Published in final edited form as:

Proteomics. 2011 October ; 11(20): 3947–3961. doi:10.1002/pmic.201100100.

Pluripotent stem cell heterogeneity and the evolving role of proteomic technologies in stem cell biology

Rebekah L. Gundry^{1,2}, Paul W. Burrridge³, and Kenneth R. Boheler⁴

¹Department of Biochemistry, Medical College of Wisconsin, Milwaukee, WI, USA

²Department of Biotechnology and Bioengineering Center, Medical College of Wisconsin, Milwaukee, WI, USA

³Institute for Cell Engineering, The Johns Hopkins University School of Medicine, Baltimore, MD, USA

⁴Molecular Cardiology and Stem Cell Unit, National Institute on Aging, National Institutes of Health, Baltimore, MD, USA

Abstract

Stem cells represent obvious choices for regenerative medicine and are invaluable for studies of human development and drug testing. The proteomic landscape of pluripotent stem cells (PSCs), in particular, is not yet clearly defined; consequently, this field of research would greatly benefit from concerted efforts designed to better characterize these cells. In this concise review, we provide an overview of stem cell potency, highlight the types and practical implications of heterogeneity in PSCs and provide a detailed analysis of the current view of the pluripotent proteome in a unique resource for this rapidly evolving field. Our goal in this review is to provide specific insights into the current status of the known proteome of both mouse and human PSCs. This has been accomplished by integrating published data into a unified PSC proteome to facilitate the identification of proteins, which may be informative for the stem cell state as well as to reveal areas where our current view is limited. These analyses provide insight into the challenges faced in the proteomic analysis of PSCs and reveal one area – the cell surface subproteome – that would especially benefit from enhanced research efforts.

Keywords

Cell biology; Cell surface; Heterogeneity; Pluripotent stem cell

1 Introduction

Stem cells are specialized cells in complex organisms that retain the ability to self-renew (i.e. undergo cell division in an undifferentiated state) indefinitely or to differentiate into one or many types of specialized cell types. Because of these traits, stem cells are useful for studies of mammalian development, drug testing, toxicology, and for cell replacement therapies and regenerative medicine [1]. However, several challenges currently impede the use and study of most stem cell subpopulations. First, the inherent heterogeneity that exists

within populations and among lines complicates our ability to phenotypically characterize and isolate these cells as well as to distinguish them from differentiating progeny. Second, despite the rapid progress in stem cell biology over the past few years, researchers currently rely heavily on indirect methodologies to establish both stem cell identity and function in vitro and in vivo. These challenges are relevant for both basic and translational research applications.

The overall goal for this review is to discuss these challenges in a context apt for the proteomics community. Specifically, we provide an overview of stem cell potency, highlight the types and practical implications of heterogeneity in pluripotent stem cells (PSCs), and provide a unique, detailed analysis of the current view of the pluripotent proteome. Considering these results in the context of PSC heterogeneity, the importance of continued efforts for characterizing the PSC cell surface subproteome emerges.

2 Stem cell potency

Three broad categories of naturally occurring stem cells have been classified: embryonic (blastocyst, epiblast derived or germ cell derived), adult (tissue-specific or cord blood), and cancer stem cells [2]. The differentiation potential or potency of these cell types serves functionally to define the type of stem cell. Pluripotent embryonic stem cells (ESCs) display the broadest developmental potential [1–3] and can differentiate into all cells of a developing embryo (Fig. 1A). Perhaps most importantly, mouse ESCs (mESCs) can be used to generate animal chimeras, in which the ESC genotype can be passed through the germline. An entire stem cell-derived embryo, excluding some extra-embryonic tissues, can also be formed from some high-quality mESCs. For ethical reasons, human ESCs (hESCs) cannot be tested for these two abilities, but based on other criteria (differentiation potential in vitro and teratoma formation), these cells display the entire pluripotent phenotype typical of mESCs. Another pluripotent cell type, primordial germ cells, produce gametes in vivo and have multilineage differentiation capacity in vitro [1]; however, these cells have undergone varying degrees of demethylation and remethylation, which limit their use in chimera formation and germline transmission.

Adult stem cells possess a more restricted developmental potential and are generally considered to be unipotent to multipotent. These cells typically produce only progeny of a closely related family, providing “new” cells for replenishing specialized cells in the adult that have been damaged or lost. The archetype of restricted differentiation potential are hematopoietic stem cells (HSCs), which give rise to all of the myeloid and lymphoid cell types present in blood via progenitors [4–6]. As is typical for progenitor cells, myeloid and lymphoid progenitors have a limited self-renewal capacity, but a very high proliferation capacity. This latter property permits the production of a large number of differentiated progeny in a relatively short period of time suitable for replacing lost cells. Since human erythrocytes, for example, only survive about 100–120 days in the circulation, these cells must be constantly replaced. Other examples of adult stem cells with similar properties and hierarchies included neuronal, mesenchymal, epithelial, hepatic, intestinal, and pancreatic stem cells [6–11]. Despite the on-going process of cell replacement, the stem cell hierarchy that gives rise to specific progenitors and specialized cell types rarely ever form differentiated cell progeny in vivo outside the normal hierarchy.

Cancer stem cells are a minor population of tumor cells in a cancer that have stem cell-like properties. Cancer stem cells may arise through neoplastic changes initiated in normal self-renewing stem cells or in downstream progenitors, but are not necessarily derived from stem cells. Instead, the originating cell(s) has acquired properties of self-renewal [12–14]. Unlike normal stem cells, cancer stem cells are by definition oncogenic, have lost some regulatory

controls necessary to prevent uncontrolled proliferation or differentiation, and are therefore therapeutically unviable. Some of these cells are mutant versions of normal stem cells or progenitor cells as in certain forms of leukemia; whereas, others are mutant somatic cells that have regained the capacity to self-renew [12, 13, 15, 16]. Importantly, many cancer stem cells are defined based on the presence of normal stem cell markers. Therapeutically viable stem cells are thus normal units of tissue regeneration and development (for rev. see [1]); whereas, cancer stem cells, which show traits of normal stem cells, are abnormal and cannot be considered therapeutically viable. It is also important to note that ESCs can form tumors in vivo, but these cells are not considered oncogenic. This is because ESCs when placed in the appropriate environment (i.e. blastocyst) do not form tumors.

A fourth class of stem cells is experimentally derived from fetal or adult cells types through a process known as reprogramming. The first example of this came from the group of Shinya Yamanaka who showed that induced pluripotent stem cells (iPSCs) could be experimentally derived in vitro through transcription factor-mediated reprogramming of cells through expression of SOX2, OCT4 and in conjunction with KLF4 and c-MYC (SOKM) in mice [17]. Subsequently, simultaneous reports from Takahashi et al., [18], using SOKM, and Yu et al., [19], using SOX2, OCT4, NANOG, LIN28 (SONL), showed reprogramming of human cells. Improvements in efficiency were then demonstrated by Park et al., using SOKM+SV40T and hTERT [20] and then Lowry et al., [21] using SOKM +NANOG. Yu et al. [22] demonstrated that Sox2, Oct4, KLF4, c-MYC, NANOG, LIN28 and SV40T (SOKMNLTL) allowed for episomal (non-viral) reprogramming. Finally, others have shown that the addition of additives, such as valproic acid, or small molecule inhibitors such as PD0325901, SB431542 [23] and CHIR9902 [24] facilitate the reprogramming process such that fewer transcription factors are required.

iPSCs can be derived from almost any somatic tissue, including but not limited to embryonic and adult fibroblasts [25, 26], hepatocytes and gastric epithelial cells [27], pancreatic cells [28], neural stem cells [29, 30] and B lymphocytes [31] in mouse, and skin fibroblasts [32, 33], keratinocytes [34], adipose stem cells [35], and peripheral blood cells [36] in human. The reprogramming process from adult mouse cells to iPSCs generally takes at least 2wk but can take up to 2 months. During this time period, the overexpressed transcription factors cause a sequential and time-dependent activation of specific markers like alkaline phosphatase and surface-associated stage-specific antigens [37, 38]. The reprogramming process is “complete” when endogenous genes encoding *nanog* and *pou5f1* (*oct4*) are fully activated and have “pushed” the transcriptional hierarchy to a pluripotent compatible state. The ease of generating iPSCs in vitro has fostered the development of patient-derived cells. Unlike hESC-derived progeny, descendants of patient-derived iPSCs are immunologically compatible with the recipient. These cells and their offspring may therefore prove most apt for therapeutic interventions in humans with intractable disease states or syndromes.

