efficiently enriched for cells that express pluripotency markers, without necessarily increasing the pool of cells expressing oncogenes. In addition, examining the expression of *CD200* in the different cell pools, revealed a similar expression profile to that of *CD30* and *Podocalyxin* (Fig. 5c), suggesting that *CD200* might be another useful pluripotency marker. Conversely, *CD9*, was expressed at comparable levels in all the cell pools, confirming its limited specificity.

To verify that the hiPSCs isolated by FACS using the combination of SSEA4, TRA-1-81, and CD30 surface markers are truly pluripotent, we applied the same sorting strategy described above for the generation of 3 different hiPSC lines (FTi115, FTi116, and FTi117) from human neonatal fibroblasts (FTc91). The 3 hiPSC lines generated and maintained under feeder-free culture conditions (SMC4-supplement media) expressed high levels of SSEA4, TRA-1-81, and CD30 as assessed in post selection flow cytometry (Fig. 6a). They also homogeneously expressed all of the pluripotency markers tested and were seen to silence transgene expression (Figs. 6b–e). Karyotype analyses of the 3 hiPSC lines between passages 8 and 11, post selection, indicated that the cells maintained normal genomes under feeder-free and single cell culture (Fig. 6f). To test the tri-lineage differentiation potential, embryoid bodies were generated, seeded on Matrigel-coated plates and examined 2 weeks later. Ectoderm (cell positive for Nestin and Tuj1), mesoderm (cells positive for α SMA), and endoderm (cells positive for Fox2 and AFP) lineages were easily observed in the differentiated cultures by immunofluorescence (Fig. 6g). In addition, FTi117 iPSCs were injected subcutaneously in immunocompromised mice to test for teratoma

formation. Six weeks post injection, teratomas were dissected and found to contain tissues representative of the three germ layers (Fig. 6h). Thus hiPSCs isolated by FACS using the combined surface markers SSEA4, TRA-1-81, and CD30 are pluripotent and maintain stable genomes. These results confirm that CD30 is a highly specific marker of pluripotency with utility for the selection of bona fide hiPSCs when in combination with commonly used surface markers.

Discussion

The ability to generate hiPSCs from easily accessible somatic cell types and any genetic background opens the possibility for effective human disease modeling and eventual autologous or allogeneic cell therapy²⁸. Despite the relatively broad and quick adoption of iPSC technology, questions of efficiency (both time and cost), quality and reproducibility in iPSC derivation and selection remain. Further, the common methods of iPSC selection and characterization hinder the practical production of iPSCs from many somatic cell lines in parallel; limiting the ability to effectively model polygenic diseases or build large iPSC banks. We recently reported the use of a chemical cocktail for enhanced reprogramming efficiency, enhanced clonogenicity and survival following single cell passage of hiPSCs in feeder-free conditions⁸. These improvements facilitated the use of FACS as a tool for the isolation of rare hiPSCs from a heterogeneous reprogramming pool, enhancing reprogramming and hiPSC selection efficiency through an automated process.

