Stem Cell Reports
Article
OPEN ACCESS

A Human Pluripotent Stem Cell Surface N-Glycoproteome Resource Reveals Markers, Extracellular Epitopes, and Drug Targets

Kenneth R. Boheler,[1,2,3,*] Subarna Bhattacharya,[4] Erin M. Kropp,[4] Sandra Chuppa,[4] Daniel R. Riordon,[2] Damaris Bausch-Fluck,[10] Paul W. Burridge,[5] Joseph C. Wu,[5] Robert P. Wersto,[2] Godfrey Chi Fung Chan,[1,6] Sridhar Rao,[7,8,9] Bernd Wollscheid,[10] and Rebekah L. Gundry[4,*]

[1]Stem Cell and Regenerative Medicine Consortium, LKS Faculty of Medicine, Hong Kong University, Hong Kong, SAR
[2]National Institute on Aging, National Institutes of Health, Baltimore, MD 21224, USA
[3]Division of Cardiology, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA
[4]Department of Biochemistry, Medical College of Wisconsin, Milwaukee, WI 53226, USA
[5]Stanford Cardiovascular Institute, Stanford University School of Medicine, Stanford, CA 94305, USA
[6]Department of Pediatrics & Adolescent Medicine, Hong Kong University, Hong Kong, SAR
[7]Department of Pediatrics, Medical College of Wisconsin, Milwaukee, WI 53226, USA
[8]Department of Cell Biology, Neurobiology and Anatomy, Medical College of Wisconsin, Milwaukee, WI 53226, USA
[9]Blood Research Institute, Blood Center of Wisconsin, Milwaukee, WI 53226, USA
[10]Department of Biology, Institute of Molecular Systems Biology, Swiss Federal Institute of Technology (ETH) Zurich, Wolfgang-Pauli-Strasse 16, 8093 Zurich, Switzerland
*Correspondence: bohelerk@hku.hk (K.R.B.), rgundry@mcw.edu (R.L.G.)
http://dx.doi.org/10.1016/j.stemcr.2014.05.002

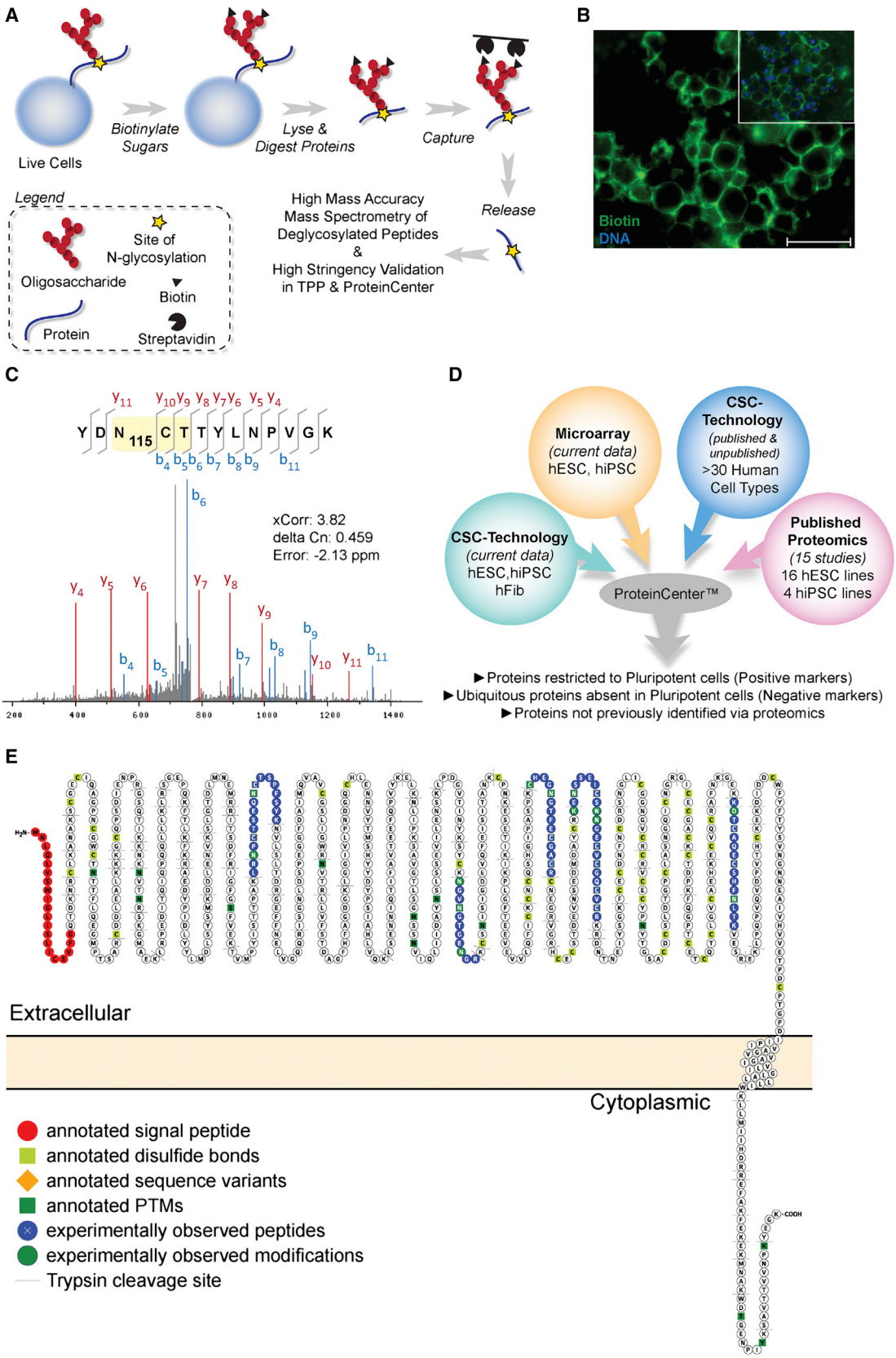
SUMMARY

Detailed knowledge of cell-surface proteins for isolating well-defined populations of human pluripotent stem cells (hPSCs) would significantly enhance their characterization and translational potential. Through a chemoproteomic approach, we developed a cell-surface proteome inventory containing 496 N-linked glycoproteins on human embryonic (hESCs) and induced PSCs (hiPSCs). Against a backdrop of human fibroblasts and 50 other cell types, >100 surface proteins of interest for hPSCs were revealed. The >30 positive and negative markers verified here by orthogonal approaches provide experimental justification for the rational selection of pluripotency and lineage markers, epitopes for cell isolation, and reagents for the characterization of putative hiPSC lines. Comparative differences between the chemoproteomic-defined surfaceome and the transcriptome-predicted surfaceome directly led to the discovery that STF-31, a reported GLUT-1 inhibitor, is toxic to hPSCs and efficient for selective elimination of hPSCs from mixed cultures.

INTRODUCTION

Human pluripotent stem cells (PSCs) can differentiate into nearly all somatic cell types present in the human body and can generate clinically relevant numbers of cells for regenerative medicine. The advent of hiPSCs, derived from somatic cells by the exogenous expression of defined transcription factors, has overcome ethical issues associated with human embryonic stem cells (hESCs) and, when derived from the patient, may avoid immunological complications. Human iPSCs have also opened new avenues of research for the study of basic disease mechanisms and development of informative model systems for drug discovery. Although promising, significant limitations to the therapeutic use of hiPSCs remain unresolved. These include interline variations ranging from inconsistent transcription factor expression and differential DNA methylation to sporadic point mutations and chromosomal defects that affect in vitro differentiation, tumorigenicity, and potential clinical applications (Feng et al., 2010; Gore et al., 2011; Robinton and Daley, 2012). Moreover, current tests of hiPSC potency rely on extensive in vitro differentiation tests, in vivo teratoma assays in rodents (Maherali and Hochedlinger, 2008; Robinton and Daley, 2012) or bioinformatic and gene expression assays (Bock et al., 2011; Müller et al., 2011), which cannot be practically implemented into high-throughput hiPSC line generation designed to limit interline variability.

The lack of suitable cell-surface marker panels and related affinity-based reagents for isolating high-quality hiPSCs and well-defined progeny significantly restricts our ability to minimize interline variability and employ hiPSCs for regenerative medicine. Although guidelines and animal-free methods have been proposed for the derivation and characterization of therapeutic and good manufacturing practice compliant hiPSCs (Buta et al., 2013; Funk et al., 2012; Maherali and Hochedlinger, 2008; Müller et al., 2010), no system is available to overcome safety and efficacy issues of hiPSCs analogous to immunophenotyping of blood lineages for identifying and isolating hematopoietic stem cells (HSCs). Although markers such as SSEA-3, SSEA-4, Tra-1-60, and Tra-1-81 aid in the identification of hPSCs, few known surface markers and application-specific antibodies are restricted to the pluripotent state (Damjanov et al., 1982; Kannagi et al., 1983; Lowry et al., 2008).



(legend on next page)

Moreover, as cell-surface proteins play critical roles in inter- and intracellular communication, a better understanding of the cell surface should inform the dynamic interplay between cells and their microenvironment that ultimately regulates how hPSCs interact with and respond to external cues and differentiate in a directed manner (Lian et al., 2013; Murry and Keller, 2008; Yan et al., 2005). Coupling this functional relevance with the fact that more than 60% of US Food and Drug Administration-approved drug therapies target membrane proteins, and 38% of disease-related proteins are membrane associated (Cheng et al., 2012; Yildirim et al., 2007), we aimed to generate a new resource derived from a targeted analytical approach, *C*ell *S*urface *C*apture (CSC) technology (Gundry et al., 2009, 2012; Hofmann et al., 2010; Wollscheid et al., 2009) that experimentally verifies extracellular domains of surface proteins. This resource, through its direct protein measurements, will foster the identification of proteins and epitopes useful for immunophenotyping and facilitate the identification of drugs that target hPSCs. The value of this resource is exemplified by the identification of proteins capable of marking live hPSCs and the identification of a small molecule inhibitor, STF-31, that allows for selective depletion of hPSCs from hESC and hiPSC-derived progeny.

