Published in final edited form as:

Nat Rev Mol Cell Biol. 2013 June ; 14(6): 357–368. doi:10.1038/nrm3584.

Origins and implications of pluripotent stem cell variability and heterogeneity

Patrick Cahan and George Q. Daley

Stem Cell Transplantation Program, Division of Pediatric Hematology and Oncology, Manton Center for Orphan Disease Research, Howard Hughes Medical Institute, Boston Children's Hospital and Dana Farber Cancer Institute, Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Harvard Stem Cell Institute, Boston, Massachusetts 02115, USA

Abstract

Pluripotent stem cells constitute a platform to model disease and developmental processes and can potentially be used in regenerative medicine. However, not all pluripotent cell lines are equal in their capacity to differentiate into desired cell types *in vitro*. Genetic and epigenetic variations contribute to functional variability between cell lines and heterogeneity within clones. These genetic and epigenetic variations could 'lock' the pluripotency network resulting in residual pluripotent cells or alter the signalling response of developmental pathways leading to lineage bias. The molecular contributors to functional variability and heterogeneity in both embryonic stem (ES) cells and induced pluripotent stem (iPS) cells are only beginning to emerge, yet they are crucial to the future of the stem cell field.

The defining characteristics of pluripotent stem cells, which are self-renewal and the capacity to differentiate into any cell type in the body, make them attractive tools to investigate development and disease, and generate hope that in the future they can be used to replace diseased or damaged cells. In this Review, we use the term 'pluripotent stem cells' to generally refer to all types of pluripotent stem cells, including induced pluripotent stem (iPS) cells, embryonic stem (ES) cells and embryonal carcinoma (EC) cells. Functional, phenotypical and molecular differences in pluripotent stem cells have been observed since their isolation^{1,2}. The origins and implications of variability (which we define as variations among cell lines or between ES cells and iPS cells) and heterogeneity (which we define as variations within clones of cells) within ES cells have been difficult to investigate because of the small number of ES cell lines that are available for a comprehensive evaluation. Now, pluripotent stem cells are more accessible with the advent of transcription factor-based cellular reprogramming³. Initially, much attention focused on the similarities of iPS cells and ES cells. However, recent comparisons of pluripotent stem cells have uncovered molecular variations that correlate with their functional variability (defined as *in vitro* differentiation capacity). A pressing question is the extent to which variability and heterogeneity will affect the usefulness of pluripotent stem cells, regardless of the derivation method, in regenerative medicine, drug screening, disease modelling and in studying developmental processes. As more investigators contemplate using pluripotent stem cells

Correspondence to G.Q.D. george.daley@childrens.harvard.edu.

Competing interests statement

The authors declare [competing financial interests](#): see Web version for details.

FURTHER INFORMATION

George Q. Daley's homepage: <https://daley.med.harvard.edu>

ALL LINKS ARE ACTIVE IN THE ONLINE PDF

(BOX 1), and because substantial comparative analyses of diverse pluripotent stem cells have recently been completed, now is an opportune moment to re-evaluate the issue of functional variability among pluripotent stem cells, especially between ES cells (which have been viewed as the gold standard in the field) and iPS cells (which hold great promise for therapeutic benefit).

In this Review, we discuss the molecular and functional variability and heterogeneity of pluripotent stem cells that have been derived from different sources and using different techniques (TABLE 1). We evaluate these data from three different perspectives. First, on the basis of our current knowledge of the molecular under-pinnings of pluripotency, we discuss the functional differences among pluripotent stem cells and speculate on factors that contribute to functional variability. Next, we describe variability in global gene expression profiles and the epigenetic status among pluripotent stem cells. Finally, we review how derivation and extended passage in culture may introduce or select for genetic changes that affect research and clinical applications of pluripotent stem cells.

Research into pluripotent stem cells encompasses many areas of focus. We refer the readers to other recent reviews on pluripotency itself^{4,5} and the history of its investigation⁶, as well as on the potential applications of stem cells in regenerative medicine^{7,8}, disease modelling^{9–12}, developmental studies^{13,14} and drug development¹⁵.

The defining properties of stem cells

To appreciate the variability and heterogeneity in pluripotent stem cells, we need to introduce their defining characteristics: self-renewal and pluripotency (BOX 2).

Self-renewal properties

Self-renewal refers to the capacity of a cell to maintain its identity after division. A cell division in which at least one of the daughter cells is identical to the parent cell is self-renewing. To expand, pluripotent stem cells must undergo symmetric self-renewal divisions in which each daughter cell maintains pluripotency. This proliferative capacity of pluripotent stem cells is appealing because it gives researchers unlimited material that can be used over a long period of time and across laboratories. Pluripotent stem cells must fulfil two requirements to be self-renewing: they must maintain the pluripotent state following division (as described below) and they must continue to divide. Although continuing to divide is a defining feature of self-renewal, there is no evidence that it contributes to functional differences in derived pluripotent stem cell lines, and therefore we do not explore the topic further but we refer the reader to another review¹⁶.

Factors contributing to pluripotency

Pluripotency, which is the ability to give rise to all lineages of the embryo proper, only occurs transiently in the early embryo and is achieved through the coordination of diverse molecular pathways. Over the past 30 years, considerable efforts and resources have culminated in a persuasive yet incomplete model of the molecular underpinnings of pluripotency in mouse pluripotent stem cells. The hallmark of this model is a gene regulatory network governed by the autoregulated transcription factors OCT4 (also known as POU5F1), SOX2 and NANOG (collectively referred to as OSN), which, together with Polycomb group (PcG) chromatin regulators^{17,18}, maintain lineage specification programmes in a silent yet poised state^{19,20} (FIG. 1). There is an emerging view that these core factors separately promote specific fates, and their concurrent expression in pluripotent stem cells serves to mutually antagonize commitment to differentiation along any lineage^{21,22}. The pluripotency network also maintains signalling pathways competent to direct cells towards mesendodermal or ectodermal fates. OCT4 and SOX2 promote the

expression of fibroblast growth factor 4 (FGF4)²³, which directly phosphorylates MAPK1 (also known as ERK2). MAPK1 in turn ‘primes’ cells to respond appropriately to differentiation signals^{24,25}. Leukaemia inhibitory factor (LIF; which is supplied by feeder cells or added in purified form to feeder-free cultures) and bone morphogenetic protein 4 (BMP4; which is present in serum) counteract MAPK1-enabled differentiation and thereby promote self-renewal by acting through STAT3 (signal transducer and activator of transcription 3)²⁶, SMAD1 and ID (inhibitor of differentiation)²⁷. OSN promote the expression of signalling components that are required for early cell fate choice; in this way, these signalling pathways are kept in a poised state. For example, OSN promote the expression of Notch 1 (REF. 28), which enables neural differentiation of mouse ES cells²⁹. OSN also regulate the expression of WNT3 (REF. 28), and the WNT3 downstream effector β -catenin is required for mesoderm differentiation of mouse ES cells³⁰. However, unlike the Notch signalling pathway, the WNT pathway also has an active role in the self-renewal of mouse ES cells^{31,32} by interfering with T cell factor 3 (TCF3)-mediated repression of OSN target genes^{33,34}. Non-coding RNAs (ncRNAs) are important players in the pluripotency network, as OSN factors activate the pluripotent stem cell-specific microRNAs (miRNAs) miR-290 and miR-302 (REF. 35), which in turn inhibit the expression of pluripotency repressors³⁶. At the same time, OSN inhibit the expression of lineage-specific miRNAs such as miR-9 (REF. 37), which promotes the proliferation of neural progenitors derived from human ES cells.