In the present study our aim was to build on this FACS-based platform through the identification of additional surface markers for use in high-quality hiPSC selection, preferably cell surface antigens that could distinguish not only between reprogramming cells and differentiated cells but that could additionally select bona fide hiPSC clones from partially reprogrammed intermediates. To this end, we observed that while the commonly used pluripotency surface markers SSEA4 and TRA-1-81/TRA-1-60 can distinguish fibroblasts from hiPSCs, they were insufficient for the separation of the bona fide hiPSCs from some reprogramming intermediates. We screened a panel of CD molecules (CD9, CD30, CD50, CD90, and CD200) that have been implicated in pluripotency or are expressed in hPSCs. The expression profiles of these surface markers were explored by qRT-PCR in a wide array of cell types including hESCs and hiPSCs (cultured on feeder cells or on feeder-free conditions), partially reprogrammed cells, primary somatic cells, and hPSC-derived differentiated cells. The expression analyses revealed CD30 as a potential marker for the isolation of bona fide hiPSCs. Additionally, analysis of these surface markers across a panel of normal adult tissues revealed clear differences in expression patterns; CD9 displayed the broadest range whereas CD30 displayed the most restricted range of expression. CD30 was expressed in lymphocytes (peripheral blood), which confirms previous observations reporting its expression in activated lymphocytes²⁴. In contrast, CD200 and Podocalyxin were expressed in the majority of tissues examined. Interestingly, the CD30 and Podocalyxin expressions did not overlap in the same tissues, suggesting that their combined use for the selection of hiPSCs could provide added specificity. CD30 expression was also tested by qRT-PCR, flow cytometry and by fluorescence microscopy in reprogramming cell culture pools and in established, partially differentiated, hiPSC clones. In these cultures, CD30 positive staining displayed a strong correlation to the endogenous expression of *NANOG*, a definitive marker of pluripotency. Combining CD30 with traditional markers SSEA4 and TRA-1-81 significantly increased specificity and enhanced the efficiency of the FACS-based selection process. The hiPSC clones generated using this system maintained pluripotency and genomic stability over many single cell enzymatic passages in feeder-free culture systems. Our results validate the feasibility and efficiency of FACS as a tool for the automated isolation of pluripotent cells, and highlight the significance of the combinatorial use of multiple pluripotency surface markers for the