Of all the stem cells currently available for possible therapeutic interventions, PSCs seem to have the greatest potential to provide cures for degenerative diseases through direct cell replacement, particularly in cases where limited numbers of tissue-specific stem cells are available or the cells are difficult to isolate. The generation of iPSCs derived from a patient's own cells means that the risk of immune rejection is substantially lowered. iPSCs thus may supplant hESCs as the primary source of pluripotent cells; however, the advantages of iPSCs are counterbalanced by unresolved questions involving safety and efficacy – areas of active investigation and rapid progress.

3 PSC heterogeneity

One aspect of PSC biology that is often underappreciated is the heterogeneity that exists between stem cell lines and even among cells from the same line. Theoretically, clonally derived cells should all be identical; however, dividing cells will by definition show transcriptional and protein differences based on their passage through the cell cycle. Additionally, genetic abnormalities have been described, and PSCs may undergo a spontaneous mutation at a rate of $10(-9)$ per nucleotide. Moreover, relative to early-passage hESC lines, late-passage lines have one or more genomic alterations, including aberrations in copy number (45%), mitochondrial DNA sequence (22%), and gene promoter methylation (90%) [39]. Problems associated with X-inactivation have also been reported in hESC lines [40], and other epigenetic defects are likely to exist. While some defects can be detected by genomic hybridization, even when karyotyping fails to detect them, the identification of spontaneous genetic and epigenetic abnormalities is a major challenge in the field [41]. In vitro, all PSCs also undergo some degree of spontaneous differentiation in vitro (Fig. 2), which ultimately results in cell commitment, a loss of potency, and the establishment of robust cell cycle checkpoints [2]. More surprisingly, however, cells within an ESC colony show remarkable cell-to-cell heterogeneity. For example, the transcription factor POU5F1 (OCT4) is highly abundant in all undifferentiated ESCs; however, other transcription factors like STELLA1, REX1, and NANOG are differentially expressed in these same cells. STELLA1 and NANOG are present in approximately 20–30 and 80% of ESCs, respectively [42]. Cells that are STELLA1⁺ and NANOG⁺ are more inner cell mass-like; whereas, those that are NANOG⁻, GATA6⁺ and HEX1⁺ are pre-disposed to generate primitive endoderm [43, 44]. Similarly, cells that have lost REX1 (FGF5⁺) are pre-disposed to generate primitive ectoderm [45]. These pre-differentiation states are however dynamic, and the lack of a signal to differentiate causes these cells to cycle between a more inner cell mass-like state and one with a predisposition to differentiation. A summary of currently understood transcription factor heterogeneity in PSCs is presented in Fig. 1B.

Pluripotent iPSCs are even more heterogeneous than ESCs [46]. This is due in part to the reprogramming process, which proceeds differentially among transfected cells and partially as a consequence of the cell of origin. iPSCs thus show significant reprogramming variability and remarkable interline variation. Differences among lines range from different quantities of essential transcription factors associated with pluripotency, unique differentiation inclinations, variable degrees of tumorigenicity, and in the case of mouse, differential chimerism and low germline transmission [47, 48]. iPSCs also show variability in somatic memory and aberrant reprogramming of DNA methylation, including differentially methylated regions proximal to centromeres and telomeres, and differences in cytosine- guanine methylation and histone modifications [49]. Several reports also indicate that hiPSC compared to hESC progeny senesce prematurely [50, 51] and have lower telomerase activity than hESCs, which may account for the impaired proliferative potential of some iPSC derivatives [52]. Moreover, it is clear from “Dolly” and nuclear transfer experiments that adult somatic cells retain cell autonomous defects that adversely affect the viability of reprogrammed cells [53]. Aging, therefore, is likely to negatively affect the quality of hiPSCs that are generated in vitro from “old” somatic cells.

Given the extent of heterogeneity within and among stem cell lines, the question is therefore how best to identify, isolate, and characterize these cells. From a basic science perspective, the introduction of sequences that confer antibiotic resistance (e.g. neomycin, puromycin, hygromycin, and herpes simplex virus thymidine kinase) for clonal selection or of reporter genes (e.g. green fluorescent protein (GFP/EGFP)) to identify specific cell lineages have proved most useful, and there are numerous studies that have exploited this system to answer fundamental questions related to stem cell biology [1]. These techniques are not

however viable for therapeutic applications of stem cells, because of possible side effects associated with DNA integration, including oncogenesis, or toxicity associated with the presence of a transgene. Thus, methods that can exploit naturally occurring, accessible moieties are highly desirable for the isolation of PSCs and/or their progeny for therapy. In an effort to understand how proteomic technologies can address these evolving needs, we examined published proteomic studies of PSCs and present a global overview of the current pluripotent proteome. Considering the current needs in the stem cell field, this analysis reveals that while proteomic technologies have the potential to identify novel targets for immunophenotyping PSCs, more efficient discovery of cell surface proteins combined with robust functional studies of populations identified by these surface markers will be required. We therefore discuss practical strategies for efficiently identifying cell surface markers and highlight several challenges faced in their practical implementation.

4 Current global view of the pluripotent proteome

Numerous ($n > 50$) proteomic-based studies of PSCs involving a broad range of strategies have been published. However, there is currently no consensus regarding the global implications of these results. For example, which proteins are required for the maintenance of pluripotency; which signaling pathways are important; how do these proteins interact to promote differentiation or maintain potency? To begin to address these questions, we have performed a broad-based comparison of the results from published proteomic studies of mouse and human PSCs. Of the >50 studies available, those reports that exclusively examined secreted proteins (e.g. conditioned media), did not provide a publicly available list of accession numbers for proteins explicitly identified in the undifferentiated state, and data from differentiation experiments within these reports were excluded. Data from the remaining 34 studies (Table 1A) were imported in ProteinCenter (Proxeon/ThermoFisher). All data were clustered by 85% sequence homology to remove protein redundancy and gene ontology terms and transmembrane predictions were obtained via ProteinCenter. The analysis included 34 publications (Table 1A) and resulted in a total of 7487 and 7295 proteins identified in mouse and human, respectively, of which 3088 were found in both species (Fig. 3A). The list of identified proteins for each species is contained in Supporting Information Table S1, and this represents the first large-scale compilation of published data on human and mouse PSCs presented in an easily-accessible format.

