npg

Open

ORIGINAL ARTICLE

Specific lectin biomarkers for isolation of human pluripotent stem cells identified through array-based glycomic analysis

Yu-Chieh Wang[1, 2], Masato Nakagawa[3], Ibon Garitaonandia[1, 2], Ileana Slavin[1, 2], Gulsah Altun[1, 2], Robert M Lacharite[1, 2], Kristopher L Nazor[1, 2], Ha T Tran[1, 2], Candace L Lynch[1, 2], Trevor R Leonardo[1, 2], Ying Liu[1, 2, 7], Suzanne E Peterson[1, 2], Louise C Laurent[1, 2, 7], Shinya Yamanaka[3, 4, 5, 6], Jeanne F Loring[1, 2]

[1]The Scripps Research Institute, Department of Chemical Physiology, La Jolla, CA 92037, USA; [2]The Scripps Research Institute, Center for Regenerative Medicine, 10550 N. Torrey Pines Rd., SP-3021, La Jolla, CA 92037, USA; [3]Center for iPS Cell Research and Application (CiRA), Kyoto University, Kyoto, Japan; [4]Institute for Integrated Cell-Material Sciences, Kyoto University, Kyoto, Japan; [5]Japan Science and Technology Agency, Yamanaka iPS Cell Special Project, Kawaguchi, Japan; [6]Gladstone Institute of Cardiovascular Disease, San Francisco, CA 94158-2261, USA; [7]Department of Reproductive Medicine, University of California, San Diego, La Jolla, CA, USA

Rapid and dependable methods for isolating human pluripotent stem cell (hPSC) populations are urgently needed for quality control in basic research and in cell-based therapy applications. Using lectin arrays, we analyzed glycoproteins extracted from 26 hPSC samples and 22 differentiated cell samples, and identified a small group of lectins with distinctive binding signatures that were sufficient to distinguish hPSCs from a variety of non-pluripotent cell types. These specific biomarkers were shared by all the 12 human embryonic stem cell and the 14 human induced pluripotent stem cell samples examined, regardless of the laboratory of origin, the culture conditions, the somatic cell type reprogrammed, or the reprogramming method used. We demonstrated a practical application of specific lectin binding by detecting hPSCs within a differentiated cell population with lectin-mediated staining followed by fluorescence microscopy and flow cytometry, and by enriching and purging viable hPSCs from mixed cell populations using lectin-mediated cell separation. Global gene expression analysis showed pluripotency-associated differential expression of specific fucosyltransferases and sialyltransferases, which may underlie these differences in protein glycosylation and lectin binding. Taken together, our results show that protein glycosylation differs considerably between pluripotent and non-pluripotent cells, and demonstrate that lectins may be used as biomarkers to monitor pluripotency in stem cell populations and for removal of viable hPSCs from mixed cell populations.

Keywords: lectins; glycosylation; pluripotency biomarkers; human pluripotent stem cell; induced pluripotent stem cells; embryonic stem cells

Introduction

Human pluripotent stem cells (hPSCs) hold great promise for the treatment of human disease. These cells,

Correspondence: Jeanne F Loring
Tel: +1-858-784-7362; Fax: +1-858-784-7211
E-mail: jloring@scripps.edu
Website: http://www.scripps.edu/loring/

which include human embryonic stem cells (hESCs) and induced pluripotent stem cells (hiPSCs), have two defining characteristics: the ability to self-renew indefinitely and the ability to differentiate into all cell types in the body. These qualities make hPSCs potential sources of large quantities of any cell type that can then be used for research and clinical applications, including disease modeling, drug screening, and cell replacement therapies for degenerative disease. However, a major concern for the use of hPSCs for all of these applications, especially for transplantation therapy, is the need for a pure popula-

tion of the cell type of interest. For cell therapy, residual undifferentiated hPSCs are potentially tumorigenic and thus must be removed [1]. In addition, use of impure cell populations can lead to inaccurate evaluations of drug efficacy and toxicity, differentiation potential, or other cellular responses. Thus, reliable techniques for identifying pluripotent cells in populations are critically needed, not only for purification of differentiated hPSCs intended for clinical use, but also for validation of cell identity in basic research. There are a few antibodies that target pluripotency-associated cell surface antigens [2-7] and can be used to identify pluripotent cells in preparations of viable cells. But the generation of cell type-specific surface antibodies is challenging, and optimization of cell purification strategies would be enhanced by additional methods to identify, isolate, and remove pluripotent cells.

Pluripotency-associated antigens on the cell surface are often glycoproteins or glycolipids [2-4, 7], suggesting that specific glycosylation patterns of proteins may be a hallmark of pluripotency. Glycosylation is the most widespread post-translational modification, occurring in over 50% of all cellular proteins [8]. Dynamic changes in protein glycosylation have been observed in tumorigenesis [9, 10], cell differentiation [11], cell adhesion [12], and host-pathogen interactions [13, 14], and communications between the intracellular signaling pathways and the extracellular environment are frequently mediated through oligosaccharide modifications of membrane-bound proteins and their ligands [15, 16]. Though the importance of protein glycosylation in hPSCs has been suggested by several previous studies using cell biology and biochemical methods [16-18], information regarding the global protein glycosylation status in a broad spectrum of hPSCs and their differentiated derivatives is lacking. Advances in microarray technologies have led to the development of lectin microarrays, which use lectins as probes to determine the relative amounts of specific glycan marks among different cell populations. These arrays make sensitive, high-throughput characterization of protein glycosylation feasible [19, 20]. The utility of lectin arrays for characterizing cells has been illustrated by two very recent studies [21, 22].

To identify lectins that are useful for identifying and separating hPSCs in mixed cell populations, we systematically profiled the glycoproteins in a broad spectrum of cells (Table 1) using lectin microarrays. These cells included hESCs, human induced pluripotent stem cells (hiPSCs), human dermal fibroblast (HDF) cells, human epidermal melanocytes (HEM), malignant melanoma cell lines, and other differentiated cell types. We identified a subset of lectins that distinguish pluripotent cells from non-pluripotent cells, regardless of cell origins and culture methods. In addition, we demonstrated the feasibility of using a pluripotency-associated lectin to identify and remove viable pluripotent cells from heterogeneous populations containing differentiated cells. We also show correlations between expression of specific glycosyltransferases and pluripotency-associated protein glycosylation.

Results

Identification of pluripotency-associated lectins

To identify lectins recognizing glycoproteins on the cell surface, we first used the lectin arrays to determine the lectin-binding patterns for the hydrophobic protein fraction. This fraction should be highly enriched for membrane-bound proteins because of the abundance of hydrophobic amino-acid residues in transmembrane domains, while hydrophilic proteins are more likely to be intracellular proteins. The profiling results were subjected to a supervised clustering analysis, and the 13 lectins that showed highly significant differential binding between pluripotent and non-pluripotent cell types ($P < 0.05$, t-test) were identified for a "priority lectin" list (Supplementary information, Table S1). As shown in Supplementary information, Figure S1, hierarchical clustering based on the glycan signatures of hydrophobic proteins determined by these 13 priority lectins shows that the pluripotent cells clustered together, regardless of their type (hESC or hiPSC), source (somatic cell type used for reprogramming), or location of the lab of origin (La Jolla or Kyoto). With the exception of human leukocytes, whose profile overlapped with the pluripotent cells, the tested non-pluripotent cells all clustered separately. The shared profiles between hPSCs and human leukocytes may be due to the frequent presence of sialyl-Lewis X glycan on the leukocyte surface; reactivity of fucose-specific and sialic acid-specific lectins has been reported in studies of human leukocytes [23, 24].