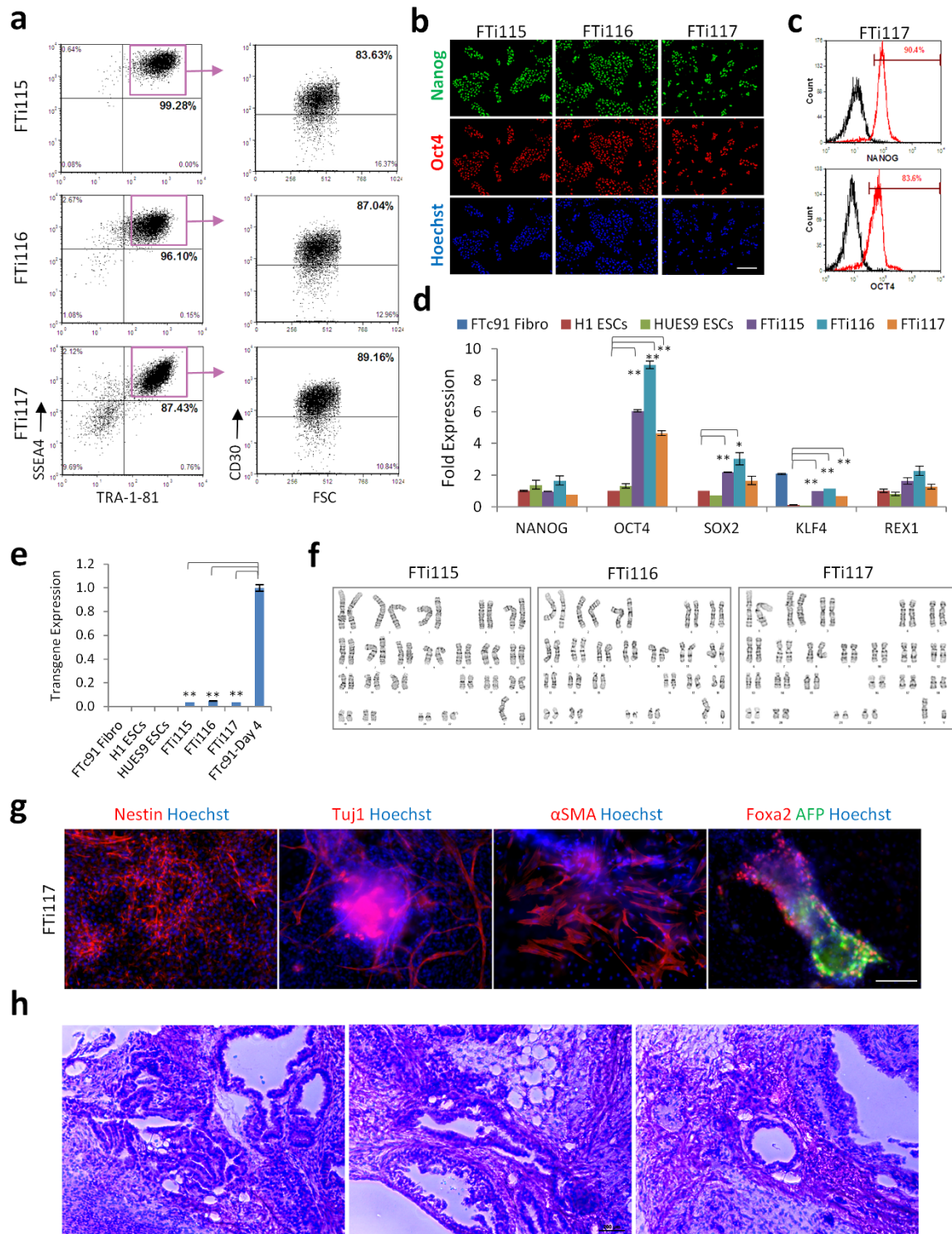


Figure 6 | hiPSC selection by sorting SSEA4⁺ TRA-1-81⁺ CD30⁺ cells yields pluripotent feeder-free hiPSC lines with stable genome and tri-lineage differentiation potential. Four weeks post initiation of reprogramming of FTc91 human fibroblasts, hiPSCs were separated by FACS and using the SSEA4, TRA-1-81, and CD30 markers as described above. Three iPSC lines FTi115, FTi116, and FTi117 were established and characterized. (a) Flow cytometry analyses of indicated hiPSCs. (b) Indicated iPSC lines were immunostained for expression of NANOG (green) and OCT4 (red). Nuclei were stained with Hoechst dye (blue). Scale bar is 200 μ m. (c) Nanog⁺ and Oct4⁺ FTi117 iPSCs were quantified by intracellular flow cytometry. (d), (e) qRT-PCR analysis of gene expression of pluripotency markers (d) and transgene (e) in indicated hiPSC lines. The lentivirus used for reprogramming expressed the transgenes as a single polycistronic cassette. Transgene levels were measured by a TaqMan primer-probe set within the viral WPRE element. Error bars represent standard deviation of duplicates. The asterisks denote corresponding p-values (* < 0.05, ** < 0.001). (f) FTi115 (p11), FTi116 (p10), and FTi117 (p8) maintained a normal karyotype (46, XY) over extended period of culture in SMC4 and on Matrigel coated plates. (g) Differentiation potential of FTi117 was tested by seeding 5-day old embryoid bodies on Matrigel-coated plate and staining cells 2 weeks later for markers of ectoderm (Nestin and TuJ1), mesoderm (α SMA), and endoderm (Foxa2 and AFP). Scale bar is 100 μ m. (h) Histological sections of teratoma derived from FTi117 iPSCs. Panels show neuroepithelia (left), adipocytes (middle), gut epithelia (right).



efficient and specific isolation of true hiPSCs. This process is easily adaptable to high throughput platforms, requiring little or no manual manipulation, facilitating its use in wider academic, industrial and eventually clinical applications.

Methods

Cell culture. hiPSCs were cultured in conventional media containing DMEM/F12 (Mediatech), 10 ng/mL bFGF (Invitrogen), 20% v/v knockout serum replacement (Invitrogen), 1% v/v non-essential amino acids (Mediatech), 2 mM L-glutamine (Mediatech) and 100 μ M β -mercaptoethanol. For most experiments described in this study, hiPSCs were cultured on Matrigel (BD Biosciences)-coated plates. Under feeder-free conditions, the media was supplemented with the SMC4 chemical cocktail to enhance survival and inhibit differentiation during and after FACS as described previously. The SMC4 chemical additives are PD0325901 (0.4 μ M), CHIR99021 (1 μ M), Thiazovivin (5 μ M) and SB431542 (2 μ M). Upon confluency, feeder-free hiPSCs were split using Accutase (3–5 min at 25°C; Millipore). Alternatively, and whenever indicated, hiPSCs were cultured (without SMC4) on a layer of mitomycin C treated mouse embryonic fibroblasts (Millipore). All cell cultures were maintained in a humidified incubator at 37°C and 5% CO₂.