RESULTS

Human Pluripotent Cell-Surface N-Glycoproteome

We employed the chemoproteomic CSC technology to directly measure and define the average N-glycoprotein surface proteome of hiPSCs, hESCs, and human fibroblasts (hFibs) (Figure 1A). Biotin labeling of glycoproteins was specific to the cell surface (Figure 1B), and all reported glycoproteins detected by this technology were identified by peptides containing a deamidation (resulting from PNGaseF cleavage after glycosylation-based enrichment) at the consensus amino acid sequence motif (NxS/T) for N-glycosylation (Figure 1C). Altogether, 496 N-glycoproteins were identified with a false discovery rate (FDR) $\leq$5% on the cell surface of hPSCs (Figure 2; Table S1 available online), composed of 92 cluster of differentiation (CD) and 411 non-CD molecules, including predicted transmembrane (92%), Glycosylphosphatidylinisotol (GPI)-linked (4%), and extracellular matrix proteins (4%). A complete catalog of extracellular glycopeptides and glycosites is provided in Tables S2, S3, and S4, and data were compiled into ProteinCenter (Figure 1D). These data provide direct experimental evidence regarding surface localization, which is critical for proteins whose subcellular localization is unknown or ambiguous, and transmembrane orientation for proteins whose structure has not yet been resolved. LRRN1, a protein with generic "membrane" annotation but not currently linked to gene ontology (GO) terms for the cell surface or plasma membrane (Figure 1E), and FAM216A, a protein with previously unannotated subcellular localization (UniProt), are examples of where CSC technology data provide new experimentally derived localization information, and immunostaining corroborates the surface localization of these proteins, thus confirming the validity of CSC-technology-based annotations. The discovery approach used in this study is limited to the identification of N-glycosylated peptides; therefore, it is not uncommon for proteins to be identified by a single N-glycopeptide (e.g., proteins with only a single N-glycosite in their amino acid sequence). Additional confidence in all protein identifications is provided by the use of high mass accuracy mass spectrometry during data acquisition and the implementation of three independent search algorithms for data processing. Examples of proteins that were identified by a single N-glycopeptide, but for which we provide positive antibody-based evidence include ADCYAP1R1, EFNA3, FAM216A, HTR2C, IL27RA, NPR1, PVRL2, and SLC7A3. Compared to published proteomic studies of hPSCs (i.e., compilation of 15 studies; herein referred to as the "PSC proteome" (Gundry et al., 2011), 158 proteins (31% of the current data) have not been reported among the >9,500 proteins previously described in hPSCs, and 106

Figure 1. CSC Technology Workflow

(A) Simplified schematic workflow of the CSC technology.

(B) Immunofluorescence image of hESC colony after cell-surface biotin labeling via the CSC technology, illustrating selective extracellular biotinylation. Blue, DNA (Hoechst); green, Biotin (Streptavidin-FITC). Scale bar, 200 μm.

(C) Annotated MS/MS spectrum of one peptide from Interleukin-17 receptor D illustrating the deamidation (N_{115}) within the N-glycosite sequence motif (highlighted), which collectively represents the data "tag" used for filtering out non-cell-surface contaminants from the final protein list.

(D) Bioinformatics workflow for merging multiple data types for identifying cell-surface proteins of interest for downstream characterization.

(E) Graphical representation of the coverage of N-glycopeptides and confirmation of extracellular domain for LRRN1, a previously "generically" annotated membrane protein without GO terms linking it to the cell surface. Image generated using Protter (http://wlab.ethz.ch/protter) (Omasits et al., 2013).

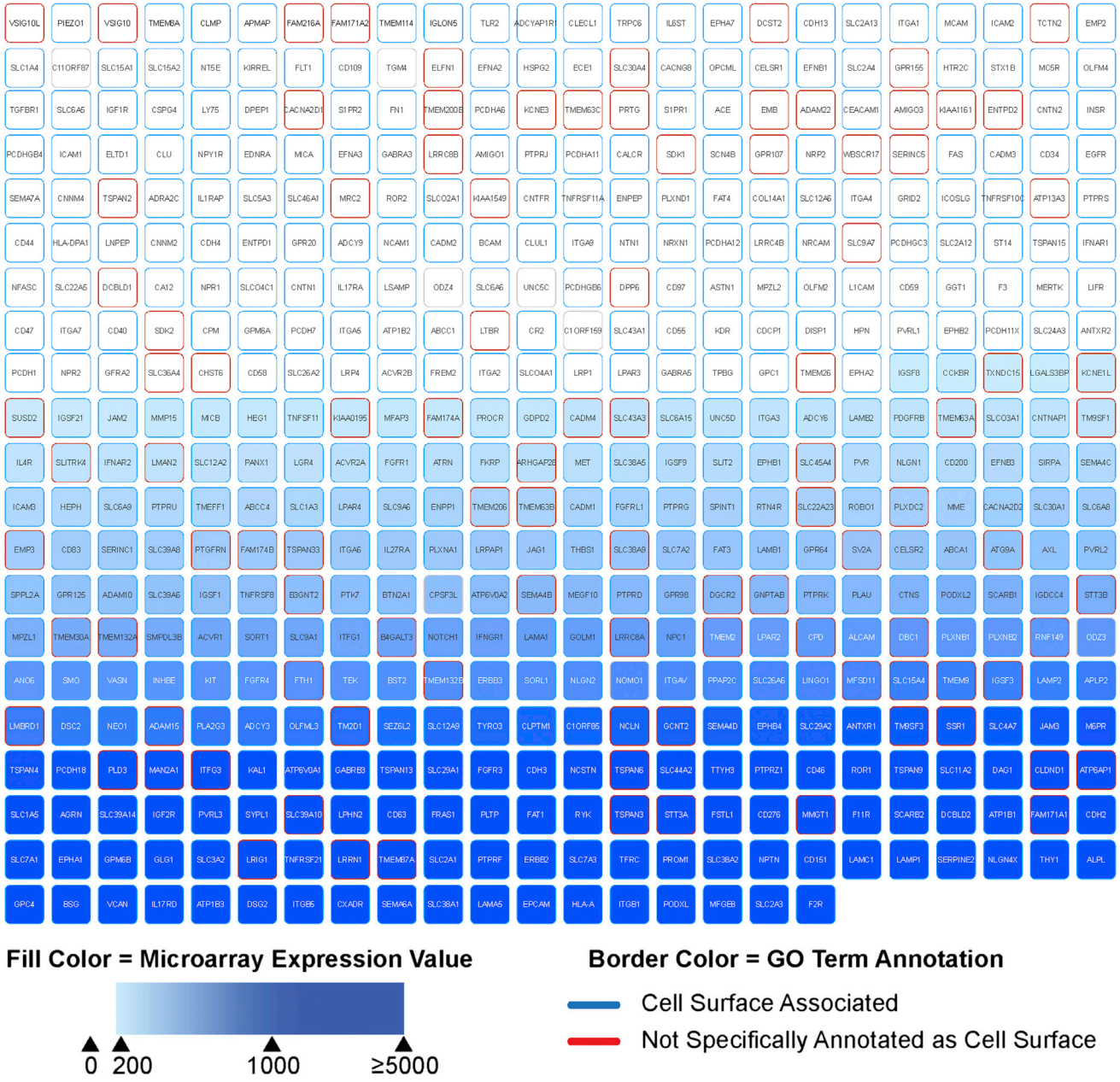


Figure 2. The hPSC Cell-Surface N-Glycoproteome

All 496 N-glycoproteins identified in hPSCs are listed in order of their increasing mRNA expression (top left to right bottom). For simplicity, the microarray values represented here are an average among all replicates of H9 hESC and KB3 hiPSC. The fill color of each box is representative of transcript abundance (unlogged microarray value), where white indicates values below 200. The border color indicates Gene Ontology (GO) protein annotations of cell-surface localization. Proteins represented by white boxes with red borders are examples of proteins identified via the CSC technology but would otherwise be unlikely to be categorized as potential surface markers due to their low transcriptional expression and lack of surface localization GO annotation.

proteins (21% of the current data) were identified in only a single other study within the PSC proteome (Gundry et al., 2011). Consistent with the PSC phenotype, EPCAM and ALPL were observed exclusively in hPSCs, whereas others such as Thy1 (CD90) were also observed in fibroblasts.

Surface Proteome Comparisons Reveal Putative hPSC-Restricted Markers

Putative markers for positive selection were identified by comparing the surface proteome of hPSCs with those from hFibs, nondiseased somatic cell types (n = 12) and

cancer cell types (n = 38) present in the Cell Surface Protein Atlas (see Table S5). As selection criteria, putatively restricted PSC markers were required to be absent from hFibs, present in no more than three somatic cell types, but were not limited by the number of cancer cell types in which they were observed, as similarities between cancer and pluripotent cells are known (Kim et al., 2010; Ratajczak et al., 2013; Riggs et al., 2013). Two hundred proteins met these selection criteria, and 90 of these were uniquely identified in this study, whereas 47 others were reported in only one other proteomic study (Table S6). Of these 200 PSC-restricted proteins, 48 were uniquely identified in hPSCs when compared to other cells within the Atlas (Table 1). We provide surface immunostaining for eight of the 90 hPSC-restricted surface proteins uniquely identified here (ADCYAP1R1, EFNA3, FAM216A, FGFR3, HTR2C, IL17RD, NPR1, and OPCML; Figures 3 and 4), as well as for ten hPSC-restricted proteins observed among the 15 previous proteomic reports (CXADR [seven other studies], EPHA1 [five]; IL27RA [one]; LINGO1 [one]; LRRN1 [three]; PTPRZ1 [four]; SEMA6A [six]; SPINT1 [three]; SLC7A3 [four]; and UNC5D [one]). Eight additional cell-surface proteins (ACVR2A, F3, ICAM3, ITGA6, NRCAM, PCDH1, PVRL2, and UNC5C) not hPSC-restricted are also illustrated in Figures 3 and 4 and S1. In all cases, antibody-based analyses were consistent with surface localization determined by the CSC technology. Moreover, costaining of fixed cells illustrates that surface expression colocalizes to OCT4-positive cells (Figure 3A), and altogether we demonstrate the ability of 28 antibodies to successfully stain live hPSCs either by immunocytochemistry (Figures 3B and S1) or flow cytometry and confirm the absence of hPSC-restricted markers on hFibs (Figure 4).