To validate the reliability of glycomic profiling as a means of identifying pluripotent cells, we challenged the system by comparing nearly genetically identical pluripotent and non-pluripotent cells. We compared two newly reprogrammed hiPSC clones (iPS1.HEMd and iPS2.HEMd) with their somatic cell line of origin, a primary human melanocyte line (HEMd). These cell lines have identical genomes except for the presence of integrated sequences from the reprogramming factors used to generate the hiPSCs. The hiPSC lines were identified as pluripotent by immunofluorescence staining, differentiation *in vitro*, and teratoma generation (Supplementary information, Figure S2a and S2b). When analyzed and clustered with the other samples, the hiPSC lines

Table 1 List of cultured cells used in the study

Sample name	Registry name[1]	Note[2]
Human embryonic stem cells		
BG01	BG01	Passage 67[3]
BG02	BG02	Passage 46[3]
BG03	BG03	Passage 55[3]
WA01	WA01	Passage >50[3]
WA01	WA01	Passage 35[4]
WA07	WA07	Passage 35[3]
WA09	WA09	Passage 31[4]
WA09P90_MEF/MAN	WA09	Passage 90; MEF, manual passage[3, 5]
WA09P90_MG/MAN	WA09	Passage 90; Matrigel, manual passage[3, 5]
WA09P90_MEF/ACC	WA09	Passage 90; MEF, enzymatic passage[3, 5]
WA09P42_MG/ACC	WA09	Passage 42; Matrigel, enzymatic passage[3, 5]
WA09P90_MG/ACC	WA09	Passage 90; Matrigel, enzymatic passage[3, 5]
Induced pluripotent stem cells from human dermal fibroblasts (HDF)		
201B2 P24	CiRA201i-B2HDF	Passage 24[4]
201B6 P24	CiRA201i-B6HDF	Passage 24[4]
201B7 P24	CiRA201i-B7HDF	Passage 24[4]
201B7 P78	CiRA201i-B7HDF	Passage 78[4]
253G1 P24	CiRA253i-G1HDF	Passage 24[4]
253G4 P24	CiRA253i-G4HDF	Passage 24[4]
iPS2.HDF-ALS275	TSRI02i-ALS275	Passage 60; ALS patient 275[3, 5]
iPS5.HDF-ALS285	TSRI05i-ALS285	Passage 24; ALS patient 285[3, 5]
iPS8.HDF-ALS284	TSRI08i-ALS284	Passage 22; ALS patient 284[3, 5]
iPS1.HDF	TSRI001i-HDF	Passage 37; Matrigel, enzymatic passage[3, 6]
iPS7.HDF	TSRI007i-HDF	Passage 62; Matrigel, enzymatic passage[3, 6]
iPS11.HDF	TSRI011i-HDF	Passage 58; Matrigel, enzymatic passage[3,6]
Induced pluripotent stem cells from human epidermal melanocytes (HEM)		
iPS1.HEMd	TSRI001i-HEMD	Passage 3[3]
iPS2.HEMd	TSRI002i-HEMD	Passage 3[3]
Cells used for reprogramming		
aHDF	NA	HDF passage 6[4]
HDF51	NA	HDF passage 9; Sciencell[3]
HDF-ALS275	NA	HDF passage 7; ALS patient 275[3, 8]
HDF-ALS284	NA	HDF passage 6; ALS patient 284[3, 8]
HEMd	NA	HEM passage 5; Sciencell[3]
Other cell types		
SKMEL2	NA	Human melanoma cells, passage 45; ATCC[3]
MALME3M	NA	Human melanoma cells, passage 36; ATCC[3]
SKMEL5	NA	Human melanoma cells, passage 41; ATCC[3]
UACC257	NA	Human melanoma cells, passage 10; Charles River[3]
WA07P34MNPD29	WA07	Motor neuron progenitor cells differentiated from WA07 cells[7]
HREpic	NA	Human renal epithelial cells, passage 12; Sciencell[3]
HSkMC	NA	Human skeletal muscle cells; passage 15; Sciencell[3]
MEF	NA	Mouse embryonic fibroblast[3]

[1]Name of cell line submitted to International Stem Cell Registry (U. Mass.).
[2]Passage number, culture methods, source of cells.
[3]Samples cultured in Loring laboratory.
[4]Samples cultured in Yamanaka laboratory.
[5]Provided by I Garitaonandia, Loring laboratory.
[6]Provided by I Slavin, Loring laboratory.
[7]Provided by H Keirstead, UC Irvine.
[8]Cells provided by Solutions for ALS (SALSa).

were grouped with the other pluripotent cells, while the primary melanocytes were classified as non-pluripotent (Figure 1A). To explore this further, hESCs (WA07) were differentiated into motor neuron progenitor cells using a well-established protocol [25]. The glycomic profile of the neuronal cells was then compared to that of the undifferentiated hESCs. Similar to the somatic cell-hiPSC comparison, the differentiated motor neuron progenitors clustered with the non-pluripotent cells and the undifferentiated hESCs cells with the same genome were grouped with the pluripotent cells (Figure 1A).

Detection of hPSCs and separation of hPSCs from differentiated cells using pluripotency-associated lectins

From the 13 priority lectins, we selected 3 lectins, AOL, TJA-II, and UEA-1, that showed consistently strong signals in pluripotent cells (Figure 1B). To determine whether these lectins are specifically bound to the surfaces of hPSCs, we performed immunofluorescence staining on a mixed population of HDFs and hESCs using biotinylated lectins. The UEA-I lectin bound selectively to pluripotent hESCs that are also positive for POU5F1 (also known as OCT4) and SSEA-4, two well-established markers of pluripotency (Figure 1C). Similar staining patterns were also observed when the cells were incubated with biotinylated AOL and biotinylated TJA-II (Supplementary information, Figure S2c and S2d).

Since UEA-I showed the lowest reactivity with non-pluripotent cells among the three selected lectins in the array analysis (Figure 1B) and fluorescence staining (Figure 1C and Supplementary information, Figure S2c and S2d), we tested the utility of lectins for separation of cells from mixed hPSC/differentiated cell populations by using a magnetic submicron bead separation procedure

mediated by the UEA-I lectin. The general procedure to test hPSC-purging efficiency is illustrated in Figure 2A. As shown in Figure 2B and Supplementary information, Figure S3, depletion of hPSCs from mixed populations was very efficient; non-pluripotent (calcein-positive) cells comprised up to 99.5% of the uncaptured cell population recovered from mixtures of WA09 cells and HDF or HEMd. For pluripotent R-Olig2 cells mixed with their non-pluripotent, differentiated derivatives, the purity of calcein-positive cells in the uncaptured cell population after the cell separation was about 95% (Figure 2B and Supplementary information, Figure S3f). These data strongly suggested that the lectins identified in this study could be a useful tool to purge hPSCs from a variety of differentiated cell types.