Unexpectedly, 63% of proteins were found only in ≤ 2 data sets and less than 10% were found in $\geq 50\%$ of studies for each species (Fig. 3B). Since this limitation is possibly due to the relatively small data sets for some of the publications, the analysis was refined to include only those data sets that reported >1000 proteins ($n = 7$ mouse and 9 human data sets). This resulted in a total of 6500 mouse and 6966 human proteins, where similar to the larger comparison, a majority (61% of mouse, 64% of human) of proteins were found in ≤ 2 data sets. When considering only the larger data sets, the number of proteins identified in $\geq 50\%$ of data sets increased (27% of mouse, 19% of human). The lack of extensive overlap among these studies may be a result of the functional heterogeneity within PSC cultures as well as the significant inter-line variation already known to exist, but likely also reflects variations due to culture conditions, sampling statistics typical of any proteomics experiment, and the variety of technical approaches employed (e.g. protein enrichment, protein/peptide separation, MS instrumentation). Of course, failure to identify a protein is not conclusive evidence for the absence of a protein in the sample. Rather, qualitative differences may result from the MS-level analysis (e.g. failure to select a peptide for fragmentation, failure to generate sufficient fragmentation for identification) or inherent properties of the protein (e.g. inability of tryptic digestion to generate peptide of suitable m/z for MS analysis) and these are exacerbated in the case of low-abundance proteins. For these reasons, it is likely that

some of the proteins found in only a few studies are actually more widely present than reported.

Among all 34 studies examined, 169 proteins were found in ten or more studies for each species (Fig. 3C). It is expected that while these most commonly identified proteins represent those that are highly abundant (e.g. actin, tubulin, ribosomal proteins, heat shock proteins), they may also include those proteins important for the maintenance of pluripotency. In human, LIN28A, which is an important pluripotency factor [19] and CTNNB1, which is involved in the Wnt pathway and is important in self-renewal [54], were found in ten studies. Additionally, TRIM28, which is important for PSC self-renewal [55] was found in 11 studies. While this broad-based comparison strategy has the potential to reveal proteins with biologically relevant roles in PSC biology, it is likely to also reveal those proteins that are most abundant in PSCs, regardless of biological significance.

In addition to the studies compiled into the “PSC proteome” discussed above (Table 1A and Supporting Information Table S1) several studies have used proteomic technologies to study early stages of PSC differentiation (Table 1B), and some have been previously reviewed [56–60]. Finally, there are a large number of additional studies that examine the proteomes of PSC-derivatives and/or secreted factors. These are excellent topics for future reviews, but are beyond the scope of the current discussion.

5 Targeted surface protein discovery for sorting out heterogeneity

Owing to the types of heterogeneity described earlier, it is currently difficult to obtain “homogeneous” populations of PSCs for proteomic analysis. Consequently, important nuances inherent to the control of self-renewal or commitment are possibly diluted in studies that average an entire culture. Thus, refined analyses on more homogeneous populations will be required to assess the potentially informative protein abundance changes and/or modifications that have been identified through the qualitative and quantitative proteomic analysis of heterogeneous populations so that they may be correlated with a specific functional relevance. The obvious question is how do we obtain more homogeneous populations of PSCs for future detailed analyses? The best example of how this can be achieved is from the HSC system, where cell surface accessible markers are used to identify and isolate functionally defined subsets of bone marrow-derived HSCs and their progeny [61, 62]. However, a similar “cell surface barcode” [63] does not exist for the identification and purification of bona fide PSCs. Rather, the analysis of PSCs rely on surface markers like SSEA-1 (mouse), SSEA-3 and SSEA-4 (human), TRA-1-60, and TRA-1-81, none of which are suitable for the isolation of homogeneous PSC subpopulations or lineage-specific progeny [64, 65]. While it will likely not be possible to identify surface markers specific to every cell type present, the ability to identify and isolate functionally defined subpopulations of PSCs without the use of genetic manipulations would be beneficial to both therapeutic and research efforts.

As in business, a barcode is an optical and readable representation of data that is unique to each product. By analogy, a “cell surface barcode” should be uniquely informative of a biological state for specific cell types. Although the need for a unique classification of cells based on biological markers is well appreciated, the cell surface remains less well characterized than other subproteomes and there are currently only a limited number of candidate surface markers for PSCs. This can be attributed to the challenges faced when analyzing cell surface proteins, including the difficulty in solubilizing transmembrane proteins and enriching for a lesser abundant subproteome. Thus, the application of proteomic technologies that are especially well suited for the study of cell surface proteins is poised to make a significant impact on our understanding of how stem cells communicate

with their microenvironment and on the development of new reagents for identifying and isolating PSCs and their progeny.

Creative approaches have been developed to circumvent the challenges in identifying cell surface proteins (selected reviews [66–70]). In general, two categories of strategies have emerged, which have shown an ability to identify cell surface proteins more efficiently than the analysis of the whole cellular proteome and are expected to be useful in identifying surface markers on PSCs. These include “physical enrichment” and “affinity enrichment” of plasma membrane-(PM) associated proteins (Fig. 4A and B). Briefly, in physical enrichment strategies, differential centrifugation is used to enrich for membrane proteins and several studies reviewed here have used variations of this approach on mouse and human PSCs [71–73]. In affinity enrichment strategies, several types of chemical tags can be used to selectively label accessible regions of surface proteins and these methods are typically combined with physical enrichment strategies. The most common affinity enrichment strategies include protein biotinylation via cleavable or non-cleavable biotin onto accessible lysine residues (reviewed by [74]), and oligosaccharide biotinylation via hydrazide chemistry (CSC-Technology), first described by Wollscheid et al. [63]. One advantage of using cleavable biotin (for protein biotinylation) or the CSC-Technology (oligosaccharide biotinylation) is the data “tag” resulting from the cleaved biotin or oligosaccharide that remains on the peptide and can be identified by MS. This facilitates the data analysis by allowing for efficient removal of contaminating, non-surface proteins that may have been enriched during centrifugation, as it is not possible to purify the PM from intracellular membranes by centrifugation alone. Moreover, this data tag also provides experimental evidence regarding the extracellular domain for transmembrane proteins. Such information is critical for the future development of affinity reagents that recognize the extracellular domain, especially in instances where membrane topology predictions are ambiguous. There are technical challenges to consider for any cell surface protein enrichment strategy, including the specificity of the label for protein subsets, the need to optimize labeling conditions for each cell type examined, the potential for contamination from intracellular membranes, and the often large number of cells required. However, as the methods become more streamlined and the detection limit in modern mass spectrometers improves, this will enable the analysis of smaller, more homogeneous populations. Finally, each cell surface enrichment method targets a different characteristic of the cell surface and, consequently, they reveal complimentary views of the surface proteome [75]. It is unlikely that a single strategy will be able to fully elucidate the cell surface proteome, though integrating several approaches is expected to greatly expand our view of this important subproteome.

Physical [71, 72, 76] and affinity [63, 77, 78] cell surface enrichment strategies have been applied to mouse and human PSCs, and altogether revealed 4415 proteins. Approaches that enrich for the cell surface report 1651 unique proteins and identified proteins with multiple transmembrane domains at a higher efficiency (Fig. 4C and D) when compared to non-PM targeting strategies. As shown in Fig. 4E, it is clear that the enrichment strategies are more efficient for selectively identifying cell surface proteins over intracellular proteins than non-targeted strategies; although several studies in the “non-PM targeted” group were designed to enrich for other cellular subproteomes and thus not expected to reveal cell surface proteins.