miRNAs seem to be essential for stem cell differentiation. For example, mouse ES cells defective in DGCR8 (a double-strand RNA-binding protein that, together with Drosha, forms a complex responsible for processing primary miRNAs into precursor miRNAs) are similar to wild-type mouse ES cells in morphology, cell surface phenotype and expression of key pluripotency genes, and yet they form undifferentiated teratomas, resist *in vitro* differentiation upon withdrawal of LIF and reform ES cell colonies after days of directed differentiation at significantly higher rates than their wild-type counterparts³⁸. Similarly, mouse ES cells lacking either Dicer 1 (which cleaves pre-miRNAs) or methyl-CpG-binding domain protein 3 (MBD3, which is a part of the nucleosome remodelling and histone deacetylation (NuRD) complex) show normal expression levels of core pluripotency factors (consistent with a pluripotent state), but they fail to differentiate properly^{39,40} due to their inability to repress genetic programmes associated with pluripotency³⁶ (in the case of Dicer 1) or the trophectoderm lineage⁴¹ (in the case of MBD3).

These examples illustrate an important semantic difference between pluripotency as a state and pluripotency as a function. When pluripotency is defined as a state, then pluripotent stem cells lacking master regulators of differentiated cell types may still qualify as pluripotent even though they cannot fully differentiate. Similarly to miRNA-deficient cell lines described above, nullipotent EC cell lines (for example, F9 cells) are another example of self-renewing, pluripotent-like stem cells that are resistant to differentiation *in vitro* and form undifferentiated teratomas⁴². However, when pluripotency is defined as a function, such cells clearly are not pluripotent. A corollary of this distinction is that functional pluripotency cannot be predicted purely on the basis of markers of the pluripotent state until the integrity of all lineage specification programmes has been tested. Another implication of the distinction between pluripotency as a state versus a function is that the mechanisms that maintain the pluripotent state (that is, the networks that maintain access to lineage determination and differentiation programmes) (FIG. 1) are, to some extent, independent of the mechanisms that execute lineage determination and differentiation programmes. Describing pluripotency in this way provides a framework to link functional differences in pluripotent stem cells to potential molecular mechanisms underlying these differences and thus may open up new experimental strategies to manipulate these cells. It is likely that

these principles will apply to all cell types defined functionally by their potential rather than their steady state.

There are significant differences between human and mouse pluripotent stem cells. Human pluripotent stem cells more closely resemble mouse pluripotent stem cells that were originally derived from post-implantation embryos (termed epiblast stem cells)^{43,44} than mouse pluripotent stem cells derived from pre-implantation embryos or by reprogramming with OCT4, SOX2, Krüppel-like factor 4 (KLF4) and MYC (collectively referred to as OSKM). Some notable differences include the failure of female derived human pluripotent stem cells and epiblast stem cells to reactivate the silenced X chromosome as well as their reliance on FGF and either activin or Nodal signalling for self-renewal, which cause differentiation in mouse pluripotent stem cells rather than LIF and inhibitors of MEK and glycogen synthase 3 (GSK3) (reviewed in REF. 45). The resemblance of human pluripotent stem cells and epiblast stem cells is important with regard to functional variability because epiblast stem cells are often associated with a theoretical pluripotent stem cell state primed for differentiation⁴⁶ that exists between a self-renewing ‘naive’ state (also referred to as ground pluripotent stem cell state) and a state committed to differentiation. It is tempting to speculate that stochastic molecular events during the derivation process of human pluripotent stem cells could result in a cell line that is primed with a subtle lineage bias, and this bias can be observed in directed differentiation towards specific lineages. This fate bias might, however, be overridden by the complex signalling environment that is active in teratoma assays.

Perturbations to the pluripotency network that cause functional differences in pluripotent stem cells may either be detectable or remain silent at the gene expression level in the pluripotent state. For example, altered sensitivity to differentiation signals may be caused by variable expression of components of the lineage commitment signalling networks. Alternatively, genetic or epigenetic disruptions to any commitment or differentiation programme, leading to an altered differentiation propensity, will remain largely silent in the pluripotent state. Similarly, genetic or epigenetic disruption of programmes that repress the pluripotency network upon differentiation initiation (which creates a resistance to differentiate) will also be silent in the pluripotent state. Evans and Kaufmann invoked a “degree of differentiative ability rather than selective restriction” (REF. 47) as one possible explanation for the distinct cell type distributions in teratomas that develop from mouse ES cell subclones⁴⁷, anticipating, 30 years ago, the concept of stem cell variability.

Functional variability

Pluripotent stem cells differ in the capacity to form distinct cell types in teratomas; the capacity to produce different cell types is more easily quantifiable by directed differentiation *in vitro* (BOX 3). Above, we outlined possible perturbations to the pluripotent state that would have functional outcomes. Here, we discuss pluripotency as a function and summarize three studies that examined the *in vitro* differentiation propensities of large numbers of human pluripotent stem cell lines. One of the first studies to characterize *in vitro* lineage potential of these cells⁴⁸ evaluated spontaneous or directed differentiation to cardiomyocytes and pancreatic cell types in 17 distinct human ES cell lines. In this study, the lineage potential was assessed by the expression of genes that are indicative of the pluripotent state in the three germ layers and eight organs at four time points. Unlike pluripotency genes, which showed uniform expression in all cell lines at all time points, the expression of most germ layer markers and their derivatives varied significantly, suggesting a marked diversity in inherent differentiation propensity. There was a general correlation between germ layer-specific gene expression and derivative cell type marker expression, although some cell lines preferentially expressed markers specific for a cell subtype (for

example, some ectoderm-specific lines expressed only neural markers and not skin markers). This could be explained by a perturbation to differentiation programmes subsequent to ectoderm commitment. In an independent set of experiments, six of the 17 cell lines described above exhibited significant differences in motor neuron differentiation⁴⁹, suggesting that functional differences in pluripotent stem cells cannot be solely attributed to laboratory-specific effects.

In another study, neural differentiation ability was determined for five human ES cell lines and 12 human iPS cell lines (which were generated by either integrating or transgene-free strategies)⁵⁰. Differentiation to neuroepithelial cells, neural progenitors and motor neurons was quantified by whole-cell patch-clamp recordings and the expression of genes specific for each cell type (that is, paired box 6 (PAX6), orthodenticle homolog 2 (OTX2), homeobox 9 (HB9), ISL LIM homeobox 1 (ISL1) and ISL2). iPS cells showed significantly lower neuroepithelial differentiation capacity than ES cells, regardless of donor cell type and method of reprogramming factor delivery. Interestingly, neuroepithelial differentiation of several iPS cell lines was significantly improved by media supplementation with FGF2 or inhibition of SMAD, suggesting that altered expression of lineage commitment signalling networks changed the normal sensitivity of these lines to FGF2 and SMAD. However, the differentiation of two cell lines into neuroepithelial cells was unchanged by either FGF2 or SMAD inhibition, suggesting distinct causes for differentiation resistance and that these two cell lines had acquired silent disruptions to the pluripotency network that impeded neural commitment. Although iPS cells differentiated into neuroepithelial cells with lower efficiency, their ability to form fully differentiated TUJ1 (which is a neuron specific class III β -tubulin) expressing neurons was equivalent to ES cells. This implies that although the initial propensity to commit to the ectoderm lineage was reduced in iPS cells, the neural differentiation programmes had remained largely intact. When coaxed to form motor neurons, iPS cell-derived neuroepithelial cells formed significantly more OLIG2 (oligodendrocyte transcription factor 2) positive motor neuron progenitors and fewer HB9 positive post-mitotic motor neurons. It is interesting that this phenotype cannot be corrected by reduced exposure to sonic hedgehog (SHH) or retinoic acid, which are known to promote the differentiation of OLIG2 positive progenitors. This suggests that the distinct motor neuron differentiation profiles of human iPS cell- and human ES cell-derived neuroepithelial cells is attributable to epigenetic perturbations to the neural differentiation programmes, as other disruptions would inhibit the initial ectoderm fate choice.