To determine whether viable hPSCs could be isolated from mixed cell populations using lectins, a mixture of WA09 hESCs and fibroblasts was sorted using UEA-I-linked magnetic beads. Results are shown in Supplementary information, Figure 3a and 3b. Both bound and unbound populations were replated in culture dishes and stained with SSEA-4 to identify pluripotent cells in each fraction. In the viable cell population captured from the mixed population by UEA-I beads, calcein fluorescence (fibroblasts) was mostly absent, and most cells were positive for the pluripotency-associated markers SSEA4 (Supplementary information, Figure S3b), POU5F1, NANOG, SOX2, Tra-1-60, and Tra-1-81 (Supplementary information, Figure S3c). The captured cells retained the ability to differentiate *in vitro* into all three germ layers (Supplementary information, Figure S3c). In contrast, most of the cells in the unbound fraction were fibroblasts (calcein-positive) and negative for SSEA-4 (Supplementary information, Figure S3a), indicating that the lectin-

Figure 1 Glycomic profiling of hPSCs and non-pluripotent cells using lectin microarrays with validation of lectin binding to hPSCs. **(A)** Hydrophobic proteins isolated from the cells were labeled and analyzed by lectin microarrays. Supervised hierarchical clustering based on the array data of 13 priority lectins for the hydrophobic proteins of 26 samples of hPSCs (12 samples of hESC and 14 samples of hiPSCs, including 2 clones of hiPSCs newly reprogrammed from human primary melanocytes) and 22 samples of mammalian non-pluripotent cells (4 samples of HDFs, 4 samples of human melanoma cells, 1 sample of primary human melanocytes, 1 sample of mouse embryonic fibroblasts, 2 samples of human leukocytes, 1 sample of human adult brain tissue, 1 sample of human fetal liver tissue, 1 sample of motor neuron precursor cells differentiated from WA07 cells, 1 sample of cardiac muscle, 1 sample of adipose tissue, 2 samples of skeletal muscle, 1 sample of adrenal gland, 1 sample of lung tissue, and 1 sample of renal epithelium) was performed using the NIA Array Analysis Tool. In the dendrogram, samples acquired at the Scripps Research Institute and Kyoto University are indicated by yellow and blue solid squares, respectively. Pluripotent and non-pluripotent cells are indicated by red dots and green dots, respectively. Some of the samples are highlighted by colored lines to emphasize their relationships. Purple line: WA07 cells. Pink line: motor neuron precursor cells differentiated from WA07 cells. Black line: human primary melanocytes. Orange line: hiPSCs derived from primary melanocytes. **(B)** Fluorescent intensity of UEA-1, AOL, and TJA-II lectin binding with hydrophobic cell extracts based on microarray data. RFU: relative fluorescence units. **(C)** WA09 cells were spiked into a cell suspension of HDFs to create a mixed cell population. The cells were fixed on adhesion slides and then subjected to staining with anti-POU5F1 and SSEA-4 antibodies as well as biotinylated UEA-I lectin. WA09 cells with expression of pluripotent biomarkers POU5F1 or SSEA-4 are indicated by arrowheads and positive UEA-I staining. Insets, magnified images of the double-positive cells.
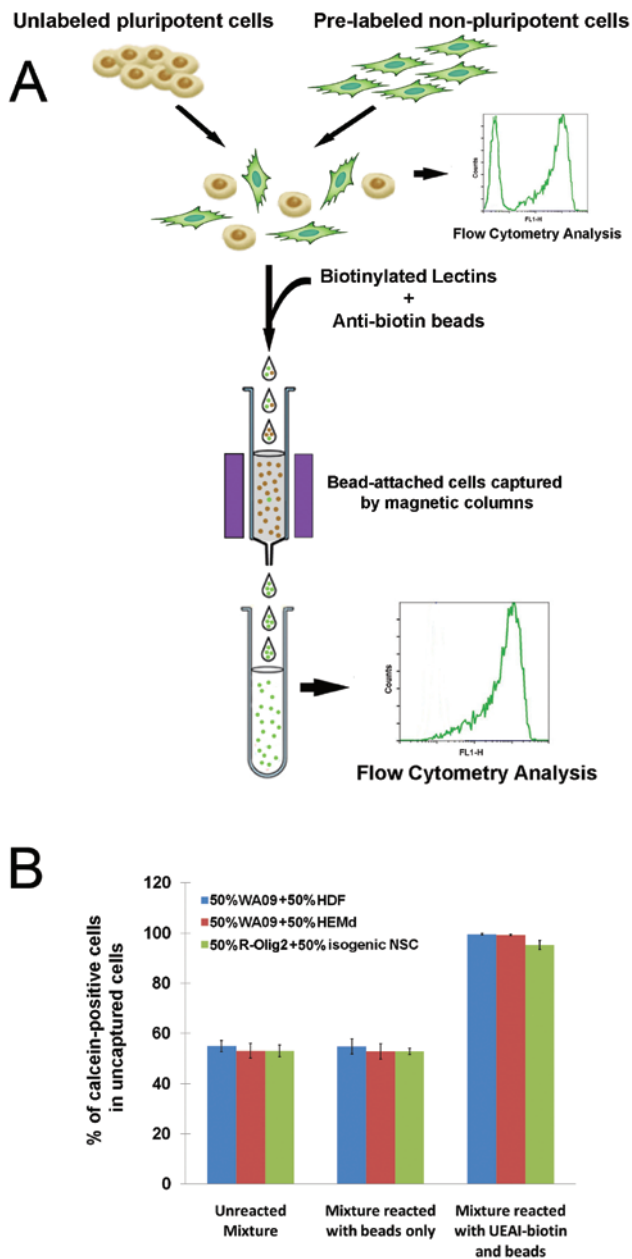
Figure 2 Lectins can be used to purge pluripotent cells from mixed populations. **(A)** Diagram of depletion experiments. hESCs (unlabeled) were mixed with different Calcein AM-labeled differentiated cells (green) at a ratio of 1:1. Cells were incubated with UEA-I magnetic beads. The cells captured by the magnetic beads were isolated and the unbound, eluted cells were analyzed for Calcein AM fluorescence by flow cytometry. **(B)** Results of depletion experiments using mixtures of WA09 and HDFs, WA09 and HEMd, and R-Olig2 hESCs and their NSC derivatives. For each sample, the unreacted mixture and the mixture reacted with beads showed about 50% Calcein AM-positive cells because half of the cells in the 1:1 mixture were labeled. After magnetic separation of the beads, nearly all of the remaining cells were Calcein AM-positive, indicating that the pluripotent cells had been removed from the mixture.

bound beads effectively separated viable pluripotent and non-pluripotent cells.

To quantify the sensitivity and specificity of the binding of UEA-I lectin in hPSCs, we used UEA-1-mediated fluorescence staining in conjunction with flow cytometry analysis. Approximately 95% of WA09 cells were strongly positive for UEA-I binding, while less than 5% of HDF cells were dimly positive (Figure 3A). We found that UEA-I was rendered easily removable from the cell surface by washing in a fucose-containing buffer (data not shown). Flow cytometric analysis of multiple hPSCs lines co-stained with SSEA-4 antibody and UEA-I lectin (Figure 3B) indicated that UEA-I is a comparable biomarker to SSEA-4 for detecting cellular pluripotency with high sensitivity and specificity.

Comparison of lectin-binding patterns in hydrophobic and hydrophilic proteins extracted from hPSCs and differentiated cells

The results shown so far describe the glycocomponents of hydrophobic proteins expressed in hPSCs. To determine whether the glycomic profiles of the hydrophilic protein fraction also reflected the pluripotent state of cells, we selected hydrophilic protein samples extracted from 11 samples of hPSCs and 10 samples of non-pluripotent cells for profiling on the lectin arrays. As was observed for the hydrophobic proteins, hierarchical clustering of the glycomic profiles of hydrophilic proteins separated the pluripotent cells from the non-pluripotent cells (Supplementary information, Figure S4a).