Surface Proteome Comparisons Reveal Negative hPSC Markers

Negative selection markers are not well established for hPSCs, even though positive selection carries the risk of functional changes with antibody binding (e.g., CD3 antibody binding promotes T cell activation; Berg et al., 1998). To identify putative markers for negative cell selection, we focused on proteins present in hFib, present in six or more nondiseased cell types in the CSPA, and which were not observed in hPSCs via the CSC technology. This led to the classification of 15 proteins as potential negative markers of selection of hPSCs from coculture with hFibs (Table 1). Flow cytometry confirmed the absence of three selected proteins (DPP4 [CD26], NRP1 [CD304], COLEC12) on hESCs and hiPSCs that were highly immunoreactive on hFibs (Figure 4C). Moreover, we classified aminopeptidase N (ANPEP, CD13) as a negative marker of pluripotency, consistent with a recent report showing successful fluorescence activated cell sorting (FACS) of hiPSCs after reprogramming (Kahler et al., 2013). These data also reveal discrepancies between mouse and human PSCs. We previously identified NRP1, COLEC12, and CD13 in mouse PSCs (Gundry et al., 2012), but, based on CSC technology and antibody-based analyses here and (Kahler et al., 2013), these surface proteins are absent from hPSCs.

Theoretical Surface Proteome versus CSC Technology

A comparison of the theoretical surface proteome based on transcriptomic data to that defined by the CSC technology underscores why proteomic strategies are essential for providing direct evidence for the location and quantity of cell-surface proteins (Figure 2). For example, proteins such as ADCYAP1R1, EFNA3, HTR2C, NPR1, and UNC5D would have been overlooked as potential surface markers if relying solely on transcriptomic approaches. By microarray, these transcripts were detected at only background/threshold levels (Table S1), even though robust protein levels were detected using Ab-based techniques. Moreover, proteins such as FAM216A may be overlooked as potential surface markers as their GO terms do not specifically indicate they are localized to the plasma membrane. Surface proteins are affected by nontranscriptional mechanisms of regulation (e.g., surface proteins shuttle among cellular compartments, protein half-life independent of mRNA expression timeline (Kristensen et al., 2013; Vogel and Marcotte, 2012), and these types of posttranscriptional regulatory mechanisms may be broadly important as other surface proteins identified here with similarly low mRNA levels have known critical functions in early embryonic development or pluripotency (e.g., LPAR2, 3, and 4 [Liu and Armant, 2004] and TLR2 [Taylor et al., 2010]). Additionally, the microarray data may highlight a number of other PSC-restricted proteins with functional relevance to pluripotency due to their relatively high mRNA levels (e.g., APLP2, FRAS1, FTH1, GPC4, GCNT2, NLGN4X, PTPRZ1, SEMA6A, SLC38A2, SLC7A3, and VCAN). Although several cell-surface-annotated proteins with microarray values above 5,000 were not detected via CSC technology, further inspection reveals that GO term assignments for these can be misleading (e.g., TUBB3, MYH9 have GO terms including plasma membrane, even though these are not cell-surface-accessible proteins), and others do not contain an N-glycosylation site (e.g., ENO1, KARS, and SETP2) explaining why they were not detected in the current study, though alternative targeted proteome strategies such as Cys-CSC and K-CSC (Hofmann et al., 2010) would likely capture them if they are truly located at the cell surface.

hPSC Cell-Surface Protein Resource Enables Rational Selection of Drug Targets

The facilitated glucose transporter family is one subset of proteins that came to light as a result of the perceived

Table 1. Cell-Surface Proteins of Interest

Gene Symbol for Protein	Mouse ESC	Mouse iPSC	hESC	hiPSC	Human Fibroblasts	No. of Nondisease Human Cell Types in Cell Surface Protein Atlas (out of 14)	No. of Human Cancer Cell Types in CSPA (out of 38)	No. of Human Pluripotent Stem Cell Proteomic Studies (out of 15; Gundry et al., 2011)	Promoter Occupancy		
									Nanog	Oct4	Sox2
Reference Markers											
ALPL	●	●	●	●		5	5	7	●		
EPCAM (CD326)	●	●	●	●		1	4	7	●	●	
THY1 (CD90)	●	●	●	●	●	6	22	5		●	
Pluripotency Restricted											
ADAM22			●	<		1	13	2			
ADCYAP1R1			●	●		0	0	0			
AMIGO3			●	●		0	0	1			
APLP2			●			0	22	3		●	
ARHGAP28			●			0	0	3			
ASTN1			●	●		0	0	0	●	●	
ATP13A3	●	●		●		0	0	2			
ATP6V0A2	●		●			1	16	5			
B4GALT3	●	●	●			1	11	1			
C11ORF87				●		0	0	0			
CACNG8		●	●	●		0	0	0		●	
CALCR	●			●		1	0	0	●		
CCKBR			●	●		1	0	1	●		
CD83				●		0	17	0			
CDH3	●		●	●		1	1	3			
CELSR2	●		●	●		1	6	1			
CHST6			●			1	6	0			
CLECL1			●			0	0	0			
CLUL1			●			0	0	0		●	
CNTFR	●	●	●	●		1	3	0		●	
CNTN2	●			●		1	0	0			
CPSF3L				●		1	0	1			
DBC1			●			0	0	0	●		
DPEP1			●	●		1	2	0		●	
DPP6			●	●		0	0	0	●	●	
EFNA3	●		<[a]	●		0	0	0			
ELFN1			●	●		1	9	1			

(*Continued on next page*)

Table 1. ***Continued***

Gene Symbol for Protein	Mouse ESC	Mouse iPSC	hESC	hiPSC	Human Fibroblasts	No. of Nondisease Human Cell Types in Cell Surface Protein Atlas (out of 14)	No. of Human Cancer Cell Types in CSPA (out of 38)	No. of Human Pluripotent Stem Cell Proteomic Studies (out of 15; Gundry et al., 2011)	Promoter Occupancy		
									Nanog	Oct4	Sox2
ENTPD2		●	●	●		1	3	1			
EPHA7				●		1	0	0		●	
FAM216A				●		0	0	0			
FGFR4 (CD334)			●	●		1	5	1			
FKRP	●	●	●			0	11	2			
FLT1			●	●		0	1	1		●	
FTH1			●			1	0	1			
GABRA3			●	●		0	0	0			
GABRA5				●		0	0	1		●	
GABRB3			●	●		0	1	3			
GCNT2			●			0	0	3	●	●	●
GDPD2			●	●		1	0	0	●	●	
GNPTAB	●	●	●			1	22	4	●		
GPM6A				●		0	0	0		●	
GPR20			●	●		0	0	0			
GPR64			●	●		1	2	0			
GPR98			●	●		0	0	0	●	●	●
GRID2			●	●		0	0	0	●	●	●
HEPH			●	●		1	0	0			
HPN			●			0	0	0			
HTR2C			●	●		0	0	0			
IGSF1			●	●		1	2	2	●		
IGSF21			●	●		0	0	1			
IGSF9			●	●		0	0	2		●	
IL17RD	●	●	●	●		1	5	0	●	●	●
IL27RA			●	●		0	9	1			
IL4R (CD124)			●	●		1	7	2			
INHBE			●	●		1	2	0			
KAL1			●	●		0	0	0			
KCNE1L			●	●		1	0	0			
KCNE3			●	●		0	3	0			
KDR (CD309)			●	●		1	0	3			

(*Continued on next page*)

Table 1. *Continued*

Gene Symbol for Protein	Mouse ESC	Mouse iPSC	hESC	hiPSC	Human Fibroblasts	No. of Nondisease Human Cell Types in Cell Surface Protein Atlas (out of 14)	No. of Human Cancer Cell Types in CSPA (out of 38)	No. of Human Pluripotent Stem Cell Proteomic Studies (out of 15; Gundry et al., 2011)	Promoter Occupancy		
									Nanog	Oct4	Sox2
KIT (CD117)	●		●	●		1	1	2		●	
LAMA1	●	●	●	●		1	9	6			●
LAMB2	●		●	●		1	2	2		●	
LGR4	●			●		1	0	2		●	
LINGO1			●	●		1	0	1	●	●	●
LPAR3			●	●		0	0	0	●		●
LPAR4	●	●	●	●		1	0	1	●		
LRIG1				●		1	0	4	●	●	●
LRP4	●	●	●	●		1	4	3	●	●	
LRRN1			●	●		1	4	3	●	●	●
MC5R				●		0	0	0			
MEGF10				●		0	0	0	●	●	
MFSD11			●	●		0	0	0			
MICB			●	●		1	9	1	●		
MMGT1			●	●		0	1	5			
MMP15			<	●		1	0	0	●		●
NCLN	●	●	●			0	1	6			
NLGN4X			●	●		0	3	2			
NPR1	●		●	●		1	1	0			
NPY1R			●	●		1	1	1			
NRXN1			●	●		0	0	0		●	
NTN1				●		1	0	0	●	●	●
OLFM2			●			0	0	0		●	
OLFM4			●			0	0	0			
OLFML3			●	●		0	4	1	●	●	●
PCDH11X			●	●		0	0	1			
PLA2G3				●		0	1	1		●	
PLXDC2	●		●	●		0	0	1		●	
PODXL2			●	●		1	6	1			
PTPRD				●		0	3	3			
PTPRU				●		0	0	2		●	
PTPRZ1	●		●	●		0	5	4		●	●

(*Continued on next page*)