A third type of study compared the motor neuron differentiation capacity of 16 human iPS cells, including six from patients with amyotrophic lateral sclerosis (ALS) and five human ES cell lines in two independent laboratories⁵¹. The cell lines varied significantly and reproducibly in differentiation efficiency, but these differences were not correlated to the source (ES cells or iPS cells), karyotype or residual *trans*-gene expression. Despite line-to-line variation in differentiation efficiency, the apparent quality of the motor neuron derivatives was equivalent as measured by electrophysiological recordings. Supporting a model in which the more recalcitrant cell lines have greater sensitivity to mesendoderm inducing signals, SMAD inhibition rescued the differentiation deficiency, and all cell lines formed well-differentiated teratomas.

Variations in differentiation efficiency into other lineages are likely to occur. Indeed, variation in erythroid, myeloid and osteoclast differentiation was noted in three human ES cell lines⁵² and in haematopoietic differentiation in five human ES cell lines⁵³.

Taken together, these studies illustrate several important principles. First, even pluripotent stem cells that pass the qualitative teratoma assay can vary substantially in their *in vitro* differentiation capacity. Clearly, *in vitro* differentiation does not attain the signalling

complexity inherent to the teratoma formation assay, much less in the spatially and temporally dynamic milieu of the developing embryo. In these environments, pluripotent stem cells are overwhelmed by redundant and reinforcing signalling cues that push them towards recognizable cell types. Second, the contributors to functional variation are multifactorial and not solely attributable to cellular source (embryonic or somatic), laboratory-specific effects, karyotype, passage number or other technical factors. Third, differentiation defects can be overcome by finding appropriate culture conditions.

Gene expression heterogeneity and variation

The stem cell field has been at the forefront of applying gene expression profiling technology to better understand the mechanisms of lineage potential and self-renewal, to identify crucial and irrelevant differences among pluripotent stem cell lines and to explore the role of population heterogeneity in stem cell function.

The search for a stemness signature

The early days of genome-wide expression profiling witnessed a flurry of studies aimed at defining a transcriptional signature of stemness. The original large-scale studies compared cDNA arrays from mouse ES cells and cells induced to differentiate by retinoic acid exposure. These studies were first-in-class experiments⁵⁴; for example, they first reported that mouse ES cells actively express proteins that repress homeobox (HOX) expression to maintain pluripotency⁵⁵. As different types of stem cells (such as embryonic, neural, retinal and haematopoietic stem cells (HSCs)) share the properties of self-renewal and multipotency (albeit different degrees of potency), a simple and compelling hypothesis emerged that molecular mechanisms of stemness would be conserved across cell types. To test this hypothesis and identify a stemness signature, several groups independently performed gene expression profiling experiments in stem cells and their differentiated progeny^{56–58}. All groups used a similar approach to define the stemness signature, and they all identified genes that were specifically upregulated in stem cells. If all stem cells did utilize the same mechanisms to maintain stemness, then genes encoding factors responsible for stemness should have been common to all stem cell signatures. Surprisingly, however, the stemness signatures did not overlap, whereas common lineage-specific genes (for example, neural progenitor cell signature) were consistently detected in multiple studies. This suggests that if a stemness programme that is common to these cell types indeed exists, it is not detectable using these methods and is much more subtle than expected. This finding was confirmed by thorough characterization of the expression profiles of genetically identical mouse ES cells and neural progenitors⁵⁹. These data suggest that different types of stem cells use separate cellular machineries to dictate self-renewal and to maintain multilineage potential. Microarray approaches were also used to identify transcriptional differences that would explain distinct differentiation capacities between trophectoderm and ES cells⁶⁰ and the distinct signalling requirements of mouse and human ES cells^{61,62}. Microarray experiments were also performed to determine whether distinct molecular programmes govern pluripotency and tumour formation in human ES cells, EC cell lines and seminomas⁶³.

Variations among cell lines

Variability in expression profiles of pluripotent cell lines became apparent for the first time in studies that identified genes that distinguish human ES cells from differentiated cells^{64,65}. The first study reported thousands of probesets that are uniquely detected in three human ES cell lines⁶⁶, although it is difficult to reconcile the extent of these differences with many subsequent studies in which orders of magnitude fewer genes were found to be uniquely expressed in specific pluripotent stem cell lines. More recently, the determination of differences between different ES cell lines has been overshadowed by the comparison of iPS

cells and ES cells. The initial wave of iPS cell profiling invoked the global similarities between ES cells and iPS cells when compared to fibroblasts or other somatic cells as one line of evidence that reprogrammed cells transitioned to a state closely resembling ES cells^{3,67–70}. One of the most extensive analysis of gene and miRNA expression profiles, DNA copy number (by comparative genomic hybridization) and histone methylation profiles (trimethylation of Lys27 of H3 (H3K27me3) and H3K4me3)⁷¹ in human pluripotent stem cells identified a gene signature that distinguished reprogrammed iPS cells from ES cells. Two other studies also reported substantial differences in expression profiles between human ES cells and human iPS cells^{72,73}, fuelling a debate in the field as to whether the reported gene expression differences were experimental artefacts or represented fundamental and consistent differences between human ES cells and human iPS cells that might affect their function.

A subsequent analysis of the expression profiles of 27 human ES cell and iPS cell lines revealed only four differentially expressed genes between these two types of cell lines and suggested that the apparent iPS cell signature was a result of laboratory-specific effects rather than reflecting an underlying consistent difference between ES cells and iPS cells⁷⁴, which is in agreement with an independent analysis⁷⁵. It is important to note that both studies found gene expression differences between iPS cell and ES cell lines, consistent with previous reports of expression variability, but that no differences could consistently and robustly distinguish iPS cells from ES cells. Notably, both studies examined H3K27me3 and H3K4me3 and showed that there were no consistent differences in the epigenetic profiles that could explain the observed expression variations. To specifically address the question of residual donor cell gene expression in human iPS cells, one study used a more sensitive approach than previous studies. In this study, the overall correlation between the p-values of the differential expression between donor cells and ES cells as well as between iPS cells and ES cells were determined as an alternative to comparing lists of differentially expressed genes or gene set enrichment analysis⁷⁶. They found a statistically significant correlation, supporting the hypothesis of both residual expression and incomplete expression of pluripotency factors, but it should be noted that the individual p-values of the residually expressed genes seemed to be non-significant.