Comparison of the profiles of hydrophilic and hydrophobic proteins (Supplementary information, Figure S4) revealed that the glycans that differed between hPSCs and non-pluripotent cells were different in the two protein fractions. In the hydrophobic fraction, fucose-specific (UEA-I, AOL, AAL and TJA-II) and sialic acid-specific (SNA, SSA and TJA-I) lectins bound preferentially to the proteins isolated from the hPSCs. In the hydrophilic fraction, lectins that have high affinity for glycans containing *N*-acetylgalactosamine (GalNAc), such as BPL, Jacalin, WFA, MAP, and HPA, preferentially interacted with the hPSC proteins (Supplementary information, Figure S4c). Our results suggest that glycoproteins from both the hydrophilic and the hydrophobic fractions could be used to discriminate between pluripotent and non-pluripotent cells. We observed an interesting trend in the non-pluripotent cells: cells of similar lineages tended to be clustered more closely together in their hydrophobic protein glycan profiles than in their hydrophilic protein glycan profiles (e.g., melanocytes with melanoma, fibroblasts from different sources). We detected no such lineage-associated patterns among the pluripotent cells.
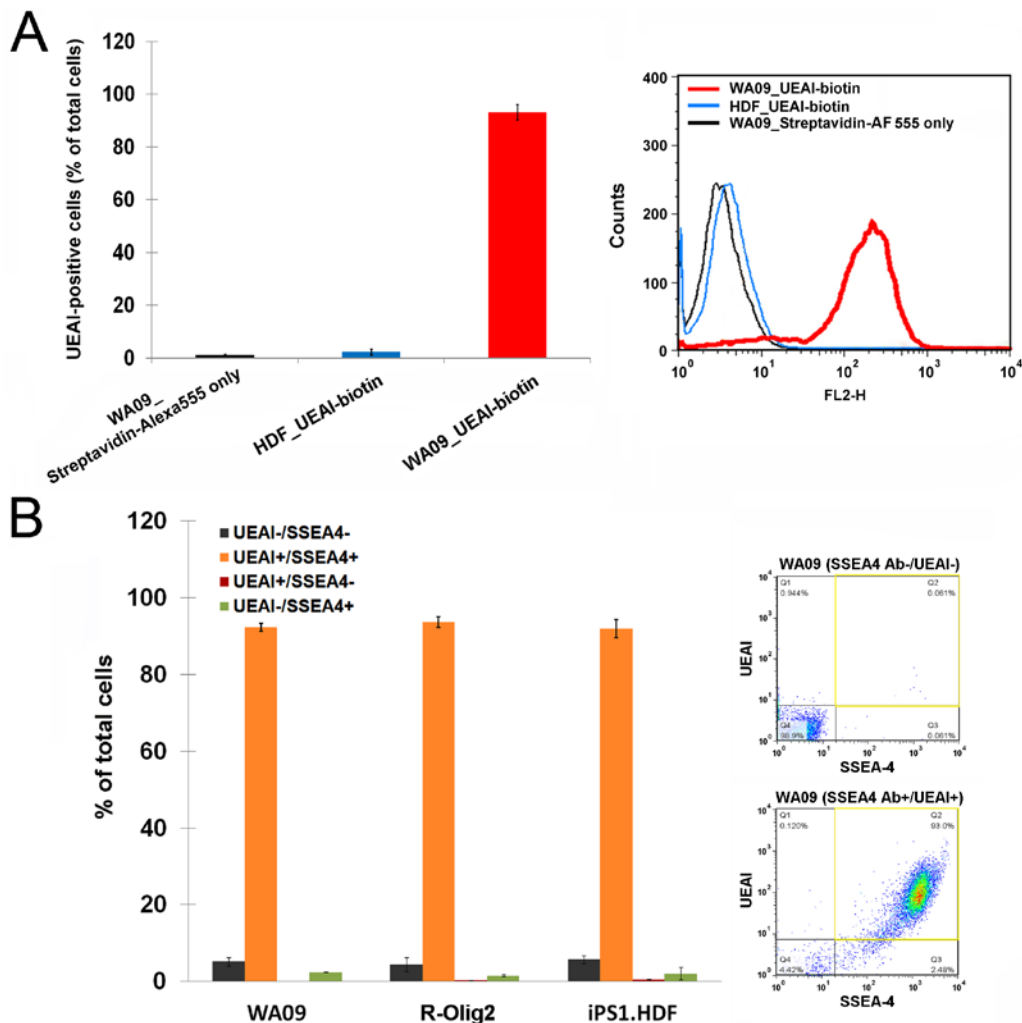
Figure 3 Lectin binding to pluripotent cells. **(A)** WA09 hES cells were incubated with streptavidin-AF 555 only or with strepta-vidin-AF 555 and biotinylated UEA-I. Human dermal fibroblasts (HDFs) were incubated with streptavidin-AF 555 and biotiny-lated UEA-I. Fluorescence intensity was analyzed by flow cytometry. As expected, WA09 cells incubated with streptavidin-AF 555 alone (negative controls) as well as HDFs incubated with streptavidin-AF 555 and biotinylated UEA-I both showed mini-mal levels of fluorescence, while WA09 cells incubated with streptavidin-AF 555 and biotinylated UEA-I showed high levels of fluorescence. **(B)** WA09, R-Olig2, and iPS1.HDF pluripotent cells were incubated with secondary antibody and streptavidin-AF 555 only (negative control; upper right) or SSEA-4 antibody, secondary antibody, UEA-I biotinylated lectin and streptavid-in-AF 555 (treated cells, lower right), and subjected to flow cytometry. The negative control cells show minimal fluorescence, but more than 95% of the treated cells in all three tested hPSC lines show either double-positive or double-negative staining. This indicates that biotinylated UEA-I lectin can be used in flow cytometry and that it labels a similar percentage of pluripotent cells as SSEA-4, a well-recognized biomarker of human cell pluripotency.

Expression analysis of fucosyltransferase and sialyl-transferase genes in hPSCs and differentiated cells

The differences in glycoprotein signatures between pluripotent and non-pluripotent cells suggested that dif-ferential expression of glycotransferase genes may be in-volved in the establishment and maintenance of distinct glycoprotein profiles. To address this idea, we examined relative mRNA expression levels of 13 fucosyltrans-

ferase genes and 19 sialyltransferase genes using array-based gene expression analysis of 22 samples of hPSCs and 27 samples of non-pluripotent cells. The hierarchi-cal clustering of the tested samples revealed that hPSCs share similar expression patterns of fucosyltransferase and sialyltransferase genes, which differ considerably from those expressed by non-pluripotent cells (Figure 4A). Differential gene expression analysis indicated that

FUT1, *FUT2*, *FUT10*, *ST3GAL2*, and *ST6GAL1* genes had higher expression levels in the hPSCs, whereas *POFUT1*, *POFUT2*, *ST3GAL1*, *ST3GAL3*, *ST3GAL4*, *ST3GAL5*, *ST3GAL6*, *ST6GALNAC6*, *ST8SIA1*, and *ST8SIA4* genes were preferentially expressed in the non-pluripotent cells (Figure 4B). Consistent with our data, the major products generated by the fucosyltransferases FUT1 and FUT2, which were highly expressed in pluripotent cells, are known to bind specifically to the lectin UEA-I that we used to discriminate pluripotent from non-pluripotent cells ([26], Lectin Frontier Database at http://riodb.ibase.aist.go.jp/rcmg/glycodb/LectinSearch). Also, the major product generated by the sialyltransferase ST6GAL1 has a high affinity for other lectins, SNA, SSA, and TJAI, that showed preferential binding to pluripotent cells ([27], Lectin Frontier Database). While the expression levels of mRNAs may not directly reflect the abundance or activity of these glycosyltransferases, they suggest that changes in glycosyltransferase gene expression may be involved in the regulation of pluripotency and differentiation of human cells.