Derivation, selection and analysis of hiPSCs by FACS. Human iPSCs were generated as described before⁸. Briefly, human fibroblasts were infected with a lentivirus containing a polycistronic cassette expressing human OCT4, SOX2 and KLF4. Virus-containing media was supplemented with 4 μ g/mL polybrene. Two to three days post-infection, fibroblast media was switched to SMC4-supplemented hiPSC media. For the initial bulk enrichment, reprogramming culture were split about 3 weeks post-infection and SSEA4/TRA-1-81 double positive were sorted by FACS. Immunostaining of cells was carried out using manufacturers' recommended antibody dilution in staining buffer containing Hanks' Balanced Salt Solution (Invitrogen), 4% fetal bovine serum (Invitrogen) and 10 mM Hepes (Invitrogen). Flow cytometry sorting was performed on FACS Aria II (BD, Biosciences), and all primary antibodies used were purchased from BD Biosciences. Sorting at clonal cell densities was performed by sorting cells directly into 96-well plates. Intracellular staining was performed using Cytofix/Cytoperm kit (BD Biosciences). Briefly, 5 \times 10⁵ cells were fixed and permeabilized in fix/perm buffer. Cells were subsequently stained with primary antibodies Oct3/4 (Santa Cruz) and Nanog (Santa Cruz) followed by secondary antibodies AlexaFluor 488 or 647 (Invitrogen). Cells were washed two times in perm/wash buffer after each staining step. Cells were then resuspended in staining buffer prior to analysis. Flow cytometry analysis was performed on Guava EasyCyte 8HT (Millipore).

RT-PCR analysis. Total RNA was isolated using the PicoPure RNA Isolation kit (Life Technologies), and used to generate first strand cDNA using the iScript cDNA Synthesis Kit (Bio-Rad). Relative gene expression levels were determined using the TaqMan Fast Universal PCR Master Mix (Applied Biosystems) and the FAM-labeled TaqMan probes listed in Supplementary Tables 1 and 2. For the analysis of gene expression in normal tissue samples (n = 48), TissueScan™ Normal Tissue cDNA Array (OriGene Technologies) was used in the qPCR analyses. The heat map was generated using the DataAssist software (Life Technologies). For each sample shown in the heat map, the middle expression level is set as the median of all of the Δ CT values from all assays for that sample. Data points for a given sample can only be compared relative to other data points for that sample.

Immunofluorescence staining. Cells were fixed using 4% v/v paraformaldehyde (Alfa Aesar), washed three times with PBS containing 0.2% v/v Tween (PBST) (Fisher Scientific) and permeabilized using 0.15% v/v TritonX-100 (Sigma-Aldrich) in PBS for 1 hr at 25°C. After permeabilization, cells were blocked with 1% v/v BSA (Invitrogen) in PBST (PBSTB) (Fisher Scientific) for 30 min at 25°C. After gentle removal of PBSTB, cells were incubated with primary antibody in PBSTB overnight at 4°C. The primary antibodies TRA-1-60, TRA-1-81, SSEA4 and CD30 were from BD Biosciences, β -III Tubulin and Foxa2 were from R&D Systems, AFP was from Dako, α -Smooth Muscle Actin was from Sigma, Oct4 was from Santa Cruz, and Nanog from Abcam. After the overnight incubation, cells were washed three times with PBST and stained with secondary antibody (Life Technologies) diluted 1:200 in PBSTB for 1 hr at 25°C. The cells were washed three times in PBST and stained with Hoechst dye or DAPI (Life Technologies).