Since these reports, the number of human pluripotent stem cell lines that have been profiled has nearly tripled, providing an opportunity to revisit the question of whether iPS cells represent a distinct subtype of pluripotent stem cells. A recent analysis of 272 microarrays of human ES cells and iPS cells from 19 experiments revealed that ES cells are globally as similar to iPS cells as they are to other ES cells (as determined by Pearson correlation coefficient), but that the inter-iPS cell correlations are more varied than the inter-ES cell correlations (as determined by the Kolmogorov–Smirnov test) (P.C. and G.Q.D., unpublished data). Such analysis does not support the hypothesis that iPS cells are a distinct subtype of pluripotent stem cells. The data only show that they exhibit more heterogeneous expression profiles, which is consistent with data from a study analysing the of expression profiles of 12 human iPS cell lines and 20 ES cell lines⁷⁷. This is not surprising given the highly varied origins of iPS cells, including different donor cell types, donors of varying age, different genetic backgrounds and diverse methods of reprogramming. We expect that as more cell lines are derived under less-than-ideal conditions (such as from aged donors) and from more diverse cell types, this wide distribution of gene expression will persist. In fact, a recent study comparing RNA sequencing data from human ES cells and iPS cells that have been reprogrammed using the episome method reported hundreds of genes that are differentially expressed⁷⁸. Moreover, the expression of nine genes identified in this study were shown to be consistently differentially methylated in human iPS cells and ES cells, and this was sufficient to distinguish iPS cells and ES cells⁷⁹, albeit not perfectly. Nonetheless, it is clear that ES cells and iPS cells can be virtually indistinguishable at a global

transcriptional level, which is consistent with the hierarchical model of the pluripotency network in which ectopic expression of OSKM induces the stable expression of endogenous OSN and results in a highly robust execution of downstream programmes that collectively maintain the poised and self-renewing state of pluripotent stem cells (FIG. 1). This model suggests that small sets of genes may be differentially expressed, and this does not substantially perturb the steady-state expression profile of pluripotent stem cells (for example, genes that encode components of lineage commitment signalling networks). Consistent with this model, it was discovered that miR-371-3 expression levels in undifferentiated human pluripotent stem cells were sufficient to accurately predict neural differentiation propensity⁸⁰, although the mechanisms behind this link are unknown.

Heterogeneity within individual cell lines

Several pluripotency-related transcription factors are heterogeneously expressed in mouse ES cell lines, including NANOG⁸¹, T box 3 (TBX3), KLF4 (REF. 82) and zinc-finger protein 42 (ZFP42; also known as REX1)⁸³, and similar expression heterogeneity has been observed in human iPS cell lines⁸⁴. Importantly, mouse ES cells with low or no NANOG expression are not merely differentiating cells because, when isolated as individual clones and allowed to expand, they reconstitute a population of pluripotent stem cells with a similar distribution of NANOG-expressing and non-expressing cells⁸⁵. This has led to the view that mouse ES cells maintained in serum and media supplemented with LIF are ‘meta-stable’, in that they transition between a ground state of pluripotency and a state primed to differentiate. Indeed, the transcript levels of lineage-associated transcription factors increase in single cells that have lost NANOG expression, and re-expression of NANOG silences lineage-specific genes⁸⁶. In contrast to the primed state, ground state mouse ES cells (maintained under more stringent, chemically defined conditions favouring self-renewal) do not exhibit fluctuations in pluripotency-related factors, nor do they transiently express lineage commitment factors or express components of differentiation signalling cascades that characterize the primed state (for example, MAPK)²⁴. Furthermore, lineage-specific transcription factors, for example PAX9, are not bivalent (defined as simultaneously having histone marks associated with transcriptional activation (H3K4me3) and repression (H3K27me))¹⁷ in ground state cells²⁴. What regulates this meta-stability is unclear, however, recent data indicate that it might be transcription factors such as PR domain zinc-finger protein 14 (PRDM14), which interferes with the activation of FGF by OSN and represses the expression of DNA methyltransferases in mouse ES cells⁸⁷. More work is required to assess whether a primed state is a requisite transition state, and, if not, how the phenomenon of lineage priming, that is, sporadic expression of lineage-associated genes in pluripotent stem cells, affects *in vitro* differentiation propensities of these cells.

Although the contribution of transcriptional noise at the single cell level to the pluripotent state has yet to be determined, it is clear that in the ideal experimental setting, iPS cells can be generated that are globally indistinguishable from ES cells. Thus, given that the diversity of cell sources will increase as iPS cells become more widely used in biomedical research, improved derivation methods are needed to ensure that the end product is as consistent as possible, and facile but comprehensive standards for defining the pluripotent state must be accepted by the community.

Epigenetic variability

Variations in expression levels between iPS cells and ES cells can be due to epigenetic and genetic factors, and these may constitute latent differences in cell potency that manifest only upon differentiation. Epigenetic factors may contribute to differences among pluripotent stem cells and may partially explain the experimental contexts in which iPS cells exhibit wider variation in gene expression as compared to ES cells. DNA methylation and

chromatin modifications as well as the unique permissive chromatin state of pluripotent stem cells have been reviewed extensively elsewhere^{88,89} and will not be discussed here. Epigenetic differences have been most comprehensively investigated by comparing ES cells and iPS cells. Most of these studies have explored the possibility of iPS cells maintaining an epigenetic memory, that is, iPS cells having an epigenetic state reminiscent of the donor cell type because of incomplete or aberrant reprogramming of the donor genome. Studies of DNA methylation in mouse iPS cells⁹⁰ and human iPS cells⁹¹ provided evidence of such an epigenetic memory, which correlated with distinct *in vitro* differentiation propensities: iPS cells derived from blood had a greater *in vitro* blood-forming capacity than fibroblast-derived iPS cells and neural progenitor-derived iPS cells. These differences were partially due to activating DNA methylation marks at the *Wnt3a* locus, which were present in blood-forming mouse iPS cells but not in cell lines that were less efficient at forming blood. Moreover, the addition of WNT3A to cell lines that did not differentiate into blood cells restored their blood forming potential, suggesting a disruption of lineage commitment signalling networks in these cell lines. These findings are largely consistent with another study, in which blood-derived mouse iPS cells had significantly higher *in vitro* blood-forming potential than mouse iPS cells derived from fibroblasts or smooth muscle⁹², and this correlated with differences in global gene expression and DNA methylation profiles. Moreover, fibroblast-derived mouse iPS cells formed significantly smaller embryoid bodies, indicating a disruption of differentiation programmes that normally repress the pluripotency network. However, these functional and molecular differences between mouse iPS cell lines diminished with continued passage. A study examining epigenetic memory in human iPS cells revealed residual H3 acetylation at β -cell-specific genes in β -cell-derived iPS cells but not in other pluripotent stem cells, and DNA methylation patterns were similar between β -cells and β -cell-derived iPS cells but not similar to iPS cells derived from other cell types. Moreover, differentiated cells from β -cell-derived iPS cells showed increased expression of pancreatic genes compared with cells originating from other pluripotent stem cells⁹³. A comparison of the DNA methylation state of 13 genomes from human iPS cells, the corresponding donor cells, ES cells and iPS cell-derived trophoblast cells⁹⁴ uncovered that although ES cells and iPS cells are globally highly similar, approximately half of the 300–600 differentially methylated CG regions in ES cells compared with iPS cells reflected epigenetic memory, and the other half were aberrant epigenetic reprogramming events. Approximately 130 differentially methylated CG regions were common to all five iPS cell lines examined, representing ‘hotspots’ of aberrant epigenetic reprogramming. These results are consistent with targeted CpG bisulphite sequencing experiments in 17 human iPS lines from seven founder cell types and seven ES cell lines, which revealed that 20–56% of the regions that are differentially methylated between iPS and ES cells are residual from the donor cell (that is, the donor cell has a similar methylation state)⁷⁹. It should be noted that the proportion of differentially methylated CpG sites between iPS cells and ES cells is minor with 0.5–3% when examining multiple donor cell types⁷⁹ and <1% when limited to fibroblast-derived human iPS cells⁷⁷.