Discussion

In this study, we identified lectin biomarkers for pluripotent cells, and show that specific lectins can be used to identify, sort, and separate human pluripotent cells (hPSCs) in mixed populations with differentiated cells. Our results suggest that lectins can be integrated into strategies for purifying populations of hPSC-derived differentiated cells destined for drug screening, studies of human disease, and cell replacement therapy. It is especially critical to purge undifferentiated hPSCs from cell preparations used for clinical transplantation, because the greatest concern for using hPSC-derived cells is that residual undifferentiated cells may be tumorigenic. To our knowledge, this is the first report showing that lectins can be used to isolate and remove hPSCs from mixed cell populations. Historically, lectins have been known for their affinity for glycans on many different cell types. The idea that lectins may react with pluripotent cells was initiated in early 1980s by studies showing the ability of lectins to bind to or kill embryonal carcinoma and germ cell tumor cells [28-32]. However, a comprehensive, large-scale analysis of lectin binding in hPSCs was not tangible until recently [21, 22]. Consistent with our findings, hPSC-preferential binding of lectins that have high affinity to glycans containing sialic acid α2-6 and fucose α1-2 structures was observed in a study using a large quantity but relatively limited types of cell samples [21]. The presence of unique fucosylation signatures on proteins in hPSCs has also been detected and associated with

the pluripotent status of the cells by proteomic analysis [17]. Although differential lectin binding to proteins extracted from pluripotent stem cells using array-based approaches has been shown by two recent reports, neither study attempted to apply this information for developing methods for isolation of pluripotent cells from mixed populations or for purging differentiated populations of residual pluripotent cells [21, 22].

The data presented here show that the fucose-specific lectin UEA-I is a potential tool for isolating viable hPSCs with sustained pluripotency from heterogeneous populations (Supplementary information, Figure S3c). Currently, separation of cell subpopulations relies in large part on antibodies to cell surface epitopes, and our research suggests that cell type-specific lectins can complement existing methods. In some cases, lectins may offer certain advantages over antibodies. They can be easily removed from cells without damaging them, by adding an oligosaccharide or monosaccharide that competes with the glycan to which the lectin binds. Also, since lectins are significantly less expensive than antibodies, they may offer advantages for procedures in which cost is an important factor, such as large-scale hPSC purifications for cell therapy and drug screening.

Cell surface antibodies are often species-specific; for example, SSEA-4 antibody, a commonly used antibody for human PSCs, does not recognize mouse PSCs, and instead binds to differentiated mouse cells. In contrast, our data suggest that the same lectins that can be used to identify human PSCs can also be used for pluripotent cells from other species, including mouse (Supplementary information, Figure S5) and a diversity of other animals (not shown). Thus, for labs working with pluripotent cells from more than one species, the same lectin can be used for all pluripotency monitoring and isolation applications. Also, the cross-species lectin recognition of PSCs implies that certain types of glycosylation are hallmarks of the pluripotent state; this suggests that studies of specific glycosylation may reveal novel information about transitions that occur during cellular reprogramming and differentiation of PSCs.

Many of the lectins that showed preferential association with the hydrophobic proteins of hPSCs bind to fucosylated and sialylated glycans, suggesting that fucosylation and sialylation are common characteristics of PSC glycoproteins. Fucosylation and sialylation are two common types of glycosylation on the cell surface [26, 33], and these modifications have been shown to be critical in normal embryogenesis and somatic stem cell differentiation [16, 34-36]. Our gene expression analysis suggests that the protein fucosylation and sialylation signatures in hPSCs may be due to differential expression
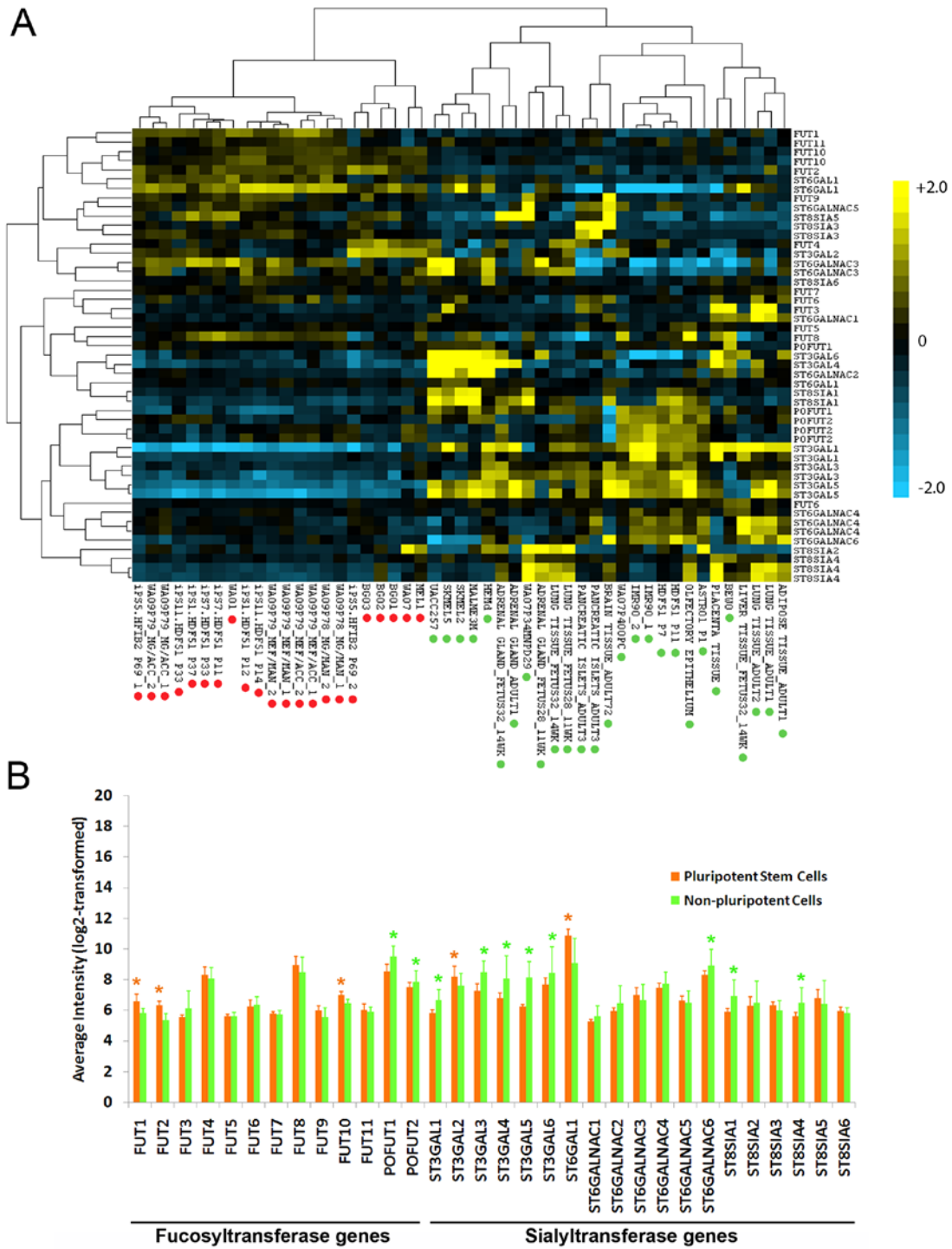
Figure 4 Gene expression profiles of human fucosyltransferases and sialyltransferases in hPSCs and non-pluripotent cells. **(A)** RNA was hybridized to Illumina arrays as described in Materials and Methods. The heat-map representation of the array data indicates the relative expression levels of specific glycosyltransferases in the cells. The results of the hierarchical clustering of expression patterns in different cell types are illustrated as a dendrogram. Pluripotent and non-pluripotent cells are indicated by red and green dots, respectively. **(B)** The expression levels of 13 fucosyltransferases and 19 sialyltransferases in 22 samples of hPSCs (8 samples of hiPSCs and 14 samples of hESCs) and 27 samples of human non-pluripotent cells (14 samples of organ tissues and 13 samples of cultured cells). The fluorescence intensities of each gene probe were averaged, log2-transformed and plotted. Columns, mean of log2-transformed fluorescence intensities; bars, standard deviation. *$P < 0.01$, fold change of expression > 1.4 (differential expression analysis performed by GenePattern using non-transformed values of intensity).