With the implementation of these targeted strategies, the view of the cell surface proteome of PSCs is beginning to emerge, but it is not clear that the cell surface has been completely mapped and functional studies using any putatively informative markers are pending. However, optimism for the future significant impact of proteomic analyses of PSCs is well-founded. For example, the “cell surface barcode” useful for identifying cells in the HSC system currently employs a number of cluster of differentiation (CD) molecules. Figure 5

illustrates a qualitative CD molecule barcode, including 146 CD molecules, identified thus far in human and mouse PSCs based upon the studies reviewed here. Additionally, using the CSC-Technology on mouse PSCs (preliminary data), our labs have identified additional CD molecules and non-CD molecules not described in the data reviewed here. We expect that this CD molecule barcode will be a useful starting point for immunophenotyping PSCs using commercially available antibodies, and integrating quantitative measures and specific protein modification information in the future is likely to provide additional specificity to this evolving “cell surface barcode”.

We envision that an iterative strategy combining targeted surface accessible protein discovery and quantification, affinity reagent development, live cell sorting, and downstream functional characterizations of the sorted subpopulations will contribute to a better understanding of the extent of heterogeneity present in PSC cultures. Moreover and analogous to HSCs, antibodies against accessible cell surface proteins should assist in the immunophenotyping of these cells as well as the isolation of specific populations suitable for developmental studies or regenerative medicine. To illustrate a practical example of how surface markers can be used to sort subpopulations of PSCs, we sorted live SSEA-1^{pos}(CD15) and SSEA-1^{neg} cells from ESCs with suboptimal colony morphology (Fig. 5B). To achieve this, live ESCs were obtained using a non-enzymatic cell dissociation solution (Millipore) to dissociate the colonies and were subsequently stained with SSEA-1 (eBioscience, E025199) prior to sorting. Sorted cells were plated and expanded and the SSEA-1^{pos} colonies showed better colony morphology and higher OCT4 abundance than the SSEA-1^{neg} population (Fig. 5B). While SSEA-1 is neither a protein nor is it specific to the pluripotent state, this example highlights the ability of sorting live PSCs by surface accessible markers for subsequent expansion and downstream functional characterization. In another example, CD326 was recently shown to be useful for the isolation and subsequent expansion of live iPSCs [79]. Thus, it is reasonable to predict that the application of appropriate proteomic technologies to expand the repertoire of cell surface markers for PSCs will lead to the development of affinity-based strategies useful for sorting defined subpopulations without relying on genetic modifications.

Finally and based upon these meta-analyses, a better view of the cell surface landscape is beginning to emerge for PSCs. However, several key challenges and questions still remain. First, to implement the newly identified surface markers into a cell sorting strategy, it is likely that new affinity reagents suitable for live cell sorting will need to be developed, which is especially difficult and time-intensive for integral membrane proteins. Second, the true value of any stem cell marker (natural or transgenic) will ultimately be determined by its ability to isolate defined stem cell populations with known differentiation potentials. Only through this process will it be possible to isolate therapeutically relevant stem cell populations useful in regenerative medicine. Thus, new and improved methods and reagents must be employed to facilitate the visualization and functional characterization of normal single stem cells within a population, and this is one area where proteomic technologies are expected to make a significant impact. Especially, proteomic strategies that provide insight into surface accessible protein domains are expected to facilitate these efforts by accelerating the epitope design and affinity reagent development for live cell immunophenotyping. Third, from a biological perspective, it may not be possible to develop a panel of surface markers that can isolate a truly homogeneous population of PSCs due to the extensive heterogeneity that is continually being discovered, but the development of antibody panels for live cell sorting is likely to be informative in evaluating heterogeneity among lines and isolating more therapeutically viable iPSCs. Finally, as cell surface proteins play critical roles in cell signaling, proliferation, and adhesion (reviewed in [80–82]), and represent a large percentage of current drug targets [83], the analysis of cell surface proteins is expected to greatly facilitate our understanding of how stem cells interact with their

microenvironment as well as facilitate the study of stem cell-derived disease models, apart from their role in developing surface marker panels for sorting and identifying PSC populations.

6 Conclusions

Heterogeneity within and among PSC colonies, cultures, and cell lines present significant challenges to interpreting metadata such as genome-wide and proteome-wide analyses. Recognition of these challenges throughout the experimental design and analysis will enable proteomic technologies to make significant contributions to the stem cell community, especially as the strategies and analytical instrumentation improve and we are routinely able to identify and quantify proteins from smaller amounts of starting material. As our view of the functional heterogeneity associated with PSCs continues to evolve, the proteomics community will adapt existing and emerging technologies to further address the changing needs of the stem cell community. Specifically, targeted proteomic strategies that provide insight into the cell surface proteome will facilitate a better understanding of stem cell biology as well as foster the development of novel and practical molecular tools necessary for the study of PSCs. Moreover, it is likely that an extended analysis of the accessible PSC proteome will lead to new insights and possibly exclusionary data that will be useful in the harvesting and characterization of therapeutically viable iPSCs as opposed to partially reprogrammed “iPSCs” that retain traits of the originating cell type. Proteomic technologies thus have the potential of playing a critical role in revolutionizing the stem cell field.

Abbreviations

CD	cluster of differentiation
ESC	embryonic stem cell
hESC	human ESC
HSC	hematopoietic stem cell
iPSC	induced pluripotent stem cell
mESC	mouse ESC
PM	plasma membrane
PSC	pluripotent stem cell

Acknowledgments

The authors are supported by 4R00HL094708-03 (R.L.G.), the Innovation Center at the Medical College of Wisconsin (R.L.G.), Maryland Stem Cell Research Fund Postdoctoral Fellowship (P.W.B.), the Intramural Research Program of the NIH, National Institute on Aging (K.R.B.), and NIH Induced Pluripotent Stem Cell Center (NiPSCC) Pilot Study Award (K.R.B.).

References

1. Wobus AM, Boheler KR. Embryonic stem cells: prospects for developmental biology and cell therapy. *Physiol. Rev.* 2005; 85:635–678. [PubMed: 15788707]
2. Boheler KR. Stem cell pluripotency: a cellular trait that depends on transcription factors, chromatin state and a checkpoint deficient cell cycle. *J. Cell Physiol.* 2009; 221:10–17. [PubMed: 19562686]
3. Boheler KR. Pluripotency of human embryonic and induced pluripotent stem cells for cardiac and vascular regeneration. *Thromb. Haemostasis.* 2010; 104:23–29. [PubMed: 20458433]
4. Bryder D, Rossi DJ, Weissman IL. Hematopoietic stem cells: the paradigmatic tissue-specific stem cell. *Am. J. Pathol.* 2006; 169:338–346. [PubMed: 16877336]