Together, the above studies suggest that, first, both epigenetic memory and aberrant epigenetic states can be consequences of reprogramming, which therefore is not always robust. Second, epigenetic memory or aberrant epigenetic states can translate into limited *in vitro* differentiation capacity and may persist even after directed differentiation to certain lineages. Third, aberrant epigenomes may be ameliorated or corrected by drug treatment or extended passage in culture. Finally, the potential functional consequences of epigenetic variability in pluripotent stem cells can be mitigated by modifying differentiation protocols. In particular, experimental contexts, the technical limitations and infidelity of reprogramming can result in definable distinctions among different sources of iPS cells and between iPS cells and ES cells. Moreover, the genetic effects of reprogramming transgenes (for example, the insertion sites or the order in polycistrons) may cause non-optimal factor

stoichiometry that has functional consequences as exemplified by the epigenetic silencing of the *Dlk1-Dio3* locus in mouse iPS cells^{95,96}. Thus, although in an ideal generic case iPS cells and ES cells are not readily distinguishable in molecular terms, in any specific experimental context they may behave quite distinctly.

Variations in genetic background

The effect of genetic variants on *in vitro* differentiation is probably amplified due to fewer redundant signalling pathways being activated in the dish compared with signalling during development. Therefore, beyond concerns about genetic mutations acquired during the selection process and extended culture entailed in pluripotent stem cell isolation (either from embryos or through reprogramming), it seems plausible that ‘normal’ population variation contributes substantially to the variability of *in vitro* differentiation. In fact, in one study the donor genetic background accounted for more of the functional differences between human iPS cells than did the donor cell type or derivation method, supporting the role of genetic background differences in pluripotent stem cell function⁹⁷.

Experimental introduction of genetic alterations

Recently, there has been intense scrutiny of DNA alterations that accompany reprogramming. A survey of DNA copy number variation (CNV) burden in 106 human ES cells and iPS cells using single nucleotide polymorphism (SNP) arrays reported no overall association between passage number and CNV frequency. Frequent duplications of parts of chromosomes 12 and 20 in both ES cells and iPS cells were found, and an increased mean number of amplifications and heterozygous deletions in iPS cells compared with non-pluripotent stem cells (9 versus 6 amplifications and 3.4 versus 1.9 deletions in iPS cells and non-pluripotent stem cells, respectively) were reported⁹⁸. Although some genes that are deleted in iPS cells are associated with tumour suppression and some genes that are amplified in these cells are associated with proliferation or oncogenesis, it is unclear whether the overlaps are significant or would be expected by chance. Another study used SNP arrays to identify DNA copy number variants in 22 iPS cell lines, three donor fibroblasts and 17 ES cell lines to determine whether human pluripotent stem cells select for common CNVs⁹⁹. They found that reprogrammed populations at very early passages have extensive genetic mosaicism, which read out as an increased CNV load, and that continued passage reduced CNV frequency, which could be due to a selection process. Another study used array comparative genomic hybridization (aCGH) to detect CNVs in 25 human iPS cells and their parental fibroblasts¹⁰⁰. As expected, some CNVs in iPS cells were indistinguishable from CNVs in the parental fibroblasts. However, some CNVs in the iPS cells were unique to the iPS cells. These CNVs could have been generated *de novo* as part of the reprogramming process, or could have been present in a minor, yet undetectable, population of the donor fibroblasts. Moreover, this study also reported three recurrent CNVs in independent iPS cell lines and, consistent with previous studies, that these CNVs were lost during continued passage.

Recently, next-generation sequencing has enabled high-resolution identification of mutations in pluripotent stem cells. One study described the exome sequences of 22 human iPS cells and their donor fibroblasts and discovered approximately six mutations per exome (half of which are putatively function-altering) and an enrichment of mutations in cancer-associated genes⁸⁸. However, whole genome sequencing of three human iPS cell lines that have been generated by non-integrating episomal vectors did not show an enrichment of single nucleotide variants (SNVs) in cancer-associated genes¹⁰¹, nor were any CNVs or SNVs shared between these cell lines. A comparison of the exomes of five human iPS cells with parental fibroblast cell lines estimated that the mutation rate is ninefold higher in iPS

cells¹⁰². However, whether mutations in iPS cells result from the reprogramming process or from clonal selection and expansion is still unclear, but there is mounting evidence that heterogeneity in the founder populations has a significant role. Whole-genome sequencing in nine mouse iPS cell lines and their donor cells revealed that most SNVs were detectable as rare alleles in the founder population and that the hundreds of genome-wide SNVs and coding SNVs were not enriched in any pathway (cancer-related or otherwise)¹⁰³. Similarly, careful sequencing analysis of *de novo* CNV incidences in 20 human iPS cell lines and founder fibroblasts showed that at least 50% of the apparent human iPS cell-specific CNVs were actually present in the founder fibroblasts at low frequency¹⁰⁴. Finally, the donor cell type does not seem to have a substantial effect on the number of SNVs in protein-coding genes in human iPS cells¹⁰⁵.

As a decade ago, when the first gene expression profiles of stem cells were published, the field is still in the early stages of characterizing mutations in pluripotent stem cells. Some of the pressing issues that are emerging include whether the isolation and propagation of pluripotent stem cells select for consistent genetic changes, whether such changes have oncogenic potential and which mutations affect pluripotent stem cell function. Due to the differences inherent to array-based and sequencing technologies, and the distinct areas of focus of each study, definitive answers to these issues await more comprehensive studies. However, we can summarize this first generation of high-density array and sequence-based investigations of pluripotent stem cells as follows. First, the act of reprogramming itself does not seem to induce a high rate of mutations in common genes or pathways. However, when examining large numbers of CNV profiles, there are at least some recurrent variations in pluripotent stem cells that have been discovered in more than one study, including a minimal duplication of chromosome 20 that contains the locus for ID1 and other genes^{98,106}. Such duplications are likely to be a result of adaption to tissue culture. Second, there are approximately 5–10 coding mutations per iPS cell genome compared with donor cells, and most of these mutations already exist in the donor cell population and are captured and amplified by clonal selection. Third, the possibility that tumour-suppressors and oncogenes are not infrequently altered in iPS cells means that any potential therapeutic use of an iPS cell line will require exhaustive DNA screening.