of specific fucosyltransferases and sialyltransferases in pluripotent cells. This supports the idea that post-translational glycosylation of cell surface proteins plays a role in the regulation of transitions between pluripotent and differentiated states that occur during reprogramming of somatic cells and differentiation of pluripotent cells. It is also interesting to note that hydrophilic glycoproteins, which are primarily intracellular proteins, also differed between pluripotent and differentiated cells. In particular, we observed differences in GalNAc-containing glycoconjugates; GalNAc usually acts as a core component of *O*-linked glycans, which are frequently found on serine and threonine residues in mucins and other secreted proteins [37]. This suggests that *O*-linked glycosylation of serine and threonine in secreted proteins may be involved in the regulation of pluripotency, possibly by controlling autocrine or paracrine feedback loops.

Our work shows that lectin-based approaches are sensitive and reliable tools to detect pluripotency-associated protein glycosylation. The most important application of lectins to the basic and clinical applications of hPSCs is likely to be removal of residual pluripotent cells from differentiated populations, and to our knowledge, our report is the first to show the feasibility of using lectins to remove undifferentiated pluripotent cells from mixed populations. Additionally, our work suggests that glycosylation changes may also provide clues to the mechanisms underlying reprogramming, pluripotency, and differentiation. We anticipate that the use of lectins as biomarkers will become a valuable tool for stem cell research, especially for translational applications such as purifying precursor populations or purging pluripotent cells from differentiated populations.

Materials and Methods

Cell culture

Cells obtained from the Loring lab, HDF51 and MEF cells, were cultured in DMEM containing 10% fetal bovine serum (FBS) at 37 °C. HEMd cells (ScienCell Research Laboratories, Carlsbad, CA, USA) were cultured in melanocyte medium (MelM; ScienCell Research Laboratories). The methods for culturing all the hPSCs provided by the Loring lab are listed in Supplementary information, Table S2.

Cells obtained from the Yamanaka lab, adult HDF (aHDF) cells harvested from a human skin biopsy specimen, were cultured in DMEM containing 10% FBS, 50 units/ml penicillin and 50 µg/ml streptomycin. The methods for culturing all the hPSCs provided by the Yamanaka lab are listed in Supplementary information, Table S2.

Dr Hans Keirstead generously provided WA07 hESCs and WA07P34MNPD29 motor neuron progenitors differentiated from WA07 hESCs. The culture methods and the differentiation protocols were previously reported [25].

Generation of hiPSCs

In the Loring lab, HEMd and HDF cells were seeded onto a six-well plate at a density of 1.5×10^6 cells per well one day before viral transduction. Cells were given equal volumes of viral supernatants containing retroviruses carrying each of four pMX expression vectors for *POU5F1*, *Sox2*, *Klf4*, *MYC* genes [38] on day 0 and day 1. On day 2, the transduced cells were transferred onto radiation-inactivated MEF feeder cells at a density of 1×10^4 cells per well of a six-well plate and cultured in DMEM/F12 media with L-glutamine containing 20% KnockOut™ Serum Replacement, 100 µM non-essential amino acids, 100 µM β-mercaptoethanol (all from Life Technologies, Carlsbad, CA, USA), 12 ng/ml FGF2, and 500 µM valproic acid (VPA; both from Stemgent, Cambridge, MA, USA) at 37 °C. The medium was changed every day for 14 days before VPA was withdrawn. The hiPSC colonies were manually picked 3 weeks after transduction and transferred onto new plates covered by the MEF feeder cells. In the Yamanaka lab, the hiPSCs were generated from aHDF cells using a previously reported reprogramming protocol [38].

Directed neural differentiation of hPSCs

R-Olig2 hESCs were subcloned from BG01 hESCs [39] and cultured in StemPro® hESC serum-free medium (Life Technologies). Non-pluripotent, SOX1-positive neural stem cells (NSCs) were differentiated from R-Olig2 hESCs as previously described [40] with minor modifications. Briefly, hESC colonies were harvested using collagenase and cultured in suspension as embryoid bodies (EBs) for 8 days in hESC medium without bFGF. EBs were then cultured for an additional 2-3 days in suspension in neural induction medium containing DMEM/F12 with Glutamax, non-essential amino acids, N2, and bFGF prior to attachment on Geltrex-coated cell culture dishes. After 2-3 days of adherent culture, neural rosettes were formed and manually isolated, dissociated into single cells with Accutase, and replated onto culture dishes coated with Geltrex (Life Technologies). The NSC population was then expanded in StemPro® NSC serum-free medium (Life Technologies) containing bFGF (20 ng/ml) for 8-10 passages. Expression of NSC biomarkers SOX1 and NESTIN in the cells was confirmed by immunofluorescence staining.

Glycomic profiling of proteins isolated from human tissues and cultured cells

The proteins of human tissues and cultured pluripotent and non-pluripotent cells were fractionated using CelLytic™ MEM protein extraction kit (Sigma-Aldrich, St Louis, MO, USA), and the manufacturer's protocol. One microgram of the hydrophobic proteins or the hydrophilic proteins of each sample was labeled with 50 µg of Alexa Fluor (AF) 555 succinimidyl ester (Life Technologies) in 20 µl of PBS containing 1% Triton X-100 (Sigma-Aldrich). The labeled proteins were purified using Zeba™ desalt spin columns (7K MWCO, Thermo Scientific, Waltham, MA, USA). The concentration of the purified proteins was adjusted to 500 ng/ml with Probing solution (GPBiosciences, Yokohama, Japan) to form chip-ready samples. LecChip™ lectin microarray chips consisting of 45 selected lectins (Supplementary information, Table S1) that exhibit binding specificities for a variety of different oligosaccharide structures were incubated with 100 µl of the chip-ready samples at 4 °C for 16 h. Before scanning the chips with GlycoStation™ Reader 1200 (GPBiosciences), the samples on the chips were re-

placed with 100 µl of fresh probing solution. Image digitizing and integration of fluorescence intensities of lectin spots from each sample were conducted using Signal Capture1.0 and Glycostation Tool Pro1.0. Since the lectin DSA has a broad specificity to oligosaccharide structures of $(GlcNAc\beta1-4)_n$ and $Gal\beta1-4$ GlcNAc commonly existing in bi-, tri- and tetra-antennary complexes of *N*-liked glycans and generates strong and relatively unchanged fluorescence intensities in most of the tested cells, we chose DSA as the standard to normalize the fluorescence intensities of 45 lectins among all samples. Hierarchical clustering analysis to determine the similarity in regard to the glycoprotein compositions among cells was performed using NIA Array Analysis Tool (National Institute on Aging, http://lgsun.grc.nia.nih.gov/ANOVA/index.html). Student's *t* test was used as the statistical method to identify the lectins showing significant binding specificity between different cell types.