Table 1. ***Continued***

Gene Symbol for Protein	Mouse ESC	Mouse iPSC	hESC	hiPSC	Human Fibroblasts	No. of Nondisease Human Cell Types in Cell Surface Protein Atlas (out of 14)	No. of Human Cancer Cell Types in CSPA (out of 38)	No. of Human Pluripotent Stem Cell Proteomic Studies (out of 15; Gundry et al., 2011)	Promoter Occupancy		
									Nanog	Oct4	Sox2
SEMA4B	●	●	●	●		1	16	1		●	
SEMA6A			●	●		0	0	6	●	●	
SLC15A1	●		●	●		0	0	1			
SLC15A2			●	●		0	1	0	●		●
SLC24A3				●		1	0	0	●	●	●
SLC2A12			●			0	1	0	●	●	●
SLC30A4			●			0	0	0			
SLC38A2	●	●	●	●		1	0	4			
SLC38A5			●			0	0	2			
SLC38A9	●		●	●		1	2	2			
SLC43A1	●	●		●		1	6	1			
SLC6A5			●	●		0	0	0		●	
SLC7A3		●	●	●		1	0	6			
SLITRK4			●	●		0	3	2			
SPINT1	●	●	●	●		1	0	3			
ST14	●	●	●	●		0	6	1			
STX1B			●	<		1	0	2			
TLR2 (CD282)	●		<	●		1	6	0		●	
TMEM114			●	●		1	0	0	●		
TMEM132B			●			1	1	1	●		●
TMEM200B				●		1	3	1	●	●	
TMEM63A				●		1	3	3			
TMEM63C			●	●		1	3	0			
TNFRSF10C (CD263)			●			0	0	0			
TNFRSF21				●		0	4	1		●	●
TNFRSF8 (CD30			●	●		1	10	2			
TNFSF11 (CD254)			●	●		0	0	0			
UNC5D			●	●		0	0	1		●	
VSIG10	●	●	●	●		1	5	2	●		
VSIG10L			●	●		0	0	0			
WBSCR17	●	●	●	<		0	1	2			

(*Continued on next page*)

Table 1. *Continued*

Gene Symbol for Protein	Mouse ESC	Mouse iPSC	hESC	hiPSC	Human Fibroblasts	No. of Nondisease Human Cell Types in Cell Surface Protein Atlas (out of 14)	No. of Human Cancer Cell Types in CSPA (out of 38)	No. of Human Pluripotent Stem Cell Proteomic Studies (out of 15; Gundry et al., 2011)	Promoter Occupancy		
									Nanog	Oct4	Sox2
Pluripotency: Putative Negative Markers											
ANPEP (CD13)	●	●			●	7	14	2			
BTN3A3					●	9	24	0			
COLEC12	●	●			●	7	11	1			
DPP4 (CD26)					●	9	16	1	●		●
EMR2 (CD312)					●	7	21	0			
FAP					●	6	12	1	●		
FBN1					●	6	8	2			
GPR126	●				●	6	13	0			
ICAM5					●	6	23	0			
NRP1 (CD304)	●	●			●	8	19	0		●	
PLAUR					●	6	20	1	●		
PRNP (CD230)	●	●			●	10	37	2			
PTPRM	●				●	7	12	1			

Pluripotency restricted were identified in one or fewer nondiseased cell types from the Cell Surface Protein Atlas. Negative markers were identified in human fibroblasts and six or more nondiseased cell types from the Atlas, but not in human pluripotent cells. Mouse ESC and iPSC data are from Gundry et al. (2012).
[a]For the hESC and hiPS surface protein data from the current study, "<" indicates the protein was identified with a probability of <0.5 in that data set.

discrepancies observed between the transcriptome analyses and the CSC technology data. We observed four members of this family (SLC2A1 [GLUT1], SLC2A3 [GLUT3], SLC2A4 [GLUT4], and SLC2A12 [GLUT12]) in hESCs and hiPSCs via the CSC technology, of which three are summarized in Figure 5A. Surface expression of GLUT1 and GLUT4 were observed by flow cytometry or immunocytochemistry (Figures 5B and 5C). By microarray, GLUT3 mRNA is three to five times higher than that of GLUT1 (Figure 5A), an observation validated by quantitative real-time PCR (Figure S2); however, significantly more MS/MS spectra for GLUT1 (487 MS/MS spectra in hESC) were observed than for GLUT3 (41 spectra) and GLUT4 (32 spectra) (Figure 5A). Protein sequences for GLUT1, GLUT3, and GLUT4 each contain a single site of N-glycosylation (UniProt) and were identified by one unique peptide sequence, whereas GLUT12 was identified by a single peptide containing three of the four possible N-glycosylation sites. Of course, the comparison of the total number of MS/MS spectra identified is only a crude measure of relative protein abundance, and it is typically inappropriate to apply spectral counting methods for relative abundance among different proteins within a sample (rather than one protein among multiple samples) due to variation in number of tryptic peptides that could result from different protein sequences. Moreover, changes in glycan structure could affect our ability to efficiently capture N-glycopeptides. Despite these caveats and in contradiction to the transcriptomic data, the relative number of spectra identified for GLUT1, GLUT3, and GLUT4 may reflect an authentic difference in surface abundance among these proteins because only a single peptide sequence is expected and explains why this protein family came to our attention.

Knowing that GLUT1 is critical in early mouse embryonic development (Chi et al., 2000; Mobasheri et al., 2005), these observations prompted us to address whether surface expression of GLUT1 is required for hPSC biology using the chemical STF-31, a reported GLUT1 inhibitor (Chan et al., 2011). The media from STF-31-treated hPSCs do not show visible changes in pH during the first 48 hr of treatment, relative to vehicle controls, confirming that overall cell metabolism is reduced (Figure S2). Although human mesenchymal stem cell (MSCs), hESC-cardiomyocytes (CMs), and hFibs exhibit no adverse effects to STF-31 treatment at any time point examined (Figure 5D), treatment with STF-31 results in 80% cell death in hPSC by

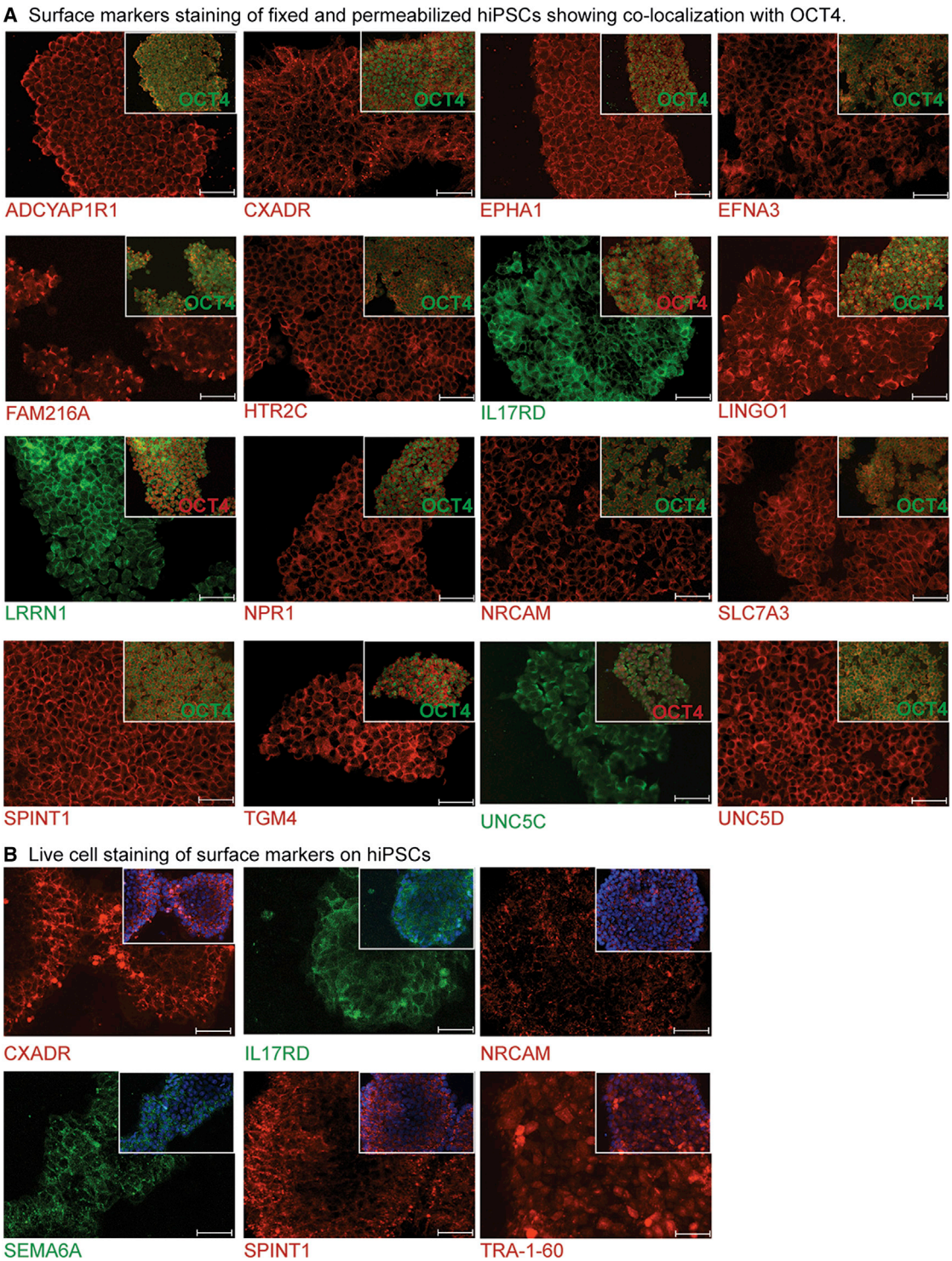


Figure 3. Immunocytochemistry of DF6-6-9T hiPS Cells Stained with Surface Markers

(A) Surface marker labeling on fixed hiPSCs with inset showing overlay with OCT4.

(B) Surface markers labeling on live hiPSCs with inset showing overlay with Hoechst. Scale bar, 50 μm.