Conclusions

Cellular reprogramming has opened up the possibility of generating patient-specific pluripotent cell lines that can be used as a source of material for cell replacement therapy, to investigate the molecular underpinnings of and to develop treatments for most human diseases. However, the development of robust differentiation protocols has been hindered by the functional variability of pluripotent stem cells. In the ideal case, generic iPS cells that have been faithfully reprogrammed seem indistinguishable from embryo-derived ES cells in molecular and functional terms. However, in certain experimental contexts, such as when multiple iPS cells are derived from specific tissue types and under suboptimal reprogramming conditions, epigenetic and genetic variations can influence the use of these cells in disease modelling and represent safety concerns for clinical applications. We suggest that future work addressing the issue of functional variability and heterogeneity should follow two complimentary paths. First, we should identify molecular markers, most likely epigenetic markers, that are indicative of the differentiation capacity for specific lineages. It is possible that these markers are associated with the interference of differentiation programmes or signalling networks that access differentiation programmes, and, as such are silent in the undifferentiated state. Second, inspired by the observations that much functional variability can be ameliorated by altering the differentiation conditions, we anticipate that the application of high-throughput approaches to develop directed differentiation conditions will yield protocols that are more robust for pluripotent stem cells.

Acknowledgments

G.Q.D. is supported by grants from the US National Institutes of Health (NIH) (UO1-HL100001 Progenitor cell biology consortium, R24-DK092760, P50HG005550 and special funds from the ARRA stimulus package- RC2-HL102815, RC4DK 090913), the Roche Foundation for Anemia Research, Alex's Lemonade Stand Foundation, Doris Duke Charitable Foundation and the Ellison Medical Foundation. G.Q.D. is an affiliate member of the Broad Institute and an investigator of the Manton Center for Orphan Disease Research and the Howard Hughes Medical Institute. P.C. is supported by grants T32HL007623 and 2T32HL66987-11 from the National Heart, Lung, and Blood Institute (NHLBI). The authors would like to thank R. Zhao and A. De Los Angeles for helpful discussions.

References

1. Kahan BW, Ephrussi B. Developmental potentialities of clonal *in vitro* cultures of mouse testicular teratoma. *J. Natl Cancer Inst.* 1970; 44:1015–1036. [PubMed: 5514468]
2. Evans MJ. The isolation and properties of a clonal tissue culture strain of pluripotent mouse teratoma cells. *Development.* 1972; 28:163–176.
3. 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] A landmark paper that demonstrates transcription factor-based reprogramming to pluripotent-like cells. It initiated a race to reprogram human fibroblasts to pluripotency, leading to widespread accessibility to pluripotent stem cells.
4. Young RA. Control of the embryonic stem cell state. *Cell.* 2011; 144:940–954. [PubMed: 21414485]
5. Nichols J, Smith A. Pluripotency in the embryo and in culture. *Cold Spring Harb. Persp. Biol.* 2012; 4:a008128.
6. Solter D. From teratocarcinomas to embryonic stem cells and beyond: a history of embryonic stem cell research. *Nature Rev. Genet.* 2006; 7:319–327. [PubMed: 16534514]
7. Wu SM, Hochedlinger K. Harnessing the potential of induced pluripotent stem cells for regenerative medicine. *Nature Cell Biol.* 2011; 13:497–505. [PubMed: 21540845]
8. Cohen DE, Melton D. Turning straw into gold: directing cell fate for regenerative medicine. *Nature Rev. Genet.* 2011; 12:243–252. [PubMed: 21386864]
9. Zhu H, Lensch MW, Cahan P, Daley GQ. Investigating monogenic and complex diseases with pluripotent stem cells. *Nature Rev. Genet.* 2011; 12:266–275. [PubMed: 21386866]
10. Bellin M, Marchetto MC, Gage FH, Mummery CL. Induced pluripotent stem cells: the new patient? *Nature Rev. Mol. Cell Biol.* 2012; 13:713–726. [PubMed: 23034453]
11. Cherry ABC, Daley GQ. Reprogramming cellular identity for regenerative medicine. *Cell.* 2012; 148:1110–1122. [PubMed: 22424223]
12. Robinton DA, Daley GQ. The promise of induced pluripotent stem cells in research and therapy. *Nature.* 2012; 481:295–305. [PubMed: 22258608]
13. Keller G. Embryonic stem cell differentiation: emergence of a new era in biology and medicine. *Genes Dev.* 2005; 19:1129–1155. [PubMed: 15905405]
14. Niwa H. Mouse ES cell culture system as a model of development. *Dev. Growth Differ.* 2010; 52:275–283. [PubMed: 20148924]
15. Grskovic M, Javaherian A, Strulovici B, Daley GQ. Induced pluripotent stem cells — opportunities for disease modelling and drug discovery. *Nature Rev. Drug Discov.* 2011; 10:915–929. [PubMed: 22076509]
16. Banito A, Gil J. Induced pluripotent stem cells and senescence: learning the biology to improve the technology. *EMBO Rep.* 2010; 11:353–359. [PubMed: 20379220]
17. Bernstein BE, et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell.* 2006; 125:315–326. [PubMed: 16630819]
18. Mikkelsen TS, et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature.* 2007; 448:553–560. [PubMed: 17603471]
19. Boyer LA, et al. Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell.* 2005; 122:947–956. [PubMed: 16153702]

20. Loh Y-H, et al. The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nature Genet.* 2006; 38:431–440. [PubMed: 16518401]
21. Thomson M, et al. Pluripotency factors in embryonic stem cells regulate differentiation into germ layers. *Cell.* 2011; 145:875–889. [PubMed: 21663792]
22. Teo AKK, et al. Pluripotency factors regulate definitive endoderm specification through eomesodermin. *Genes Dev.* 2011; 25:238–250. [PubMed: 21245162]
23. Yuan H, Corbi N, Basilico C, Dailey L. Developmental-specific activity of the FGF-4 enhancer requires the synergistic action of Sox2 and Oct-3. *Genes Dev.* 1995; 9:2635–2645. [PubMed: 7590241]
24. Marks H, et al. The transcriptional and epigenomic foundations of ground state pluripotency. *Cell.* 2012; 149:590–604. [PubMed: 22541430]
25. Kunath T, et al. FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. *Development.* 2007; 134:2895–2902. [PubMed: 17660198]
26. Niwa H, Burdon T, Chambers I, Smith A. Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. *Genes Dev.* 1998; 12:2048–2060. [PubMed: 9649508]
27. Ying QL, Nichols J, Chambers I, Smith A. BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. *Cell.* 2003; 115:281–292. [PubMed: 14636556]
28. Kim J, Chu J, Shen X, Wang J, Orkin SH. An extended transcriptional network for pluripotency of embryonic stem cells. *Cell.* 2008; 132:1049–1061. [PubMed: 18358816]
29. Lowell S, Benchoua A, Heavey B, Smith AG. Notch promotes neural lineage entry by pluripotent embryonic stem cells. *PLoS Biol.* 2006; 4:e121. [PubMed: 16594731]
30. Lindsley RC, Gill JG, Kyba M, Murphy TL, Murphy KM. Canonical Wnt signaling is required for development of embryonic stem cell-derived mesoderm. *Development.* 2006; 133:3787–3796. [PubMed: 16943279]
31. Ogawa K, Nishinakamura R, Iwamatsu Y, Shimosato D, Niwa H. Synergistic action of Wnt and LIF in maintaining pluripotency of mouse ES cells. *Biochem. Biophys. Res. Commun.* 2006; 343:159–166. [PubMed: 16530170]
32. Berge, ten D, et al. Embryonic stem cells require Wnt proteins to prevent differentiation to epiblast stem cells. *Nature Cell Biol.* 2011; 13:1070–1075. [PubMed: 21841791]
33. Yi F, et al. Opposing effects of Tcf3 and Tcf1 control Wnt stimulation of embryonic stem cell self-renewal. *Nature Cell Biol.* 2011; 13:762–770. [PubMed: 21685894]
34. Wray J, et al. Inhibition of glycogen synthase kinase-3 alleviates Tcf3 repression of the pluripotency network and increases embryonic stem cell resistance to differentiation. *Nature Cell Biol.* 2011; 13:838–845. [PubMed: 21685889]
35. Marson A, et al. Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. *Cell.* 2008; 134:521–533. [PubMed: 18692474]
36. Sinkkonen L, et al. MicroRNAs control *de novo* DNA methylation through regulation of transcriptional repressors in mouse embryonic stem cells. *Nature Struct. Mol. Biol.* 2008; 15:259–267. [PubMed: 18311153]
37. Delaloy C, et al. microRNA-9 coordinates proliferation and migration of human embryonic stem cell-derived neural progenitors. *Cell Stem Cell.* 2010; 6:323–335. [PubMed: 20362537]
38. Wang Y, Medvid R, Melton C, Jaenisch R, Blelloch R. DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. *Nature Genet.* 2007; 39:380–385. [PubMed: 17259983]
39. Kanellopoulou C, et al. Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. *Genes Dev.* 2005; 19:489–501. [PubMed: 15713842]
40. Kaji K, et al. The NuRD component Mbd3 is required for pluripotency of embryonic stem cells. *Nature Cell Biol.* 2006; 8:285–292. [PubMed: 16462733]
41. Zhu D, Fang J, Li Y, Zhang J. Mbd3, a component of NuRD/Mi-2 complex, helps maintain pluripotency of mouse embryonic stem cells by repressing trophectoderm differentiation. *PLoS ONE.* 2009; 4:e7684. [PubMed: 19888462]