5. Weissman IL, Shizuru JA. The origins of the identification and isolation of hematopoietic stem cells, and their capability to induce donor-specific transplantation tolerance and treat autoimmune diseases. *Blood*. 2008; 112:3543–3553. [PubMed: 18948588]
6. Bryder D, Rossi DJ, Weissman IL. Hematopoietic stem cells – the paradigmatic tissue-specific stem cell. *Am. J. Pathol.* 2006; 169:338–346. [PubMed: 16877336]
7. Orlacchio A, Bernardi G, Orlacchio A, Martino S. Stem cells: an overview of the current status of therapies for central and peripheral nervous system diseases. *Curr. Med. Chem.* 2010; 17:595–608. [PubMed: 20088765]
8. Barker N, Bartfeld S, Clevers H. Tissue-resident adult stem cell populations of rapidly self-renewing organs. *Cell Stem Cell*. 2010; 7:656–670. [PubMed: 21112561]
9. Fuchs E. Finding one's niche in the skin. *Cell Stem Cell*. 2009; 4:499–502. [PubMed: 19497277]
10. Karp JM, Teol GSL. Mesenchymal stem cell homing: the devil is in the details. *Cell Stem Cell*. 2009; 4:206–216. [PubMed: 19265660]
11. Casali A, Battle E. Intestinal stem cells in mammals and drosophila. *Cell Stem Cell*. 2009; 4:124–127. [PubMed: 19200801]
12. Dick JE. Stem cell concepts renew cancer research. *Blood*. 2008; 112:4793–4807. [PubMed: 19064739]
13. Kelly PN, Dakic A, Adams JM, Nutt SL, Strasser A. Tumor growth need not be driven by rare cancer stem cells. *Science*. 2007; 317:337. [PubMed: 17641192]
14. Williams RT, den Besten W, Sherr CJ. Cytokine-dependent imatinib resistance in mouse BCR-ABL(1), Arnull lymphoblastic leukemia. *Genes Dev.* 2007; 21:2283–2287. [PubMed: 17761812]
15. Tan BT, Park CY, Ailles LE, Weissman IL. The cancer stem cell hypothesis: a work in progress. *Lab. Invest.* 2006; 86:1203–1207. [PubMed: 17075578]
16. Weissman I. Stem cell research - paths to cancer therapies and regenerative medicine. *J. Am. Med. Assoc.* 2005; 294:1359–1366.
17. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*. 2006; 126:663–676. [PubMed: 16904174]
18. Takahashi K, Tanabe K, Ohnuki M, Narita M, et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell*. 2007; 131:861–872. [PubMed: 18035408]
19. Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, et al. Induced pluripotent stem cell lines derived from human somatic cells. *Science*. 2007; 318:1917–1920. [PubMed: 18029452]
20. Park IH, Zhao R, West JA, Yabuuchi A, et al. Reprogramming of human somatic cells to pluripotency with defined factors. *Nature*. 2008; 451:141–146. [PubMed: 18157115]
21. Lowry WE, Richter L, Yachechko R, Pyle AD, et al. Generation of human induced pluripotent stem cells from dermal fibroblasts. *Proc. Natl. Acad. Sci. USA*. 2008; 105:2883–2888. [PubMed: 18287077]
22. Yu J, Hu K, Smuga-Otto K, Tian S, et al. Human induced pluripotent stem cells free of vector and transgene sequences. *Science*. 2009; 324:797–801. [PubMed: 19325077]
23. Lin T, Ambasudhan R, Yuan X, Li W, et al. A chemical platform for improved induction of human iPSCs. *Nat. Methods*. 2009; 6:805–808. [PubMed: 19838168]
24. Li Y, Zhang Q, Yin X, Yang W, et al. Generation of iPSCs from mouse fibroblasts with a single gene, Oct4, and small molecules. *Cell Res*. 2011; 21:196–204. [PubMed: 20956998]
25. Okita K, Ichisaka T, Yamanaka S. Generation of germline-competent induced pluripotent stem cells. *Nature*. 2007; 448:313–317. [PubMed: 17554338]
26. Nakagawa M, Koyanagi M, Tanabe K, Takahashi K, et al. Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nat. Biotechnol.* 2008; 26:101–106. [PubMed: 18059259]
27. Aoi T, Yae K, Nakagawa M, Ichisaka T, et al. Generation of pluripotent stem cells from adult mouse liver and stomach cells. *Science*. 2008; 321:699–702. [PubMed: 18276851]
28. Stadtfeld M, Brennand K, Hochedlinger K. Reprogramming of pancreatic beta cells into induced pluripotent stem cells. *Curr. Biol.* 2008; 18:890–894. [PubMed: 18501604]
29. Kim JB, Zaehres H, Wu GM, Gentile L, et al. Pluripotent stem cells induced from adult neural stem cells by reprogramming with two factors. *Nature*. 2008; 454:646–654. [PubMed: 18594515]

30. Eminli S, Utikal J, Arnold K, Jaenisch R, Hochedlinger K. Reprogramming of neural progenitor cells into induced pluripotent stem cells in the absence of exogenous Sox2 expression. *Stem Cells*. 2008; 26:2467–2474. [PubMed: 18635867]
31. Hanna J, Markoulaki S, Schorderet P, Carey BW, et al. Direct reprogramming of terminally differentiated mature B lymphocytes to pluripotency. *Cell*. 2008; 133:250–264. [PubMed: 18423197]
32. Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, et al. Induced pluripotent stem cell lines derived from human somatic cells. *Science*. 2007; 318:1917–1920. [PubMed: 18029452]
33. Takahashi K, Tanabe K, Ohnuki M, Narita M, et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell*. 2007; 131:861–872. [PubMed: 18035408]
34. Aasen T, Raya A, Barrero MJ, Garreta E, et al. Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. *Nat. Biotechnol.* 2008; 26:1276–1284. [PubMed: 18931654]
35. Sun N, Panetta NJ, Gupta DM, Wilson KD, et al. Feeder-free derivation of induced pluripotent stem cells from adult human adipose stem cells. *Proc. Natl. Acad. Sci. USA*. 2009; 106:15720–15725. [PubMed: 19805220]
36. Yamanaka S. A fresh look at iPS cells. *Cell*. 2009; 137:13–17. [PubMed: 19345179]
37. Brambrink T, Foreman R, Welstead GG, Lengner CJ, et al. Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. *Cell Stem Cell*. 2008; 2:151–159. [PubMed: 18371436]
38. Stadtfeld M, Maherali N, Breault DT, Hochedlinger K. Defining molecular cornerstones during fibroblast to iPS cell reprogramming in mouse. *Cell Stem Cell*. 2008; 2:230–240. [PubMed: 18371448]
39. Maitra A, Arking DE, Shivapurkar N, Ikeda M, et al. Genomic alterations in cultured human embryonic stem cells. *Nat. Genet.* 2005; 37:1099–1103. [PubMed: 16142235]
40. Hoffman LM, Hall L, Batten JL, Young H, et al. X-inactivation status varies in human embryonic stem cell lines. *Stem Cells*. 2005; 23:1468–1478. [PubMed: 16123389]
41. Werbowetski-Ogilvie TE, Bosse M, Stewart M, Schnerch A, et al. Characterization of human embryonic stem cells with features of neoplastic progression. *Nat. Biotechnol.* 2009; 27:91–97. [PubMed: 19122652]
42. Graf T, Stadtfeld M. Heterogeneity of embryonic and adult stem cells. *Cell Stem Cell*. 2008; 3:480–483. [PubMed: 18983963]
43. Canham MA, Sharov AA, Ko MS, Brickman JM. Functional heterogeneity of embryonic stem cells revealed through translational amplification of an early endodermal transcript. *PLoS Biol.* 2010; 8 e1000379.
44. Hayashi K, Lopes SM, Tang F, Surani MA. Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states. *Cell Stem Cell*. 2008; 3:391–401. [PubMed: 18940731]
45. Rathjen J, Lake JA, Bettess MD, Washington JM, et al. Formation of a primitive ectoderm like cell population, EPL cells, from ES cells in response to biologically derived factors. *J. Cell Sci.* 1999; 112:601–612. [PubMed: 9973595]
46. Narsinh KH, Sun N, Sanchez-Freire V, Lee AS, et al. Single cell transcriptional profiling reveals heterogeneity of human induced pluripotent stem cells. *J. Clin. Invest.* 2011; 121:1217–1221. [PubMed: 21317531]
47. Osafune K, Caron L, Borowiak M, Martinez RJ, et al. Marked differences in differentiation propensity among human embryonic stem cell lines. *Nat. Biotechnol.* 2008; 26:313–315. [PubMed: 18278034]
48. Kim DS, Lee JS, Leem JW, Huh YJ, et al. Robust enhancement of neural differentiation from human ES and iPS cells regardless of their innate difference in differentiation propensity. *Stem Cell Rev.* 2010; 6:270–281. [PubMed: 20376579]
49. Lister R, Pelizzola M, Kida YS, Hawkins RD, et al. Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature*. 2011; 471:68–73. [PubMed: 21289626]