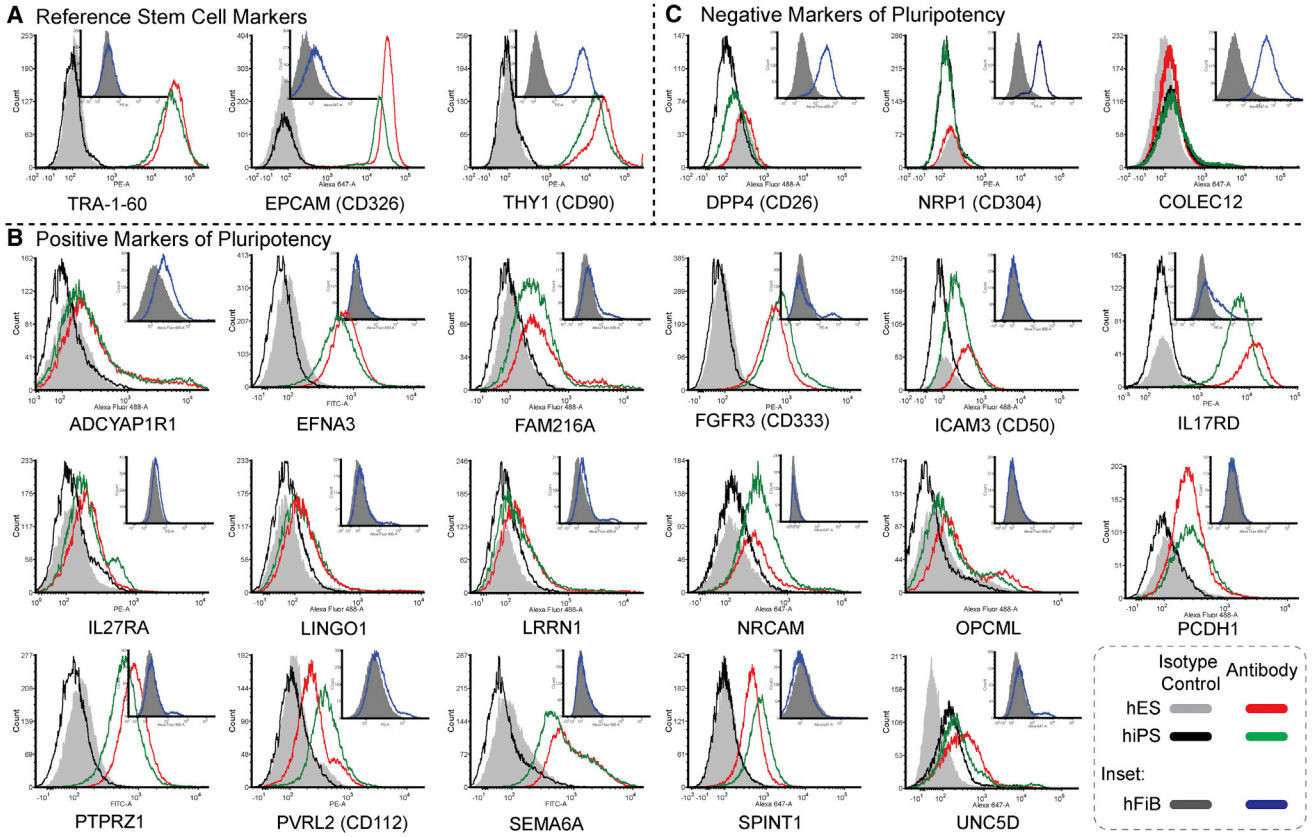


Figure 4. Flow Cytometry Histograms of Live, Unfixed, H1 hESC, DF6-6-9T hiPSC, and hFib

(A) Reference markers.

(B) Positive hPSC selection markers.

(C) Negative markers.

See also Figure S1.

48 hr and ~100% cell death by 72 hr (Figure 5E). Following 72 hr treatment, culture dishes of hiPSCs and hESCs were examined for up to 5 days post-treatment, and no viable cells were detected (data not shown). Successful elimination of hPSCs was observed regardless of cell density (30%–100% tested; data not shown). To further assess the selectivity of STF-31 for hPSCs, we established cocultured hESCs and hFibs (seeded 50:50; maintained on Matrigel for 4 days prior to treatment). Following 72 hr of STF-31 (2.5 μM) treatment, OCT4 expression was undetectable by flow cytometry and immunocytochemistry, whereas the fibroblast marker DPP4 was unaffected (Figure 5F, top and middle). Quantitative real-time PCR analyses of the treated coculture confirmed these findings, where the fibroblast markers *DPP4* and *NRP1* exhibit a relative increase (due to loss of hPSCs) and pluripotency markers (*OCT4* and *NANOG*) are reduced to background levels (Figure 5F, bottom). STF-31 could also be added to mixed cultures without killing hPSC-derived differentiated progeny. To demonstrate this, we generated monolayers of hESC-derived cardiomyocytes, which in our hands are ~98% TNNT2$^+$ by day 10 of differentiation (Figure 5G). After 72 hr of STF-31 (2.5 μM) treatment, we were unable to detect any effects on spontaneous contraction (rate or frequency) and by quantitative real-time PCR, cardiac transcripts *NKX2.5* decreased from day 10 to day 13, whereas *TNNT2* positivity remained constant. These results were indistinguishable from vehicle controls (Figure S2), thus indicating that STF-31, although toxic to hPSCs, spares a wide variety of progenitor and differentiated cells. Therefore, this resource also provides a "look-up directory" for the potential repurposing of drugs based on the direct knowledge of cell-surface protein abundance.

DISCUSSION

Until the recent development of proteomic technologies for the direct measurement of the pool of cell-surface-exposed proteins such as the CSC technology, researchers

had to rely on antibody-based probes for the sequential verification of cell-surface proteins. Microarray data partially enabled the inference of cell-surface expressed proteins, but discrepancies with flow cytometric data showed that posttranscriptional regulatory mechanisms need to be considered in order to obtain insights into the quantities of proteins in certain locations at a particular point in time. Thus, researchers lack a comprehensive and reliable resource of cell-surface exposed proteins for most human stem cells and progeny. Here, we directly measured the N-glycoprotein cell-surface proteome of hESCs and hiPSCs and provide an experimentally verified resource for the selection of lineage markers, markers for cell isolation, and the selection of reagents for the characterization of new hiPSC lines. Together with our analysis of hFibs, we have identified 719 surface proteins including 496 on hPSCs. Many of the identified proteins are barely detectable at the mRNA level depending on thresholds set forth or lack commercially available and validated antibodies that preclude their identification in antibody-based screens. Moreover, our analysis of the transcriptionally defined cell-surface proteome as predicted based on GO term annotations illustrate the limitations in relying solely on sometimes predicted annotations and transcriptional signals.

A genome-scale location analysis using published data (Gifford et al., 2013) was performed to identify surface proteins that have promoters occupied by POU5F1(*OCT4*)/*SOX2*/*NANOG* (OSN). In total, 30 (6%) of the 496 proteins identified on the surface of hPSCs have promoters occupied by OSN (Table S7). An additional 47 were occupied by two factors. Of the 206 hPSC-restricted proteins, 16 (7%) are occupied by OSN. Transcriptional dogma would suggest that promoters occupied by OSN are critical to pluripotency, and the surface data are consistent with this fact, because none of the putative negative markers were occupied by OSN and only five of the hFib restricted were occupied by OSN. Although a direct role in pluripotency has not yet been described for these 16 OSN occupied PSC-restricted proteins, several have been implicated in early developmental processes or mechanisms with relevance to pluripotency, including CXADR (Dorner et al., 2005), GPR98 (McMillan et al., 2002), IL17RD (Torii et al., 2004), LINGO1 (Mi et al., 2004), LRIG1 (Laederich et al., 2004), and LRRN1 (Hossain et al., 2008).

In the HSC field, immunophenotyping with antibodies has been critical to the evaluation of cell-surface proteins as surrogate markers of potency, function, immunological compatibility, and for the isolation of cells for therapeutic purposes. As we (Gundry et al., 2011, 2012) and others (Kahler et al., 2013; Kang et al., 2011; Manos et al., 2011) have demonstrated, accessible surface proteins are useful as markers for sorting and characterizing pluripotent cells. Application of the positive selection hPSC markers described in this resource provides the field with new targets suitable for improving the efficiency of isolation and characterization of high quality hiPSC colonies and/or those with defined differentiation potential. In addition to the positive PSC markers, these data reveal fibroblast markers (DPP4, NRP1, and COLEC12) useful for negative selection of hiPSCs from reprogrammed fibroblasts. Additional negative markers are likely to emerge as the cell types within the Cell Surface Protein Atlas expands or if less stringent data filtering are used (i.e., only require presence in hFib and absence in hPSCs, regardless of expression among other cell types). Precise mapping (time and quantity) of transcriptional and protein levels of these putative negative markers during reprogramming will ultimately influence the utility of these accessible proteins for selection purposes. The major bottleneck in using surface proteins for isolation and cell characterization is the lack of suitable antibodies or other affinity reagents that recognize native extracellular epitopes, a requirement for live cell selection. The information generated in this study partially overcome this limitation as the CSC technology experimentally verifies N-glycosylated peptides present in extracellular exposed domains of transmembrane and GPI-anchored proteins, which directly facilitates antigen design of surface epitopes useful for antibody production. Therefore, CSC technology lends itself for the immunophenotyping of cells without antibodies.

The hPSC surface proteome resource described here can be further exploited to rationally identify accessible and putative drug targets on the surface of hESCs and hiPSCs. This might be especially useful for repurposing known drugs. Although GLUT1 is present on many cell types, hPSCs, unlike many other cells that often contain multiple glucose transporters or rely on oxidative phosphorylation, are believed to be highly reliant on glycolysis, analogous to the Warburg effect in cancer cells (Folmes et al., 2011; Varum et al., 2011). Of the cells tested in this study, the reported GLUT1 inhibitor STF-31 proved to be selectively toxic to hPSCs. Recently, using a small molecule screening approach, Ben-David et al. (2013a) discovered the utility of a small molecule inhibitor of oleic acid biosynthesis, PluriSIn, with selective toxicity for hPSCs. Although it is known that metabolism influences a cell's decision to proliferate, differentiate, or remain quiescent (Shyh-Chang et al., 2013), these two studies firmly establish that small molecule-based metabolic inhibitors represent viable alternative strategies to suicide genes (Cao et al., 2007; Rong et al., 2012), SSEA-5, lactate, and Claudin-6-based approaches (Ben-David et al., 2013b; Tang et al., 2011; Tohyama et al., 2013). Furthermore, because another reported small molecule inhibitor of GLUT1 that we tested, WZB117, also proves toxic to hPSCs (data not shown), the use of

A

	# MS/MS Spectra		Microarray - Lab 1		Microarray - Lab 2	
	H9	KB3	H9	hiPSC2a	H9	KB3
SLC2A1 (GLUT1)	487	243	2544 ± 615	1586 ± 151	2953 ± 1434	1610 ± 411
SLC2A3 (GLUT3)	41	11	3423 ± 669	3151 ± 441	11196 ± 3431	7730 ± 1280
SLC2A4 (GLUT4)	32	7	6 ± 2	6 ± 2	75 ± 8	89 ± 14