42. Berstine EG, Hooper ML, Grandchamp S, Ephrussi B. Alkaline phosphatase activity in mouse teratoma. *Proc. Natl Acad. Sci.* 1973; 70:3899–3903. [PubMed: 4521215]
43. Tesar PJ, et al. New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature.* 2007; 448:196–199. [PubMed: 17597760]
44. Brons IGM, et al. Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature.* 2007; 448:191–195. [PubMed: 17597762]
45. De Los Angeles A, Loh Y-H, Tesar PJ, Daley GQ. Accessing naïve human pluripotency. *Curr. Opin. Genet. Dev.* 2012; 22:272–282. [PubMed: 22463982]
46. Nichols J, Smith A. Naïve and primed pluripotent states. *Cell Stem Cell.* 2009; 4:487–492. [PubMed: 19497275]
47. Evans MJ, Kaufman MH. Establishment in culture of pluripotential cells from mouse embryos. *Nature.* 1981; 292:154–156. [PubMed: 7242681]
48. Osafune K, et al. Marked differences in differentiation propensity among human embryonic stem cell lines. *Nature Biotechnol.* 2008; 26:313–315. [PubMed: 18278034] The first paper to systematically compare the *in vitro* differentiation capacity of human ES cells. It demonstrates that despite the equivalence of human ES cells by teratoma formation, human ES cells can still exhibit great tendencies to differentiate into specific lineages *in vitro*.
49. Di Giorgio FP, Boulting GL, Bobrowicz S, Eggan KC. Human embryonic stem cell-derived motor neurons are sensitive to the toxic effect of glial cells carrying an ALS-causing mutation. *Cell Stem Cell.* 2008; 3:637–648. [PubMed: 19041780]
50. Hu B-Y, et al. Neural differentiation of human induced pluripotent stem cells follows developmental principles but with variable potency. *Proc. Natl Acad. Sci. USA.* 2010; 107:4335–4340. [PubMed: 20160098]
51. Boulting GL, et al. A functionally characterized test set of human induced pluripotent stem cells. *Nature Biotechnol.* 2011; 29:279–286. [PubMed: 21293464]
52. Grigoriadis AE, et al. Directed differentiation of hematopoietic precursors and functional osteoclasts from human ES and iPS cells. *Blood.* 2010; 115:2769–2776. [PubMed: 20065292]
53. Chang K-H, et al. Diverse hematopoietic potentials of five human embryonic stem cell lines. *Exp. Cell Res.* 2008; 314:2930–2940. [PubMed: 18692044]
54. Kelly DL, Rizzino A. DNA microarray analyses of genes regulated during the differentiation of embryonic stem cells. *Mol. Reprod. Dev.* 2000; 56:113–123. [PubMed: 10813842]
55. Loring JF, Porter JG, Seilhammer J, Kaser MR, Wesselschmidt R. A gene expression profile of embryonic stem cells and embryonic stem cell-derived neurons. *Restor. Neurol. Neurosci.* 2001; 18:81–88. [PubMed: 11847430]
56. Ramalho-Santos M. ‘Stemness’: transcriptional profiling of embryonic and adult stem cells. *Science.* 2002; 298:597–600. [PubMed: 12228720]
57. Ivanova NB, et al. A stem cell molecular signature. *Science.* 2002; 298:601–604. [PubMed: 12228721]
58. Fortunel NO, et al. Comment on “‘Stemness’: transcriptional profiling of embryonic and adult stem cells” and “a stem cell molecular signature”. *Science.* 2003; 302:393. [PubMed: 14563990]
59. D’Amour KA, Gage FH. Genetic and functional differences between multipotent neural and pluripotent embryonic stem cells. *Proc. Natl Acad. Sci.* 2003; 100:11866–11872. [PubMed: 12923297]
60. Tanaka TS. Gene expression profiling of embryo-derived stem cells reveals candidate genes associated with pluripotency and lineage specificity. *Genome Res.* 2002; 12:1921–1928. [PubMed: 12466296]
61. Sato N, et al. Molecular signature of human embryonic stem cells and its comparison with the mouse. *Dev. Biol.* 2003; 260:404–413. [PubMed: 12921741]
62. Richards M, Tan SP, Tan JH, Chan WK, Bongso A. The transcriptome profile of human embryonic stem cells as defined by SAGE. *Stem Cells.* 2004; 22:51–64. [PubMed: 14688391]
63. Sperger JM, et al. Gene expression patterns in human embryonic stem cells and human pluripotent germ cell tumors. *Proc. Natl Acad. Sci.* 2003; 100:13350–13355. [PubMed: 14595015]