50. Feng Q, Lu SJ, Klimanskaya I, Gomes I, et al. Hemangioblastic derivatives from human induced pluripotent stem cells exhibit limited expansion and early senescence. *Stem Cells*. 2010; 28:704–712. [PubMed: 20155819]
51. Banito A, Rashid ST, Acosta JC, Li S, et al. Senescence impairs successful reprogramming to pluripotent stem cells. *Genes Dev*. 2009; 23:2134–2139. [PubMed: 19696146]
52. Vaziri H, Chapman KB, Guigova A, Teichroeb J, et al. Spontaneous reversal of the developmental aging of normal human cells following transcriptional reprogramming. *Regen. Med*. 2010; 5:345–363. [PubMed: 20230312]
53. Wilmut I, Schnieke AE, McWhir J, Kind AJ, Campbell KH. Viable offspring derived from fetal and adult mammalian cells. *Nature*. 1997; 385:810–813. [PubMed: 9039911]
54. Kelly KF, Ng DY, Jayakumaran G, Wood GA, et al. Beta-catenin enhances Oct-4 activity and reinforces pluripotency through a TCF-independent mechanism. *Cell Stem Cell*. 2011; 8:214–227. [PubMed: 21295277]
55. Hu G, Kim J, Xu Q, Leng Y, et al. A genome-wide RNAi screen identifies a new transcriptional module required for self-renewal. *Genes Dev*. 2009; 23:837–848. [PubMed: 19339689]
56. Mallanna SK, Rizzino A. Systems biology provides new insights into the molecular mechanisms that control the fate of embryonic stem cells. *J. Cell Physiol*. 2011
57. Williamson AJ, Whetton AD. The requirement for proteomics to unravel stem cell regulatory mechanisms. *J. Cell Physiol*. 2011; 226:2478–2483. [PubMed: 21792904]
58. Reiland S, Salekdeh GH, Krijgsvelde J. Defining pluripotent stem cells through quantitative proteomic analysis. *Exp. Rev. Proteomics*. 2011; 8:29–42.
59. Hutchins AP, Robson P. Unraveling the human embryonic stem cell phosphoproteome. *Cell Stem Cell*. 2009; 5:126–128. [PubMed: 19664982]
60. Wang, J.; Trowbridge, JJ.; Rao, S.; Orkin, SH. Proteomic studies of stem cells. Cambridge: StemBook, Harvard Stem Cell Institute; 2008.
61. Shizuru JA, Negrin RS, Weissman IL. Hematopoietic stem and progenitor cells: clinical and preclinical regeneration of the hematolymphoid system. *Annu. Rev. Med*. 2005; 56:509–538. [PubMed: 15660525]
62. Bryder D, Rossi DJ, Weissman IL. Hematopoietic stem cells: the paradigmatic tissue-specific stem cell. *Am. J. Pathol*. 2006; 169:338–346. [PubMed: 16877336]
63. Wollscheid B, Bausch-Fluck D, Henderson C, O'Brien R, et al. Mass-spectrometric identification and relative quantification of N-linked cell surface glycoproteins. *Nat. Biotechnol*. 2009; 27:378–386. [PubMed: 19349973]
64. Chan EM, Ratanasirinrawoot S, Park IH, Manos PD, et al. Live cell imaging distinguishes bona fide human iPS cells from partially reprogrammed cells. *Nat. Biotechnol*. 2009; 27:1033–1037. [PubMed: 19826408]
65. Stewart MH, Bosse M, Chadwick K, Menendez P, et al. Clonal isolation of hESCs reveals heterogeneity within the pluripotent stem cell compartment. *Nat. Methods*. 2006; 3:807–815. [PubMed: 16990813]
66. Dormeyer W, van Hoof D, Mummery CL, Krijgsvelde J, Heck AJ. A practical guide for the identification of membrane and plasma membrane proteins in human embryonic stem cells and human embryonal carcinoma cells. *Proteomics*. 2008; 8:4036–4053. [PubMed: 18763709]
67. Cordwell SJ, Thingholm TE. Technologies for plasma membrane proteomics. *Proteomics*. 2010; 10:611–627. [PubMed: 19834916]
68. Josic D, Clifton JG. Mammalian plasma membrane proteomics. *Proteomics*. 2007; 7:3010–3029. [PubMed: 17654460]
69. Macher BA, Yen TY. Proteins at membrane surfaces – a review of approaches. *Mol. Biosyst*. 2007; 3:705–713. [PubMed: 17882332]
70. Gundry RL, Boheler KR, Van Eyk JE, Wollscheid B. A novel role for proteomics in the discovery of cell-surface markers on stem cells: Scratching the surface. *Proteomics Clin. Appl*. 2008; 2:892–903. [PubMed: 19526049]
71. Dormeyer W, van Hoof D, Braam SR, Heck AJ, et al. Plasma membrane proteomics of human embryonic stem cells and human embryonal carcinoma cells. *J. Proteome Res*. 2008; 7:2936–2951. [PubMed: 18489135]