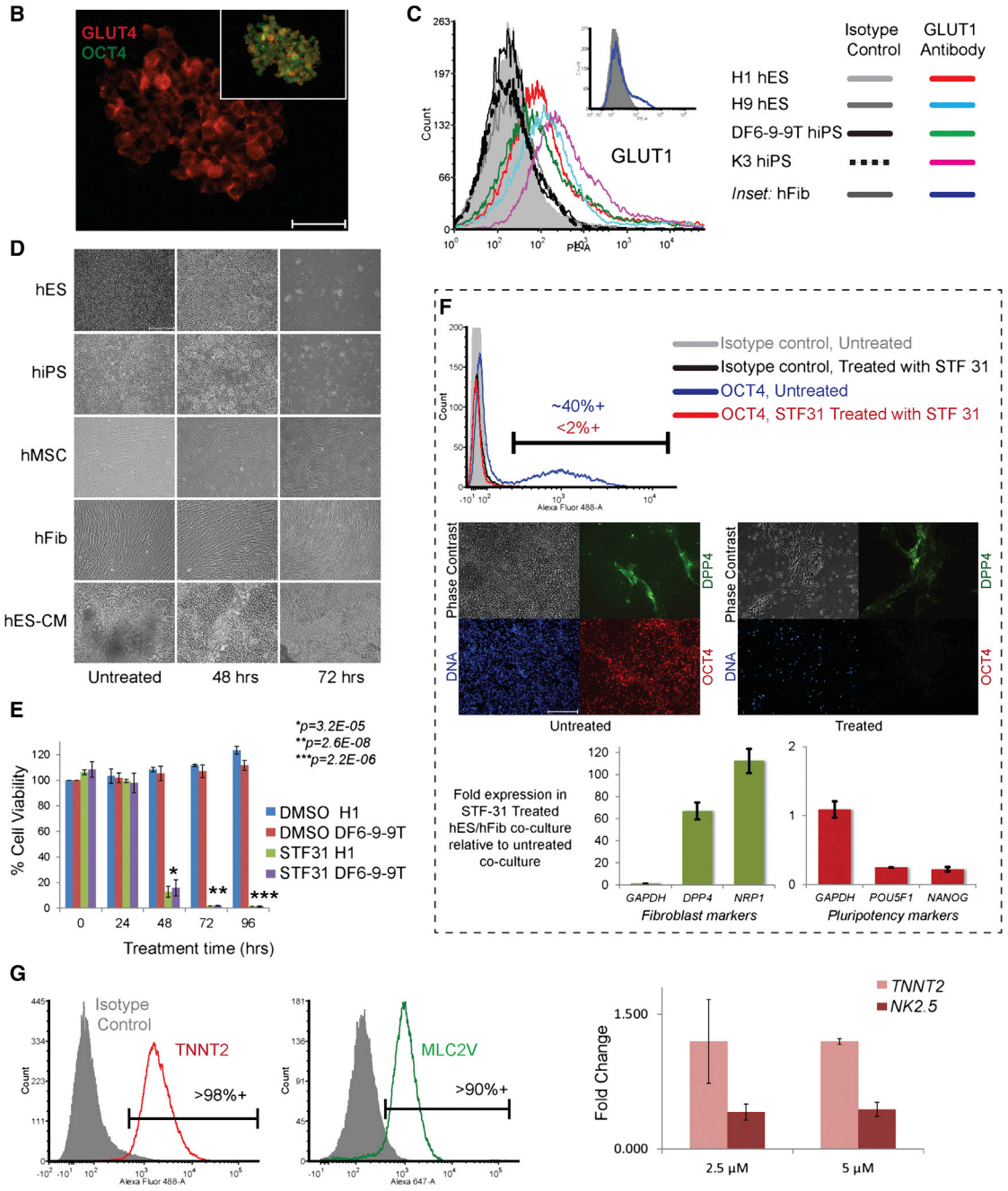


Figure 5. STF-31 Studies

(A) Table of the number of MS/MS spectra identified via the CSC technology and the microarray values for three GLUT transporters in H9 hESCs and KB3 hiPSCs.

(B) Immunofluorescence image of GLUT4 on DF6-9-9T hiPS showing surface localization, and inset shows overlay with OCT4. Scale bar, 50 µm.

(legend continued on next page)

other strategies to specifically inhibit GLUT1 may be broadly applicable for the removal of tumorigenic hPSCs from cultures of differentiating cells. In contrast to the previous reports that required alternative energy sources (Tohyama et al., 2013; Tomizawa et al., 2013), our study achieves selective hPSC elimination without alternative fuel substrates or general glucose starvation, making our findings more broadly useful to the community than previous approaches.

The comprehensive and dynamic single cell hPSC surface proteome is a work in progress. The analyses used here to identify hPSC-restricted proteins are based on information currently available in our Cell Surface Protein Atlas (D.B.-F., A. Hofmann, T. Bock, A. Frei, F. Cerciello, A. Jacobs, H. Moest, U. Omasits, R.L.G., C. Yoon, R. Schiess, A. Schmidt, P. Mirkowska, A. Härtlová, J. van Eyk, J. Bourquin, R. Aebersold, K.R.B., P. Zandstra, and B.W., unpublished data). Because this database expands and includes cell types from more developmental stages and lineages, classifications for some proteins may change, resulting in some of the hPSC-restricted proteins identified here being reclassified as nonrestricted, as well as new hPSC-restricted (positive and negative) markers coming to light. As we have learned from comparisons made possible by the Cell Surface Protein Atlas as well as the work from the hematopoietic cell field, we recognize that individual surface proteins are generally not restricted to a single cell type, and that is the reason we also considered proteins that were "relatively" restricted to hPSCs (in fewer than three other cell types) in addition to the truly restricted as a way to broaden the scope and bring more potential markers to light. Thus, this resource serves as a step toward defining surface protein and antibody panels more selective for authentic pluripotent cells. Second, as with any discovery-based mass spectrometry analysis of complex samples, the failure to identify a protein in a particular sample is not conclusive evidence that the protein is absent, due to the limitations in peptide sampling in the instrument. This is a well-recognized limitation in mass-spectrometry-based discovery approaches (Gingras et al., 2005) and can be exacerbated in the case of low abundance proteins (e.g., cell-surface proteins). Thus, the few putative differences between hESCs and hiPSCs based only on CSC technology data are not emphasized here but are included in Table S1 for completeness. Third, proteomic technologies are valuable for describing the average proteome present at a selected point in time or more simply a snapshot of proteins present on the cell surface at time of labeling. However, the presence of a protein or change in abundance of a protein with perturbation does not inherently reflect the biological relevance of the protein for a particular cell type. Nevertheless, it is predicted that proteins such as IL17RD, LINGO1, and LRRN1, which are among the most restricted to hPSCs, present in highest levels, and have promoters occupied by OSN, will be included as high-priority candidates going forward. Of note for functional relevance are LINGO1 and IL17RD as they are linked with inhibitory functions that may contribute to the maintenance of the undifferentiated state of PSCs. Moreover, although this data set is restricted to N-linked glycoproteins, it is predicted that ~90% of all cell-surface proteins are predicted to be glycosylated (Apweiler et al., 1999), and thus these data should represent a majority of the surface proteins present.

In conclusion, application of the CSC technology provides direct evidence of the pool of N-glycoproteins detectable at the hPSC surface. The bioinformatics strategy chosen based on the integration of disparate data sets provides a good filtering mechanism for the identification of sensible candidates for practical (i.e., immunophenotyping) and functional studies, not only for the hPSC populations studied here, but for a broad range of other stem

(C) Flow cytometry histograms of live cells illustrating surface localization of GLUT1 on the surface of hESCs and hiPSCs.
(D) Phase contrast images of hESCs, hiPSCs, hMSCs, hFibs, and hES-CMs after treatment with 2.5 μM STF-31, showing significant hPSC death by 48 hr and ~100% PSC death by 72 hr. Scale bar, 100 μm.
(E) Cell viability via neutral red assay of H1 hESCs and DF6-9-9T hiPSCs treated with 2.5 μM STF-31 compared to vehicle control. Data are an average of three biological replicates, and error bars represent SEM. For 48, 72, and 96 hr treatment, all p values are below 0.05, and the p values for the H1 data are shown.
(F) hESC/hFib coculture treated with STF-31 shows selective loss of hPSCs. Top: flow cytometry histograms showing OCT4 staining of untreated hESC/hFib coculture (blue) and absence of OCT4 staining in hESC/hFib coculture treated for 72 hr with STF-31 (red). Middle: bright-field and immunofluorescence images of hESC/hFib coculture untreated (left) or treated (right) with STF-31 for 72 hr, showing fibroblast marker DPP4 and pluripotency marker OCT4. Scale bar, 100 μm. Bottom: quantitative real-time PCR of *OCT4, NANOG, DPP4, NRP1*, and *GAPDH* in hESC/hFib coculture treated with STF-31 for 72 hr compared to untreated. Results from triplicate technical analyses of two biological replicates are shown, and error bars represent SEM.
(G) Left: flow cytometry histograms of day 11 hESC-CMs for TNNT2 and MLC2V, illustrating high-quality hESC-CMs obtained by in vitro differentiation of hESCs. Right: quantitative real-time PCR of cardiac genes *NKX2.5* and *TNNT2* in hESC-CMs treated for 72 hr (starting on day 10 of differentiation) with STF-31, indicating no adverse effects on the expression of these genes by the inhibitor as compared to untreated (Figure S2). Results from triplicate technical analyses of two biological replicates are shown and are normalized to untreated hESC. See also Figure S2.

cell types. Furthermore, we show that small molecules and/or inhibitors selected based on our resource could be further exploited for the depletion of cell populations in a heterogeneous cell context, such as shown here for the depletion of pluripotent cells, without the expense of a large-scale screen or gene targeting approaches. This protein-centric resource provides experimental evidence for location-specific proteins which, when further developed in the form of surface protein "barcodes" (Gundry et al., 2011), might help to delineate cellular lineages even further based on quantitative panels of cell-surface markers. A detailed understanding of the functional roles of the detected cell-surface markers in developmental and environmental interactions is certainly a prerequisite to understanding hiPSCs and ensuring their reproducible production for therapeutic purposes. Going forward, this technology and the resource described here will serve as a platform for future comparisons of stem cells of other origins. We acknowledge that extensive work will be necessary to refine the protein candidates described in this resource to arrive at the most effective and informative markers for isolating the highest quality pluripotent cells. This resource, which contains >100 proteins not previously described on the surface of hPSCs, should therefore provide a major step in a series of studies designed to define the optimum marker combinations for high throughput isolation and quality control studies of hPSCs.