Differentiation. To initiate differentiation, hiPSC were cultured in suspension in hiPSC medium without FGF. The formed EBs were transferred to Matrigel-coated plate and cultured in differentiation media for 2 weeks. Medium was changed every other day until cells were fixed and immunostained with indicated antibodies.

Teratoma formation. iPSCs were dissociated into single cells using Accutase, and 3 million cells were mixed in 100 μ L SMC4 supplemented medium and 100 μ L Matrigel and injected subcutaneously into NOD/SCID/ γ ^{null} mice. After six weeks, teratomas were harvested from mice and were immediately embedded in OCT and immersed in liquid nitrogen until complete OCT solidification (about 30–40 seconds), then promptly removed and stored in dry ice until further processing. Frozen samples were sectioned with a LEICA CM1900UV cryostat. At least 5 sections (5 μ m thickness) were obtained from different regions of each sample,

and every section was collected on superfrost/plus slides (Fisherbrand). Immediately after section collection, each slide was immersed in fixative (4% paraformaldehyde in 0.1 M phosphate buffer for 15 minutes). Staining was performed according to gold standard Harris' Hematoxylin & Eosin protocol (from Polysciences, Inc), and Cytoseal 60 was used as mounting medium. All sections were examined, interpreted and photographed using a Nikon Eclipse TS100 microscope equipped with a Nikon DS-Fi1 camera.

Karyotype analysis. Cytogenetic analysis was performed on twenty G-banded metaphase cells by WiCell Research Institute (Madison, WI).

Statistical analysis. Student's t test (two-tailed distribution) was used to calculate p-values. StepOne Software v2.2 (Life Technologies) was used to determine RQ minimum and maximum values (error bars).