72. Harkness L, Christiansen H, Nehlin J, Barington T, et al. Identification of a membrane proteomic signature for human embryonic stem cells independent of culture conditions. *Stem Cell Res.* 2008; 1:219–227. [PubMed: 19383402]
73. McQuade LR, Schmidt U, Pascovici D, Stojanov T, Baker MS. Improved membrane proteomics coverage of human embryonic stem cells by peptide IPG-IEF. *J. Proteome Res.* 2009; 8:5642–5649. [PubMed: 19899800]
74. Elia G. Biotinylation reagents for the study of cell surface proteins. *Proteomics.* 2008; 8:4012–4024. [PubMed: 18763706]
75. Hofmann A, Gerrits B, Schmidt A, Bock T, et al. Proteomic cell surface phenotyping of differentiating acute myeloid leukemia cells. *Blood.* 2010; 116:e26–e34. [PubMed: 20570859]
76. Intoh A, Kurisaki A, Yamanaka Y, Hirano H, et al. Proteomic analysis of membrane proteins expressed specifically in pluripotent murine embryonic stem cells. *Proteomics.* 2009; 9:126–137. [PubMed: 19053146]
77. Gu B, Zhang J, Wang W, Mo L, et al. Global expression of cell surface proteins in embryonic stem cells. *PLoS One.* 2010; 5:e15795. [PubMed: 21209962]
78. Nunomura K, Nagano K, Itagaki C, Taoka M, et al. Cell surface labeling and mass spectrometry reveal diversity of cell surface markers and signaling molecules expressed in undifferentiated mouse embryonic stem cells. *Mol. Cell. Proteomics.* 2005; 4:1968–1976. [PubMed: 16176923]
79. Chen HF, Chuang CY, Lee WC, Huang HP, et al. Surface marker epithelial cell adhesion molecule and E-cadherin facilitate the identification and selection of induced pluripotent stem cells. *Stem Cell Rev.* 2011; 7:722–735. [PubMed: 21305366]
80. Doherty GJ, McMahon HT. Mediation, modulation, and consequences of membrane-cytoskeleton interactions. *Annu. Rev. Biophys.* 2008; 37:65–95. [PubMed: 18573073]
81. Mukherjee S, Ghosh RN, Maxfield FR. Endocytosis. *Physiol. Rev.* 1997; 77:759–803. [PubMed: 9234965]
82. Buckley CD, Rainger GE, Bradfield PF, Nash GB, Simmons DL. Cell adhesion: more than just glue (review). *Mol. Membr. Biol.* 1998; 15:167–176. [PubMed: 10087503]
83. Yildirim MA, Goh KI, Cusick ME, Barabasi AL, Vidal M. Drug-target network. *Nat. Biotechnol.* 2007; 25:1119–1126. [PubMed: 17921997]
84. Alvarez-Manilla G, Warren NL, Atwood JA, Dalton S, et al. Glycoproteomic analysis of embryonic stem cells: identification of potential glycobiomarkers using lectin affinity chromatography of glycopeptides. *J. Proteome Res.* 2010; 9:2062–2075. [PubMed: 19545112]
85. Baharvand H, Fathi A, Gourabi H, Mollamohammadi S, Salekdeh GH. Identification of mouse embryonic stem cell-associated proteins. *J. Proteome Res.* 2008; 7:412–423. [PubMed: 18047272]
86. Baharvand H, Hajheidari M, Zonouzi R, Ashtiani SK, et al. Comparative proteomic analysis of mouse embryonic stem cells and neonatal-derived cardiomyocytes. *Biochem. Biophys. Res. Commun.* 2006; 349:1041–1049. [PubMed: 16973132]
87. Buhr N, Carapito C, Schaeffer C, Kieffer E, et al. Nuclear proteome analysis of undifferentiated mouse embryonic stem and germ cells. *Electrophoresis.* 2008; 29:2381–2390. [PubMed: 18449859]
88. Chaerkady R, Kerr CL, Kandasamy K, Marimuthu A, et al. Comparative proteomics of human embryonic stem cells and embryonal carcinoma cells. *Proteomics.* 2010; 10:1359–1373. [PubMed: 20104618]
89. Chambery A, Vissers JP, Langridge JJ, Lonardo E, et al. Qualitative and quantitative proteomic profiling of *cripto*($-/-$) embryonic stem cells by means of accurate mass LC-MS analysis. *J. Proteome Res.* 2009; 8:1047–1058. [PubMed: 19152270]
90. Elliott ST, Crider DG, Garnham CP, Boheler KR, Van Eyk JE. Two-dimensional gel electrophoresis database of murine R1 embryonic stem cells. *Proteomics.* 2004; 4:3813–3832. [PubMed: 15378706]
91. Graumann J, Hubner NC, Kim JB, Ko K, et al. SILAC-labeling and proteome quantitation of mouse embryonic stem cells to a depth of 5111 proteins. *Mol. Cell. Proteomics.* 2007; 7:672–683. [PubMed: 18045802]

92. Gundry RL, Tchernyshyov I, Sheng S, Tarasova Y, et al. Expanding the mouse embryonic stem cell proteome: combining three proteomic approaches. *Proteomics*. 2010; 10:2728–2732. [PubMed: 20512790]
93. Ho L, Jothi R, Ronan JL, Cui K, et al. An embryonic stem cell chromatin remodeling complex, esBAF, is an essential component of the core pluripotency transcriptional network. *Proc. Natl. Acad. Sci. USA*. 2009; 106:5187–5191. [PubMed: 19279218]
94. Jin J, Kwon YW, Paek JS, Cho HJ, et al. Analysis of differential proteomes of induced pluripotent stem cells by protein-based reprogramming of fibroblasts. *J. Proteome Res*. 2011; 10:977–989. [PubMed: 21175196]
95. Li QR, Xing XB, Chen TT, Li RX, et al. Large-scale phosphoproteome profiles comprehensive features of mouse embryonic stem cells. *Mol. Cell. Proteomics*. 2010; 10 M.110.001750.
96. Nagano K, Taoka M, Yamauchi Y, Itagaki C, et al. Large-scale identification of proteins expressed in mouse embryonic stem cells. *Proteomics*. 2005; 5:1346–1361. [PubMed: 15742316]
97. Pardo M, Lang B, Yu L, Prosser H, et al. An expanded Oct4 interaction network: implications for stem cell biology, development, and disease. *Cell Stem Cell*. 2010; 6:382–395. [PubMed: 20362542]
98. Prokhorova TA, Rigbolt KT, Johansen PT, Henningsen J, et al. Stable isotope labeling by amino acids in cell culture (SILAC) and quantitative comparison of the membrane proteomes of self-renewing and differentiating human embryonic stem cells. *Mol. Cell. Proteomics*. 2009; 8:959–970. [PubMed: 19151416]
99. Puente LG, Borris DJ, Carriere JF, Kelly JF, Megeney LA. Identification of candidate regulators of embryonic stem cell differentiation by comparative phosphoprotein affinity profiling. *Mol. Cell. Proteomics*. 2006; 5:57–67. [PubMed: 16188873]
100. Van Hoof D, Passier R, Ward-Van Oostwaard D, Pinkse MW, et al. A quest for human and mouse embryonic stem cell-specific proteins. *Mol. Cell. Proteomics*. 2006; 5:1261–1273. [PubMed: 16600995]
101. Wang Y, Mulligan C, Denyer G, Delom F, et al. Quantitative proteomics characterization of a mouse embryonic stem cell model of Down syndrome. *Mol. Cell. Proteomics*. 2009; 8:585–595. [PubMed: 19001410]
102. Brill LM, Xiong W, Lee KB, Ficarro SB, et al. Phosphoproteomic analysis of human embryonic stem cells. *Cell Stem Cell*. 2009; 5:204–213. [PubMed: 19664994]
103. Jaishankar A, Barthelery M, Freeman WM, Salli U, et al. Human embryonic and mesenchymal stem cells express different nuclear proteomes. *Stem Cells Dev*. 2009; 18:793–802. [PubMed: 18821827]
104. Moore JC, Fu J, Chan YC, Lin D, et al. Distinct cardiogenic preferences of two human embryonic stem cell (hESC) lines are imprinted in their proteomes in the pluripotent state. *Biochem. Biophys. Res. Commun*. 2008; 372:553–558. [PubMed: 18503758]
105. Schulz TC, Swistowska AM, Liu Y, Swistowski A, et al. A large-scale proteomic analysis of human embryonic stem cells. *BMC Genomics*. 2007; 8:478. [PubMed: 18162134]
106. Swaney DL, McAlister GC, Coon JJ. Decision tree-driven tandem mass spectrometry for shotgun proteomics. *Nat. Methods*. 2008; 5:959–964. [PubMed: 18931669]
107. Swaney DL, Wenger CD, Thomson JA, Coon JJ. Human embryonic stem cell phosphoproteome revealed by electron transfer dissociation tandem mass spectrometry. *Proc. Natl. Acad. Sci. USA*. 2009; 106:995–1000. [PubMed: 19144917]
108. Tian R, Wang S, Elisma F, Li L, et al. Rare cell proteomic reactor applied to stable isotope labeling by amino acids in cell culture (SILAC)-based quantitative proteomics study of human embryonic stem cell differentiation. *Mol. Cell. Proteomics*. 2011; 10 M110000679.

Figure 1.
Overview of PSCs. (A) Origin of adult, embryonic, and iPSCs. (B) Simplified view of the types of transcriptional heterogeneity present in PSC colonies.

Figure 2.

Examples of hESC heterogeneity. (A) Inter-line variation of transcription factors in undifferentiated hESC. The expression of the pluripotency markers POU5F1 (OCT4) and NANOG; early mesoderm marker T; early endoderm marker SOX17 and early ectoderm marker SOX1 in the hESC lines HUES-7, NOTT1, NOTT2, BG01, and H1 detected by immunohistochemistry. All lines demonstrated typical undifferentiated morphology and high levels of expression of pluripotency markers but also mixed variation in the expression of T and SOX17. Scale bar 550 mm. (B) Typical hESC colony demonstrating the three cellular phenotypes. 1, center of the colony which is very dense and showing the first signs

of neuralepithelial differentiation; 2, pluripotent “tight morphology” cells in slower growth phase; 3, normal pluripotent morphology of cells in the logarithmic growth phase.