EXPERIMENTAL PROCEDURES

Cell Culture

Generation of hiPSCs, cultivation of hiPSCs (KB3, DF6-9-9T [WiCell], hiPSC2a [Si-Tayeb et al., 2010]), hESC lines (H1, H9 [WiCell]), hFibs, bone marrow-derived mesenchymal stem cells (hMSCs), and differentiation of cardiomyocytes are described in the Supplement Information. The CSC technology was applied to H9 hESCs and KB3 hiPSCs. The flow cytometry and immunofluorescence staining were performed on DF-6-9T hiPSCs and H1 hESCs.

Cell-Surface Capture: CSC Technology

Approximately 0.6-1E8 cells per biological replicate ($n \geq 3$) of H9 hESCs, KB3 hiPSCs, and hFibs were taken through the CSC technology workflow as reported previously (Gundry et al., 2009, 2012; Hofmann et al., 2010; Wollscheid et al., 2009) with details provided in the Supplemental Information.

STF-31 Studies

Cells were treated with 0, 2.5, and 5 μM STF-31 (4-[[[[4-(1,1-Dimethylethyl)phenyl]sulfonyl]amino]methyl]-*N*-3-pyridinylbenzamide, Tocris Bioscience) for 96 hr to test cell viability with treatment. RNA was collected for quantitative real-time PCR, or cells were stained for DPP4 or OCT4. Day 10 hESC-CMs were treated for 48 hr with 0, 2.5, and 5 μM STF-31 treatment, and RNA was collected at 72 hr (day 13 of differentiation) for quantitative real-time PCR analysis of *TNNT2* and *NKX2.5*. In vitro cell toxicity was determined using a neutral red uptake assay as previously described (Steer et al., 2006).

Transcriptome Data

Microarray data were obtained in two laboratories on H9, HiPSC2a, and KB3 cells, as described in the Supplemental Information. To obtain a single expression value for visualization in Figure 2, values for a single probe were averaged among all H9 and KB3 replicates analyzed using the Illumina system; if multiple probes per gene were measured, only the highest value was included.

Quantitative Real-Time PCR, Flow Cytometry, and Immunocytochemistry

Details are provided in the Supplemental Information and Table S8.

ACCESSION NUMBERS

The Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) accession number for the microarray data reported in this paper is GSE55805.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, two figures, and eight tables and can be found with this article online at http://dx.doi.org/10.1016/j.stemcr.2014.05.002.

ACKNOWLEDGMENTS

This research was supported by NIH 4R00HL094708-03, BD Biosciences Research Grant Award, MCW Research Affairs Committee New Faculty Award, and the Kern foundation (startup funds) at the Medical College of Wisconsin (R.L.G.); the Intramural Research Program of the NIH/NIA, NIH Induced Pluripotent Stem Cell Center/Center for Regenerative Medicine Research Study Award and Research Grants Council of Hong Kong Theme-based Research Scheme T13-706/11 (K.R.B.); the Swiss National Science Foundation (grants 31003A_135805 to B.W.); Institutional Research Grant #86-004 from the American Cancer Society and the Midwest Athletes against Childhood Cancer (S.R.); U01 HL099776 and AHA Established Investigator Award (J.C.W.); AHA Postdoctoral Fellowship 12POST12050254 (P.W.B.); E.M.K. is a member of the MCW-MSTP, which is partially supported by a T32 grant from NIGMS, GM080202. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. We thank Hope Campbell at the Flow Cytometry Core of the Blood Research Institute of Wisconsin and Dr. Kate Noon, Michael Pereckas, and Xioagang Wu at the MCW Mass Spectrometry Facility for assistance with data collection. Special thanks to Dr. John Corbett (MCW) for generously providing access to the confocal microscope and the Biotechnology & Bioengineering Center (MCW) for access to the Real-Time PCR System.

Received: December 31, 2013
Revised: April 30, 2014
Accepted: May 5, 2014
Published: June 5, 2014

REFERENCES

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

Ben-David, U., Gan, Q.F., Golan-Lev, T., Arora, P., Yanuka, O., Oren, Y.S., Leikin-Frenkel, A., Graf, M., Garippa, R., Boehringer, M., et al. (2013a). Selective elimination of human pluripotent stem cells by an oleate synthesis inhibitor discovered in a high-throughput screen. Cell Stem Cell *12*, 167–179.

Ben-David, U., Nudel, N., and Benvenisty, N. (2013b). Immunologic and chemical targeting of the tight-junction protein Claudin-6 eliminates tumorigenic human pluripotent stem cells. Nat. Commun. *4*, 1992.

Berg, N.N., Puente, L.G., Dawicki, W., and Ostergaard, H.L. (1998). Sustained TCR signaling is required for mitogen-activated protein kinase activation and degranulation by cytotoxic T lymphocytes. J. Immunol. *161*, 2919–2924.

Bock, C., Kiskinis, E., Verstappen, G., Gu, H., Boulting, G., Smith, Z.D., Ziller, M., Croft, G.F., Amoroso, M.W., Oakley, D.H., et al. (2011). Reference Maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. Cell *144*, 439–452.

Buta, C., David, R., Dressel, R., Emgård, M., Fuchs, C., Gross, U., Healy, L., Hescheler, J., Kolar, R., Martin, U., et al. (2013). Reconsidering pluripotency tests: do we still need teratoma assays? Stem Cell Res. (Amst.) *11*, 552–562.

Cao, F., Drukker, M., Lin, S., Sheikh, A.Y., Xie, X., Li, Z., Connolly, A.J., Weissman, I.L., and Wu, J.C. (2007). Molecular imaging of embryonic stem cell misbehavior and suicide gene ablation. Cloning Stem Cells *9*, 107–117.

Chan, D.A., Sutphin, P.D., Nguyen, P., Turcotte, S., Lai, E.W., Banh, A., Reynolds, G.E., Chi, J.T., Wu, J., Solow-Cordero, D.E., et al. (2011). Targeting GLUT1 and the Warburg effect in renal cell carcinoma by chemical synthetic lethality. Sci. Transl. Med. *3*, 94ra70.

Cheng, F., Liu, C., Jiang, J., Lu, W., Li, W., Liu, G., Zhou, W., Huang, J., and Tang, Y. (2012). Prediction of drug-target interactions and drug repositioning via network-based inference. PLoS Comput. Biol. *8*, e1002503.

Chi, M.M., Pingsterhaus, J., Carayannopoulos, M., and Moley, K.H. (2000). Decreased glucose transporter expression triggers BAX-dependent apoptosis in the murine blastocyst. J. Biol. Chem. *275*, 40252–40257.

Damjanov, I., Fox, N., Knowles, B.B., Solter, D., Lange, P.H., and Fraley, E.E. (1982). Immunohistochemical localization of murine stage-specific embryonic antigens in human testicular germ cell tumors. Am. J. Pathol. *108*, 225–230.

Dorner, A.A., Wegmann, F., Butz, S., Wolburg-Buchholz, K., Wolburg, H., Mack, A., Nasdala, I., August, B., Westermann, J., Rathjen, F.G., and Vestweber, D. (2005). Coxsackievirus-adenovirus receptor (CAR) is essential for early embryonic cardiac development. J. Cell Sci. *118*, 3509–3521.

Feng, Q., Lu, S.J., Klimanskaya, I., Gomes, I., Kim, D., Chung, Y., Honig, G.R., Kim, K.S., and Lanza, R. (2010). Hemangioblastic derivatives from human induced pluripotent stem cells exhibit limited expansion and early senescence. Stem Cells *28*, 704–712.

Folmes, C.D., Nelson, T.J., Martinez-Fernandez, A., Arrell, D.K., Lindor, J.Z., Dzeja, P.P., Ikeda, Y., Perez-Terzic, C., and Terzic, A. (2011). Somatic oxidative bioenergetics transitions into pluripotency-dependent glycolysis to facilitate nuclear reprogramming. Cell Metab. *14*, 264–271.

Funk, W.D., Labat, I., Sampathkumar, J., Gourraud, P.A., Oksenberg, J.R., Rosler, E., Steiger, D., Sheibani, N., Caillier, S., Stache-Crain, B., et al. (2012). Evaluating the genomic and sequence integrity of human ES cell lines; comparison to normal genomes. Stem Cell Res. (Amst.) *8*, 154–164.

Gifford, C.A., Ziller, M.J., Gu, H., Trapnell, C., Donaghey, J., Tsankov, A., Shalek, A.K., Kelley, D.R., Shishkin, A.A., Issner, R., et al. (2013). Transcriptional and epigenetic dynamics during specification of human embryonic stem cells. Cell *153*, 1149–1163.

Gingras, A.C., Aebersold, R., and Raught, B. (2005). Advances in protein complex analysis using mass spectrometry. J. Physiol. *563*, 11–21.

Gore, A., Li, Z., Fung, H.L., Young, J.E., Agarwal, S., Antosiewicz-Bourget, J., Canto, I., Giorgetti, A., Israel, M.A., Kiskinis, E., et al. (2011). Somatic coding mutations in human induced pluripotent stem cells. Nature *471*, 63–67.

Gundry, R.L., Raginski, K., Tarasova, Y., Tchernyshyov, I., Bausch-Fluck, D., Elliott, S.T., Boheler, K.R., Van Eyk, J.E., and Wollscheid, B. (2009). The mouse C2C12 myoblast cell surface N-linked glycoproteome: identification, glycosite occupancy, and membrane orientation. Mol. Cell. Proteomics *8*, 2555–2569.

Gundry, R.L., Burridge, P.W., and Boheler, K.R. (2011). Pluripotent stem cell heterogeneity and the evolving role of proteomic technologies in stem cell biology proteomics. Proteomics *11*, 3947–3961.