1. Yu, J. *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science* **318**, 1917–1920 (2007).
2. Takahashi, K. *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861–872 (2007).
3. Grskovic, M., Javaherian, A., Strulovici, B. & Daley, G. Q. Induced pluripotent stem cells—opportunities for disease modelling and drug discovery. *Nat Rev Drug Discov* **10**, 915–929 (2011).
4. Rao, M. S. & Malik, N. Assessing iPSC reprogramming methods for their suitability in translational medicine. *J Cell Biochem* (2012).
5. Lowry, W. E. *et al.* Generation of human induced pluripotent stem cells from dermal fibroblasts. *Proc Natl Acad Sci U S A* **105**, 2883–2888 (2008).
6. Chan, E. M. *et al.* Live cell imaging distinguishes bona fide human iPSCs from partially reprogrammed cells. *Nat Biotechnol* **27**, 1033–1037 (2009).
7. Valamehr, B., Tsutsui, H., Ho, C. M. & Wu, H. Developing defined culture systems for human pluripotent stem cells. *Regenerative medicine* **6**, 623–634 (2011).
8. Valamehr, B. *et al.* A novel platform to enable the high-throughput derivation and characterization of feeder-free human iPSCs. *Sci Rep* **2**, 213 (2012).
9. Zhao, W., Ji, X., Zhang, F., Li, L. & Ma, L. Embryonic stem cell markers. *Molecules* **17**, 6196–6236.
10. Schopperle, W. M. & DeWolf, W. C. The TRA-1-60 and TRA-1-81 human pluripotent stem cell markers are expressed on podocalyxin in embryonal carcinoma. *Stem Cells* **25**, 723–730 (2007).
11. Wang, Y. C. *et al.* Specific lectin biomarkers for isolation of human pluripotent stem cells identified through array-based glycomics analysis. *Cell research* **21**, 1551–1563 (2011).
12. Kardel, M. D. & Eaves, C. J. Modeling human hematopoietic cell development from pluripotent stem cells. *Experimental hematology* **40**, 601–611 (2012).
13. Drukker, M. *et al.* Isolation of primitive endoderm, mesoderm, vascular endothelial and trophoblast progenitors from human pluripotent stem cells. *Nat Biotechnol* **30**, 531–542 (2012).
14. Tang, C. *et al.* An antibody against SSEA-5 glycan on human pluripotent stem cells enables removal of teratoma-forming cells. *Nat Biotechnol* **29**, 829–834 (2011).
15. Yuan, S. H. *et al.* Cell-surface marker signatures for the isolation of neural stem cells, glia and neurons derived from human pluripotent stem cells. *PLoS One* **6**, e17540 (2011).
16. Adewumi, O. *et al.* Characterization of human embryonic stem cell lines by the International Stem Cell Initiative. *Nat Biotechnol* **25**, 803–816 (2007).
17. Draper, J. S., Pigott, C., Thomson, J. A. & Andrews, P. W. Surface antigens of human embryonic stem cells: changes upon differentiation in culture. *J Anat* **200**, 249–258 (2002).
18. Mateizel, I. *et al.* Characterization of CD30 expression in human embryonic stem cell lines cultured in serum-free media and passaged mechanically. *Hum Reprod* **24**, 2477–2489 (2009).
19. Holford, L. C., Case, P. & Lawson, S. N. Substance P, neurofilament, peripherin and SSEA4 immunocytochemistry of human dorsal root ganglion neurons obtained from post-mortem tissue: a quantitative morphometric analysis. *Journal of neurocytology* **23**, 577–589 (1994).
20. Piao, J. H. *et al.* Cellular composition of long-term human spinal cord- and forebrain-derived neurosphere cultures. *Journal of neuroscience research* **84**, 471–482 (2006).
21. Barraud, P., Stott, S., Mollgard, K., Parmar, M. & Bjorklund, A. In vitro characterization of a human neural progenitor cell coexpressing SSEA4 and CD133. *Journal of neuroscience research* **85**, 250–259 (2007).
22. Kershaw, D. B. *et al.* Molecular cloning and characterization of human podocalyxin-like protein. Orthologous relationship to rabbit PCLP1 and rat podocalyxin. *The Journal of biological chemistry* **272**, 15708–15714 (1997).
23. Nielsen, J. S. & McNagny, K. M. The role of podocalyxin in health and disease. *Journal of the American Society of Nephrology : JASN* **20**, 1669–1676 (2009).
24. Kurts, C. *et al.* Signalling through CD30 protects against autoimmune diabetes mediated by CD8 T cells. *Nature* **398**, 341–344 (1999).
25. Deusch, Y. E., Tadmor, T., Podack, E. R. & Rosenblatt, J. D. CD30: an important new target in hematologic malignancies. *Leukemia & lymphoma* **52**, 1641–1654 (2011).
26. Rosler, E. S. *et al.* Long-term culture of human embryonic stem cells in feeder-free conditions. *Dev Dyn* **229**, 259–274 (2004).



27. Viswanathan, S. R. *et al.* Lin28 promotes transformation and is associated with advanced human malignancies. *Nat Genet* **41**, 843–848 (2009).
28. Saha, K. & Jaenisch, R. Technical challenges in using human induced pluripotent stem cells to model disease. *Cell Stem Cell* **5**, 584–595 (2009).

Author contributions

R.A., B.V. and M.R. conducted the experiments including surface marker expression analysis, hiPSC generation and characterization. B.R. conducted majority of the flow cytometry parameterization and sorting. F.V. performed teratoma sectioning and staining. B.V., R.A. and P.F. conceptualized the project. R.A., B.V. and P.F. wrote the paper with contribution from all authors.

Additional information

Supplementary information accompanies this paper at <http://www.nature.com/scientificreports>

Competing financial interests: All authors are current fulltime employees of Fate Therapeutics, San Diego, CA.

License: This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/3.0/>

How to cite this article: Abujarour, R. *et al.* Optimized Surface Markers for the Prospective Isolation of High-Quality hiPSCs using Flow Cytometry Selection. *Sci. Rep.* **3**, 1179; DOI:10.1038/srep01179 (2013).