Fluorescence staining of cells with specific lectins and antibodies

Lectins AOL (Tokyo Chemical Industry, Japan), and TJA-II (Seikagaku Biobusiness, Japan) were subjected to biotinylation using biotin-XX succinimidyl ester (Life Technologies). Biotinylated lectin UEA-I was purchased from Vector Laboratories (Burlingame, CA, USA). The harvested cells were resuspended in HBSS, evenly distributed on adhesion slides according to Prof. Bross (Marienfeld, Lauda-Königshofen, Germany) at 4 °C for 40 min to allow cells to attach to the slides, and then fixed in PBS containing 4% paraformaldehyde for 25 min. The cells were incubated with 6.5 µg/ml biotinylated lectins, which were pre-reacted with streptavidin-AF 488 (Life Technologies) in PBS containing 2% bovine serum albumin (BSA; Jackson ImmunoResearch, West Grove, PA, USA) at 4 °C overnight. The lectin-stained cells were permeabilized using 0.2% Triton X-100 in PBS and incubated with primary antibodies against specific pluripotent biomarkers and fluorophore-conjugated secondary antibodies. For staining for pluripotency biomarkers in iPS1.HEMd cells, colonies of the cells grown on MEF feeder cell-covered coverslips were fixed and permeabilized and incubated with primary antibodies against specific pluripotency biomarkers and fluorophore-conjugated secondary antibodies. For staining cells differentiated into each germ layer by EB formation, pluripotent cell colonies were harvested and cultured in ultra low-attachment plates with FGF-deficient DMEM/F12 medium with L-glutamine containing 20% KnockOut™ Serum Replacement, 100 µM non-essential amino acids, and 100 µM β-mercaptoethanol (all from Life Technologies) for 7 days. The medium was changed every other day. On day 8, EBs were transferred onto gelatin-coated coverslips and cultured in the same medium for an additional 7 days. EBs were then fixed, permeabilized, and incubated with antibodies against biomarkers relevant to the three germ layers. The primary antibodies used in the study were purchased from Santa Cruz (POU5F1 and SOX2), Millipore (SMA, Tra-1-60 and AFP), R&D Systems (SSEA-4 and SOX17), Stemgent (Tra-1-81), and Novus Biologicals (NG2).

Teratoma formation and histopathology

Teratomas were generated in the testis capsule according to published protocols [41] with minor modifications. Specifically, iPS1.HEMd cells were harvested using Accutase® (Life Technologies) and re-suspended in a 1:1 mixture of DMEM/F12 medium and Matrigel™ (BD Biosciences, San Diego, CA, USA). One million cells were injected into the right testis of each of two mice. H&E-stained, 4-µm sections of the tumors from two animals were evaluated microscopically by a pathologist to detect derivatives of the three germ layers. All animal procedures were performed in accordance with protocols approved by the IACUC of the Scripps Research Institute.

Lectin-mediated cell separation

Depletion of pluripotent cells from mixed cell populations was performed according to the following procedure. WA09 hESCs were mixed with HDFs or HEMd cells at a ratio of 1:1 to form a mixed cell population consisting of 1×10^6 cells in total. R-Olig2 hESCs were mixed with their differentiated NSC derivatives at the same ratio. The non-pluripotent cells were pre-labeled by 0.5 µg/ml Calcein AM (Life Technologies) in culture medium for 20 min before being mixed with the pluripotent cells. Mixed populations were incubated in biotinylated UEA-I lectin at 4 °C for 45 min, followed by addition of submicron anti-biotin beads (Miltenyi Biotec, Auburn, CA, USA) in HBSS containing 2% BSA according to the manufacturer's protocol. The cells captured by the magnetic beads were removed from the population using MACS LS Magnetic Columns (Miltenyi Biotec) to harvest the beads from the cell suspension. Residual populations were analyzed using flow cytometry as described below.

To isolate viable pluripotent cells from mixtures with fibroblasts, biotinylated UEA-I lectin was pre-incubated with streptavidin-conjugated magnetic beads (CELLection Biotin Binder Dynabeads; Life Technologies) to form UEA-I-magnetic beads. WA09 hESCs were mixed with HDFs at a ratio of 1:1 to form a mixed cell population consisting of 5×10^5 cells in total. HDFs were prelabeled with Calcein AM as described above. The mixed cells were resuspended in DMEM/F12 medium plus GlutaMax™-I (Life Technologies) containing 20 ng/ml FGF2 (Stemgent), 2% BSA, 100 µM β-mercaptoethanol (Life Technologies) and 25 µM phenylbenzodioxane carboximide (ROCK II inhibitor; Stemgent) and incubated with the UEA-I-magnetic beads at 4 °C for 3 h with continuous stirring. The cells captured by the magnetic beads were isolated using MagSweeper (Illumina, Hayward, CA, USA) to harvest the beads from the cell suspension. DNA linkers between streptavidin and beads were digested by DNase I to release the cells. The cells captured by the beads and the unbound cells left in the supernatant were resuspended separately in complete StemPro® hESC serum-free medium (Life Technologies) and placed onto plates coated with Geltrex™ matrix (Life Technologies) for overnight incubation to allow attachment of the viable cells.

Flow cytometry

For detection of residual pluripotent cells in the mixed cell populations subjected to pluripotent cell depletion, cell suspensions containing uncaptured cells were analyzed using a FACSCalibur flow cytometer (BD Biosciences) to detect Calcein AM fluorescence. For quantification of lectin binding specificity for hPSCs, 5×10^5 cells were reacted with 6.5 µg/ml UEA-I-biotin followed by streptavidin-AF 555 in HBSS containing 2% BSA. The cells were subsequently counter-stained with SSEA-4 antibody or were left unstained as controls. The cells were analyzed using a FACSCalibur flow cytometer to detect fluorescence intensity associated with the cells.

Expression profiling of glycosyltransferases by gene expression microarray

Glycosyltransferase gene expression was analyzed according to published reports [42], except that array data were normalized using the LUMI package in R with the robust spline normalization method. For differential gene expression analysis, the criterion to estimate meaningful differences was expression level differences being greater than 1.4-fold at a significance level of $\alpha_{FDR} < 0.01$.

Acknowledgments

References

1 Lawrenz B, Schiller H, Willbold E, Ruediger M, Muhs A, Esser S. Highly sensitive biosafety model for stem-cell-derived grafts. *Cytotherapy* 2004; **6**:212-222.

2 Andrews PW, Banting G, Damjanov I, Arnaud D, Avner P. Three monoclonal antibodies defining distinct differentiation antigens associated with different high molecular weight polypeptides on the surface of human embryonal carcinoma cells. *Hybridoma* 1984; **3**:347-361.

3 Kannagi R, Cochran NA, Ishigami F, *et al*. Stage-specific embryonic antigens (SSEA-3 and -4) are epitopes of a unique globo-series ganglioside isolated from human teratocarcinoma cells. *EMBO J* 1983; **2**:2355-2361.

4 Kannagi R, Levery SB, Ishigami F, *et al*. New globoseries glycosphingolipids in human teratocarcinoma reactive with the monoclonal antibody directed to a developmentally regulated antigen, stage-specific embryonic antigen 3. *J Biol Chem* 1983; **258**:8934-8942.

5 Kolle G, Ho M, Zhou Q, *et al*. Identification of human embryonic stem cell surface markers by combined membrane-polysome translation state array analysis and immunotranscriptional profiling. *Stem Cells* 2009; **27**:2446-2456.