Gundry, R.L., Riordon, D.R., Tarasova, Y., Chuppa, S., Bhattacharya, S., Juhasz, O., Wiedemeier, O., Milanovich, S., Noto, F.K., Tchernyshyov, I., et al. (2012). A cell surfaceome map for immunophenotyping and sorting pluripotent stem cells. Mol. Cell. Proteomics *11*, 303–316.

Hofmann, A., Gerrits, B., Schmidt, A., Bock, T., Bausch-Fluck, D., Aebersold, R., and Wollscheid, B. (2010). Proteomic cell surface phenotyping of differentiating acute myeloid leukemia cells. Blood *116*, e26–e34.

Hossain, M.S., Ozaki, T., Wang, H., Nakagawa, A., Takenobu, H., Ohira, M., Kamijo, T., and Nakagawara, A. (2008). N-MYC promotes cell proliferation through a direct transactivation of neuronal leucine-rich repeat protein-1 (NLRR1) gene in neuroblastoma. Oncogene *27*, 6075–6082.

Kahler, D.J., Ahmad, F.S., Ritz, A., Hua, H., Moroziewicz, D.N., Sproul, A.A., Dusenberry, C.R., Shang, L., Paull, D., Zimmer, M., et al. (2013). Improved methods for reprogramming human

dermal fibroblasts using fluorescence activated cell sorting. PLoS ONE *8*, e59867.

Kang, N.Y., Yun, S.W., Ha, H.H., Park, S.J., and Chang, Y.T. (2011). Embryonic and induced pluripotent stem cell staining and sorting with the live-cell fluorescence imaging probe CDy1. Nat. Protoc. *6*, 1044–1052.

Kannagi, R., Levery, S.B., Ishigami, F., Hakomori, S., Shevinsky, L.H., Knowles, B.B., and Solter, D. (1983). 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. *258*, 8934–8942.

Kim, J., Woo, A.J., Chu, J., Snow, J.W., Fujiwara, Y., Kim, C.G., Cantor, A.B., and Orkin, S.H. (2010). A Myc network accounts for similarities between embryonic stem and cancer cell transcription programs. Cell *143*, 313–324.

Kristensen, A.R., Gsponer, J., and Foster, L.J. (2013). Protein synthesis rate is the predominant regulator of protein expression during differentiation. Mol. Syst. Biol. *9*, 689.

Laederich, M.B., Funes-Duran, M., Yen, L., Ingalla, E., Wu, X., Carraway, K.L., 3rd, and Sweeney, C. (2004). The leucine-rich repeat protein LRIG1 is a negative regulator of ErbB family receptor tyrosine kinases. J. Biol. Chem. *279*, 47050–47056.

Lian, X., Zhang, J., Azarin, S.M., Zhu, K., Hazeltine, L.B., Bao, X., Hsiao, C., Kamp, T.J., and Palecek, S.P. (2013). Directed cardiomyocyte differentiation from human pluripotent stem cells by modulating Wnt/β-catenin signaling under fully defined conditions. Nat. Protoc. *8*, 162–175.

Liu, Z., and Armant, D.R. (2004). Lysophosphatidic acid regulates murine blastocyst development by transactivation of receptors for heparin-binding EGF-like growth factor. Exp. Cell Res. *296*, 317–326.

Lowry, W.E., Richter, L., Yachechko, R., Pyle, A.D., Tchieu, J., Sridharan, R., Clark, A.T., and Plath, K. (2008). Generation of human induced pluripotent stem cells from dermal fibroblasts. Proc. Natl. Acad. Sci. USA *105*, 2883–2888.

Maherali, N., and Hochedlinger, K. (2008). Guidelines and techniques for the generation of induced pluripotent stem cells. Cell Stem Cell *3*, 595–605.

Manos, P.D., Ratanasirintrawoot, S., Loewer, S., Daley, G.Q., and Schlaeger, T.M. (2011). Live-cell immunofluorescence staining of human pluripotent stem cells. Curr. Protoc. Stem Cell Biol. *Chapter 1*, Unit 1C 12.

McMillan, D.R., Kayes-Wandover, K.M., Richardson, J.A., and White, P.C. (2002). Very large G protein-coupled receptor-1, the largest known cell surface protein, is highly expressed in the developing central nervous system. J. Biol. Chem. *277*, 785–792.

Mi, S., Lee, X., Shao, Z., Thill, G., Ji, B., Relton, J., Levesque, M., Allaire, N., Perrin, S., Sands, B., et al. (2004). LINGO-1 is a component of the Nogo-66 receptor/p75 signaling complex. Nat. Neurosci. *7*, 221–228.

Mobasheri, A., Dobson, H., Mason, S.L., Cullingham, F., Shakibaei, M., Moley, J.F., and Moley, K.H. (2005). Expression of the GLUT1 and GLUT9 facilitative glucose transporters in embryonic chondroblasts and mature chondrocytes in ovine articular cartilage. Cell Biol. Int. *29*, 249–260.

Müller, F.J., Goldmann, J., Löser, P., and Loring, J.F. (2010). A call to standardize teratoma assays used to define human pluripotent cell lines. Cell Stem Cell *6*, 412–414.

Müller, F.J., Schuldt, B.M., Williams, R., Mason, D., Altun, G., Papapetrou, E.P., Danner, S., Goldmann, J.E., Herbst, A., Schmidt, N.O., et al. (2011). A bioinformatic assay for pluripotency in human cells. Nat. Methods *8*, 315–317.

Murry, C.E., and Keller, G. (2008). Differentiation of embryonic stem cells to clinically relevant populations: lessons from embryonic development. Cell *132*, 661–680.

Omasits, U., Ahrens, C.H., Muller, S., and Wollscheid, B. (2013). Protter: interactive protein feature visualization and integration with experimental proteomic data. Bioinformatics.

Ratajczak, M.Z., Shin, D.M., Schneider, G., Ratajczak, J., and Kucia, M. (2013). Parental imprinting regulates insulin-like growth factor signaling: a Rosetta Stone for understanding the biology of pluripotent stem cells, aging and cancerogenesis. Leukemia *27*, 773–779.

Riggs, J.W., Barrilleaux, B.L., Varlakhanova, N., Bush, K.M., Chan, V., and Knoepfler, P.S. (2013). Induced pluripotency and oncogenic transformation are related processes. Stem Cells Dev. *22*, 37–50.

Robinton, D.A., and Daley, G.Q. (2012). The promise of induced pluripotent stem cells in research and therapy. Nature *481*, 295–305.

Rong, Z., Fu, X., Wang, M., and Xu, Y. (2012). A scalable approach to prevent teratoma formation of human embryonic stem cells. J. Biol. Chem. *287*, 32338–32345.

Shyh-Chang, N., Daley, G.Q., and Cantley, L.C. (2013). Stem cell metabolism in tissue development and aging. Development *140*, 2535–2547.

Si-Tayeb, K., Noto, F.K., Sepac, A., Sedlic, F., Bosnjak, Z.J., Lough, J.W., and Duncan, S.A. (2010). Generation of human induced pluripotent stem cells by simple transient transfection of plasmid DNA encoding reprogramming factors. BMC Dev. Biol. *10*, 81.

Steer, S.A., Scarim, A.L., Chambers, K.T., and Corbett, J.A. (2006). Interleukin-1 stimulates beta-cell necrosis and release of the immunological adjuvant HMGB1. PLoS Med. *3*, e17.

Tang, C., Lee, A.S., Volkmer, J.P., Sahoo, D., Nag, D., Mosley, A.R., Inlay, M.A., Ardehali, R., Chavez, S.L., Pera, R.R., et al. (2011). An antibody against SSEA-5 glycan on human pluripotent stem cells enables removal of teratoma-forming cells. Nat. Biotechnol. *29*, 829–834.

Taylor, T., Kim, Y.J., Ou, X., Derbigny, W., and Broxmeyer, H.E. (2010). Toll-like receptor 2 mediates proliferation, survival, NF-kappaB translocation, and cytokine mRNA expression in LIF-maintained mouse embryonic stem cells. Stem Cells Dev. *19*, 1333–1341.

Tohyama, S., Hattori, F., Sano, M., Hishiki, T., Nagahata, Y., Matsuura, T., Hashimoto, H., Suzuki, T., Yamashita, H., Satoh, Y., et al. (2013). Distinct metabolic flow enables large-scale purification of mouse and human pluripotent stem cell-derived cardiomyocytes. Cell Stem Cell *12*, 127–137.

Tomizawa, M., Shinozaki, F., Sugiyama, T., Yamamoto, S., Sueishi, M., and Yoshida, T. (2013). Survival of primary human hepatocytes and death of induced pluripotent stem cells in media lacking glucose and arginine. PLoS ONE *8*, e71897.

Torii, S., Kusakabe, M., Yamamoto, T., Maekawa, M., and Nishida, E. (2004). Sef is a spatial regulator for Ras/MAP kinase signaling. Dev. Cell *7*, 33–44.

Varum, S., Rodrigues, A.S., Moura, M.B., Momcilovic, O., Easley, C.A., 4th, Ramalho-Santos, J., Van Houten, B., and Schatten, G. (2011). Energy metabolism in human pluripotent stem cells and their differentiated counterparts. PLoS ONE *6*, e20914.

Vogel, C., and Marcotte, E.M. (2012). Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. Nat. Rev. Genet. *13*, 227–232.

Wollscheid, B., Bausch-Fluck, D., Henderson, C., O'Brien, R., Bibel, M., Schiess, R., Aebersold, R., and Watts, J.D. (2009). Mass-spectrometric identification and relative quantification of N-linked cell surface glycoproteins. Nat. Biotechnol. *27*, 378–386.

Yan, Y., Yang, D., Zarnowska, E.D., Du, Z., Werbel, B., Valliere, C., Pearce, R.A., Thomson, J.A., and Zhang, S.C. (2005). Directed differentiation of dopaminergic neuronal subtypes from human embryonic stem cells. Stem Cells *23*, 781–790.

Yildirim, M.A., Goh, K.I., Cusick, M.E., Barabási, A.L., and Vidal, M. (2007). Drug-target network. Nat. Biotechnol. *25*, 1119–1126.