Figure 3.

Current global view of the pluripotent proteome. (A) Venn diagram illustrating overlap between proteins identified in mouse and human PSCs. (B) Bar graphs showing the distribution of the number of studies in which a particular protein was identified, illustrating that a majority of proteins were identified in ≤ 2 studies. (C) Heat map of the 169 proteins identified in ten or more proteomic studies of either human or mouse PSCs. Listed are the gene symbol and the number of studies in which the protein was observed for each species human (H); mouse (M).

Figure 4.

Strategies and benefits of PM targeting. (A) General proteomic workflow for PM protein enrichment. (B) Chemical-based tagging methods for affinity enrichment. (C–E) Comparison of results from seven studies [63, 71–73, 76–78] that used either physical or affinity enrichment of PM proteins on human or mouse PSCs compared to 26 other studies that did not target the PM. (C) Overlap of the proteins identified by PM targeted and non-PM targeted methods. (D) Bar graphs illustrating the distribution of the percentage of transmembrane proteins identified in each type of strategy, calculated by TMAP algorithm within ProteinCenter. (E) Percentage of proteins unique to either the PM targeted or the non-PM targeted method which are annotated with selected gene ontology terms for cellular localization.

Figure 5. Qualitative proteomic CD molecule barcode for human and mouse PSCs and an example of how a CD molecule can be used for live cell FACS of subpopulations of PSCs. (A) All protein CD molecules are listed. Solid circle indicates that the CD molecule has been identified in the human (top circle) and mouse (bottom circle) PSC proteomes reviewed here. Since not all CD molecules are annotated for both species, gray indicates CD molecules that are not predicted for that species. (B) Phase contrast image of suboptimal mES colony morphology prior to FACS sorting. FACS histogram shows isotype control (grey) and SSEA-1 (black), sorting gates are indicated by horizontal bars. Images are of resulting colonies from the two sorted populations and quantitative analysis via Western blotting of OCT4 in each sorted population.

Table 1

Summary of proteomic studies of PSCs reviewed here

First author; year	PMID/PMCID	Species	Type ^{d)}	Cell line	# Proteins reported for undifferentiated state (# proteins used for analysis ^{b)})
A. Summary of proteomic studies of undifferentiated human and mouse PSCs reviewed here and are included in the PSC proteome compilation in Supporting Information Table 1					
Alvarez-Mamilla G; 2010	19545112	M	ES	D3, R1	133
Baharvand H; 2008	18047272	M	ES	Royan B1, D3	72
Baharvand H; 2006	16973132	M	ES	Royan B1	13
Buhr N; 2008	18449859	M	ES	129/Sv ES-CK35	33
Chambery A; 2009	19152270	M	ES	R1 and Cripto-deficient R1	265 (R1) 231 (Cripto-R1)
Elliott ST; 2004	15378706	M	ES	R1	218
Graumann J; 2008	18045802	M	ES	G-olig2, R1	5111
Gu B; 2010	PMC3012103	M	ES	D3	3468 (981)
Gundry RL; 2010	PMC2934747	M	ES	R1	1895
Ho L; 2009	PMC2654396	M	ES	E14Tg2a	197
Intoh A; 2009	19053146	M	ES	D3	831
Jim J; 2010	21175196	M	ES, iPS	B6, E14, iPS	1883
Kim J; 2010	20946988	M	ES	J1	122
Li QR; 2011	21149613	M	ES	E14.1	4581
Nagano K; 2005	15742316	M	ES	E14.1	1790
Nunomura K; 2005	16176923	M	ES	D3	324
Pardo M; 2010	PMC2860244	M	ES	AB2.2	92
Puente LG; 2006	16188873	M	ES	J1	108
Wang Y; 2009	PMC2667343	M	ES	D3, D3 with trisomy 21	1661
Wollscheid B; 2009	19349973	M	ES	159-2	226
Van Hoof D; 2006	16600995	M, H	ES	HUES-1, HES-2, D3	1775 (hES), 1871 (mES)
Dormeyer W; 2008	18489135	H	ES, EC	HUES-7, NT2/D1	1077 (hES) 1261 (hEC)
Baharvand H; 2006	16758447	H	ES	Royan H2, H3, H5	685
Brill LM; 2009	PMC2726933	H	ES	H1	1602
Chaerkady R; 2010	20104618	H	ES, EC	H1, NTERA-2 clD1	1800
Harkness L; 2008	19383402	H	ES	Odense-3	1075 (157)
Jaishankar A; 2009	18821827	H	ES	H1, H9	56

First author; year	PMID/PMCID	Species	Type ^{a)}	Cell line	# Proteins reported for undifferentiated state (# proteins used for analysis) ^{b)}
McQuade LR; 2009	19899800	H	ES	SIVF001	2851
Moore JC; 2008	PMC2665880	H	ES	H1 and HES2	55
Prokhorova TA; 2009	PMC2689770	H	ES	Odense-3, KMEB1, HUES9	1556
Schulz TC; 2007	PMC2211323	H	ES	BG01	620
Swaney DL; 2008	PMC2597439	H	ES	H1	4187
Swaney DL; 2009	PMC2633571	H	ES	H1	4339
Tian R; 2010	20530636	H	ES	H1, H9, CA1	1086

First author; year	PMID/PMCID	Species	Type ^{a)}	Cell line	Brief description
B. Summary of proteomic studies that examine early differentiation stages of human and mouse PSCs, but which are not included in the “PSC proteome”					
Gao X; 2009	19275242	M	ES	P19	Quantitative changes via aggregate formation and neural differentiation
Hewel JA; 2010	20467045	M	ES	E14Tg2a	Quantification of signaling pathway members during differentiation via LIF withdrawal
Mallanna SK; 2010	20687156	M	ES	KH2 with iSOX2, iSOX21	Sox2-binding proteins with early Sox2-mediated differentiation
McHugh PC; 2008	17868153	M	ES	R1	Quantitative changes via EB formation and neural differentiation
Osman AM; 2010	20553848	M	ES	D3	Semi-quantitative changes via EMP-induced differentiation
Prudhomme W; 2004	PMC365717	M	ES	CCE	Effects of extracellular signaling factors on self-renewal and differentiation
Williamson AJ; 2008	18045800	M	ES	E14.1	Quantitative changes via EB formation
Chae JI; 2009	19206105	H	ES	CHA-hES3, H9	Cytoskeleton changes upon EB formation and neural differentiation
Phanstiel D; 2008	18326628	H	ES	H1, H7, H9	Quantification of histone isoforms upon TPA-induced differentiation
Rigbolt KT; 2011	21406692	H	ES	Odense-3, HUES9	Quantification of proteins and phosphoproteins with nondirected differentiation
Van Hoof D; 2009	19664995	H	ES	HUES-7	Quantitative phosphorylation changes upon BMP-induced differentiation

This list contains the first author of the study and publication date, PubMed identifier, species (human: H, mouse: M), cell type, cell line, and number of proteins reported by the authors for the undifferentiated state (A) or a brief description (B) [63, 71–73, 76–78, 84–108].

^{a)} EC = embryonal carcinoma cell; EB = embryoid body.

^{b)} In some instances, protein identities for all proteins in the undifferentiated state were not publicly available. In this case, the number of proteins imported into ProteinCenter is different from that stated in the original publication.