6 Lin SA, Kolle G, Grimmond SM, *et al*. Subfractionation of differentiating human embryonic stem cell populations allows the isolation of a mesodermal population enriched for intermediate mesoderm and putative renal progenitors. *Stem Cells Dev* 2010; **19**:1637-1648.

7 Pera MF, Blasco-Lafita MJ, Cooper S, Mason M, Mills J, Monaghan P. Analysis of cell-differentiation lineage in human teratomas using new monoclonal antibodies to cytostructural antigens of embryonal carcinoma cells. *Differentiation* 1988; **39**:139-149.

8 Apweiler R, Hermjakob H, Sharon N. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim Biophys Acta* 1999; **1473**:4-8.

9 Reis CA, Osorio H, Silva L, Gomes C, David L. Alterations in glycosylation as biomarkers for cancer detection. *J Clin Pathol* 2010; **63**:322-329.

10 Kakugawa Y, Wada T, Yamaguchi K, *et al*. Up-regulation of plasma membrane-associated ganglioside sialidase (Neu3) in human colon cancer and its involvement in apoptosis suppression. *Proc Natl Acad Sci USA* 2002; **99**:10718-10723.

11 Moody AM, Chui D, Reche PA, Priatel JJ, Marth JD, Reinherz EL. Developmentally regulated glycosylation of the CD8 alpha beta coreceptor stalk modulates ligand binding. *Cell* 2001; **107**:501-512.

12 Fogel AI, Li Y, Giza J, *et al*. N-glycosylation at the SynCAM (synaptic cell adhesion molecule) immunoglobulin interface modulates synaptic adhesion. *J Biol Chem* 2010; **285**:34864-34874.

13 Reid CW, Fulton KM, Twine SM. Never take candy from a stranger: the role of the bacterial glycome in host-pathogen interactions. *Future Microbiol* 2010; **5**:267-288.

14 Balog CI, Mayboroda OA, Wuhrer M, Hokke CH, Deelder AM, Hensbergen PJ. Mass spectrometric identification of aberrantly glycosylated human apolipoprotein C-III peptides in urine from Schistosoma mansoni-infected individuals. *Mol Cell Proteomics* 2010; **9**:667-681.

15 Haltiwanger RS. Regulation of signal transduction pathways in development by glycosylation. *Curr Opin Struct Biol* 2002; **12**:593-598.

16 Xia L, McDaniel JM, Yago T, Doeden A, McEver RP. Surface fucosylation of human cord blood cells augments binding to P-selectin and E-selectin and enhances engraftment in bone marrow. *Blood* 2004; **104**:3091-3096.

17 Satomaa T, Heiskanen A, Mikkola M, *et al*. The N-glycome of human embryonic stem cells. *BMC Cell Biol* 2009; **10**:42.

18 Venable A, Mitalipova M, Lyons I, *et al*. Lectin binding profiles of SSEA-4 enriched, pluripotent human embryonic stem cell surfaces. *BMC Dev Biol* 2005; **5**:15.

19 Kuno A, Uchiyama N, Koseki-Kuno S, *et al*. Evanescent-field fluorescence-assisted lectin microarray: a new strategy for glycan profiling. *Nat Methods* 2005; **2**:851-856.

20 Pilobello KT, Krishnamoorthy L, Slawek D, Mahal LK. Development of a lectin microarray for the rapid analysis of protein glycopatterns. *Chembiochem* 2005; **6**:985-989.

21 Tateno H, Toyota M, Saito S, *et al*. Glycome diagnosis of human induced pluripotent stem cells using lectin microarray. *J Biol Chem* 2011; **286**:20345-20353.

22 Toyoda M, Yamazaki-Inoue M, Itakura Y, *et al*. Lectin microarray analysis of pluripotent and multipotent stem cells. *Genes Cells* 2011; **16**:1-11.

23 Honore C, Rorvig S, Hummelshoj T, Skjoedt MO, Borregaard N, Garred P. Tethering of Ficolin-1 to cell surfaces through recognition of sialic acid by the fibrinogen-like domain. *J Leukoc Biol* 2010; **88**:145-158.

24 Marquardt T, Brune T, Luhn K, *et al.* Leukocyte adhesion deficiency II syndrome, a generalized defect in fucose metabolism. *J Pediatr* 1999; **134**:681-688.

25 Rossi SL, Nistor G, Wyatt T, *et al.* Histological and functional benefit following transplantation of motor neuron progenitors to the injured rat spinal cord. *PLoS One* 2010; **5**:e11852.

26 Becker DJ, Lowe JB. Fucose: biosynthesis and biological function in mammals. *Glycobiology* 2003; **13**:41R-53R.

27 Harduin-Lepers A. Comprehensive analysis of sialyltransferases in vertebrate genomes. *Glycobiol Insights* 2010; **2010**:29-61.

28 Draber P, Stanley P. Cytotoxicity of plant lectins for mouse embryonal carcinoma cells. *Somat Cell Mol Genet* 1984; **10**:435-443.

29 Kosmehl H, Langbein L, Katenkamp D. Lectin histochemistry of human testicular germ cell tumors. *Neoplasma* 1989; **36**:29-39.

30 Lee MC, Talerman A, Oosterhuis JW, Damjanov I. Lectin histochemistry of classic and spermatocytic seminoma. *Arch Pathol Lab Med* 1985; **109**:938-942.

31 Muramatsu T, Gachelin G, Damonneville M, Delarbre C, Jacob F. Cell surface carbohydrates of embryonal carcinoma cells: polysaccharidic side chains of F9 antigens and of receptors to two lectins, FBP and PNA. *Cell* 1979; **18**:183-191.

32 Teshima S, Hirohashi S, Shimosato Y, *et al.* Histochemically demonstrable changes in cell surface carbohydrates of human germ cell tumors. *Lab Invest* 1984; **50**:271-277.

33 Traving C, Schauer R. Structure, function and metabolism of sialic acids. *Cell Mol Life Sci* 1998; **54**:1330-1349.

34 Landmesser L, Dahm L, Tang JC, Rutishauser U. Polysialic acid as a regulator of intramuscular nerve branching during embryonic-development. *Neuron* 1990; **4**:655-667.

35 Shi SL, Stanley P. Protein O-fucosyltransferase 1 is an essential component of Notch signaling pathways. *Proc Natl Acad Sci USA* 2003; **100**:5234-5239.

36 Schwarzkopf M, Knobeloch KP, Rohde E, *et al.* Sialylation is essential for early development in mice. *Proc Natl Acad Sci USA* 2002; **99**:5267-5270.

37 Spiro RG. Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology* 2002; **12**:43R-56R.

38 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.

39 Xue H, Wu S, Papadeas ST, *et al.* A targeted neuroglial reporter line generated by homologous recombination in human embryonic stem cells. *Stem Cells* 2009; **27**:1836-1846.

40 Swistowski A, Peng J, Han Y, Swistowska AM, Rao MS, Zeng X. Xeno-free defined conditions for culture of human embryonic stem cells, neural stem cells and dopaminergic neurons derived from them. *PLoS One* 2009; **4**:e6233.

41 Gertow K, Przyborski S, Loring JF, *et al.* Isolation of human embryonic stem cell-derived teratomas for the assessment of pluripotency. *Curr Protoc Stem Cell Biol* 2007; **Chapter 1**:Unit1B.4.

42 Ku S, Soragni E, Campau E, *et al.* Friedreich's ataxia induced pluripotent stem cells model intergenerational GAATTC triplet repeat instability. *Cell Stem Cell* 2010; **7**:631-637.

(**Supplementary information** is linked to the online version of the paper on the *Cell Research* website.)