64. Zeng X, et al. Properties of pluripotent human embryonic stem cells BG01 and BG02. *Stem Cells*. 2004; 22:292–312. [PubMed: 15153607]
65. Bhattacharya B. Gene expression in human embryonic stem cell lines: unique molecular signature. *Blood*. 2004; 103:2956–2964. [PubMed: 15070671]
66. Abeyta MJ. Unique gene expression signatures of independently-derived human embryonic stem cell lines. *Hum. Mol. Genet.* 2004; 13:601–608. [PubMed: 14749348]
67. Okita K, Ichisaka T, Yamanaka S. Generation of germline-competent induced pluripotent stem cells. *Nature*. 2007; 448:313–317. [PubMed: 17554338]
68. Wernig M, et al. *In vitro* reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature*. 2007; 448:318–324. [PubMed: 17554336]
69. Maherali N, et al. Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. *Cell Stem Cell*. 2007; 1:55–70. [PubMed: 18371336]
70. Park I-H, et al. Reprogramming of human somatic cells to pluripotency with defined factors. *Nature*. 2008; 451:141–146. [PubMed: 18157115]
71. Chin MH, et al. Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. *Cell Stem Cell*. 2009; 5:111–123. [PubMed: 19570518]
72. Marchetto MCN, et al. Transcriptional signature and memory retention of human-induced pluripotent stem cells. *PLoS ONE*. 2009; 4:e7076. [PubMed: 19763270]
73. Ghosh Z, et al. Persistent donor cell gene expression among human induced pluripotent stem cells contributes to differences with human embryonic stem cells. *PLoS ONE*. 2010; 5:e8975. [PubMed: 20126639]
74. Guenther MG, et al. Chromatin structure and gene expression programs of human embryonic and induced pluripotent stem cells. *Cell Stem Cell*. 2010; 7:249–257. [PubMed: 20682450] A careful comparison of the expression profiles of human ES cells and human iPS cells, arguing that no set of genes consistently distinguishes these two cell types.
75. Newman AM, Cooper JB. Lab-specific gene expression signatures in pluripotent stem cells. *Cell Stem Cell*. 2010; 7:258–262. [PubMed: 20682451]
76. Ohi Y, et al. Incomplete DNA methylation underlies a transcriptional memory of somatic cells in human iPS cells. *Nature Cell Biol.* 2011; 13:541–549. [PubMed: 21499256]
77. Bock C, et al. Reference maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. *Cell*. 2011; 144:439–452. [PubMed: 21295703]
78. Phanstiel DH, et al. Proteomic and phosphoproteomic comparison of human ES and iPS cells. *Nature Methods*. 2011; 8:821–827. [PubMed: 21983960]
79. Ruiz S, et al. Identification of a specific reprogramming-associated epigenetic signature in human induced pluripotent stem cells. *Proc. Natl Acad. Sci. USA*. 2012; 109:16196–16201. [PubMed: 22991473]
80. Kim H, et al. miR-371-3 expression predicts neural differentiation propensity in human pluripotent stem cells. *Cell Stem Cell*. 2011; 8:695–706. [PubMed: 21624813]
81. Chambers I, et al. Nanog safeguards pluripotency and mediates germline development. *Nature*. 2007; 450:1230–1234. [PubMed: 18097409]
82. Niwa H, Ogawa K, Shimosato D, Adachi K. A parallel circuit of LIF signalling pathways maintains pluripotency of mouse ES cells. *Nature*. 2009; 460:118–122. [PubMed: 19571885]
83. Toyooka Y, Shimosato D, Murakami K, Takahashi K, Niwa H. Identification and characterization of subpopulations in undifferentiated ES cell culture. *Development*. 2008; 135:909–918. [PubMed: 18263842]
84. Narsinh KH, et al. Single cell transcriptional profiling reveals heterogeneity of human induced pluripotent stem cells. *J. Clin. Invest.* 2011; 121:1217. [PubMed: 21317531]
85. Kalmar T, et al. Regulated fluctuations in Nanog expression mediate cell fate decisions in embryonic stem cells. *PLoS Biol.* 2009; 7:e1000149. [PubMed: 19582141]
86. MacArthur BD, et al. Nanog-dependent feedback loops regulate murine embryonic stem cell heterogeneity. *Nature Cell Biol.* 2012; 14:1139–1147. [PubMed: 23103910]

87. Yamaji M, et al. PRDM14 ensures naive pluripotency through dual regulation of signaling and epigenetic pathways in mouse embryonic stem cells. *Cell Stem Cell*. 2013; 12:368–382. [PubMed: 23333148]
88. Gore A, et al. Somatic coding mutations in human induced pluripotent stem cells. *Nature*. 2011; 470:63–67. [PubMed: 21368825]
89. Gaspar-Maia A, Alajem A, Meshorer E, Ramalho-Santos M. Open chromatin in pluripotency and reprogramming. *Nature Rev. Mol. Cell Biol.* 2011; 12:36–47. [PubMed: 21179060]
90. Kim K, et al. Epigenetic memory in induced pluripotent stem cells. *Nature*. 2010; 467:285–290. [PubMed: 20644535] One of the first and most comprehensive studies to investigate the extent to which DNA methylation patterns of diverse founder cell types are detectable in reprogrammed mouse iPS cells.
91. Kim K, et al. Donor cell type can influence the epigenome and differentiation potential of human induced pluripotent stem cells. *Nature Biotechnol.* 2011; 29:1117–1119. [PubMed: 22119740]
92. Polo JM, et al. Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. *Nature Biotechnol.* 2010; 28:848–855. [PubMed: 20644536]
93. Bar-Nur O, Russ HA, Efrat S, Benvenisty N. Epigenetic memory and preferential lineage-specific differentiation in induced pluripotent stem cells derived from human pancreatic islet β -cells. *Cell Stem Cell*. 2011; 9:17–23. [PubMed: 21726830]
94. Lister R, et al. Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature*. 2011; 471:68–73. [PubMed: 21289626]
95. Stadtfeld M, et al. Aberrant silencing of imprinted genes on chromosome 12qF1 in mouse induced pluripotent stem cells. *Nature*. 2010; 465:175–181. [PubMed: 20418860]
96. Carey BW, et al. Reprogramming factor stoichiometry influences the epigenetic state and biological properties of induced pluripotent stem cells. *Cell Stem Cell*. 2011; 9:588–598. [PubMed: 22136932]
97. Kajiwarra M, et al. Donor-dependent variations in hepatic differentiation from human-induced pluripotent stem cells. *Proc. Natl Acad. Sci. USA*. 2012; 109:12538–12543. [PubMed: 22802639]
98. Laurent LC, et al. Dynamic changes in the copy number of pluripotency and cell proliferation genes in human ESCs and iPSCs during reprogramming and time in culture. *Cell Stem Cell*. 2011; 8:106–118. [PubMed: 21211785]
99. Hussein SM, et al. Copy number variation and selection during reprogramming to pluripotency. *Nature*. 2011; 471:58–62. [PubMed: 21368824]
100. Martins-Taylor K, et al. Recurrent copy number variations in human induced pluripotent stem cells. *Nature Biotechnol.* 2011; 29:488–491. [PubMed: 21654665]
101. Cheng L, et al. Low incidence of DNA sequence variation in human induced pluripotent stem cells generated by nonintegrating plasmid expression. *Cell Stem Cell*. 2012; 10:337–344. [PubMed: 22385660]
102. Ji J, et al. Elevated coding mutation rate during the reprogramming of human somatic cells into induced pluripotent stem cells. *Stem Cells*. 2012; 30:435–440. [PubMed: 22162363]
103. Young MA, et al. Background mutations in parental cells account for most of the genetic heterogeneity of induced pluripotent stem cells. *Cell Stem Cell*. 2012; 10:570–582. [PubMed: 22542160]
104. Abyzov A, et al. Somatic copy number mosaicism in human skin revealed by induced pluripotent stem cells. *Nature*. 2012; 492:438–442. [PubMed: 23160490] A sequence-based mapping of DNA copy number differences between founder cells and reprogrammed human iPS cells. Rather than being induced by reprogramming, most CNVs are already present as rare alleles in the founder population.
105. Ruiz S, et al. Analysis of protein-coding mutations in hiPSCs and their possible role during somatic cell reprogramming. *Nature Commun.* 2013; 4:1382. [PubMed: 23340422]
106. International Stem Cell Initiative. Screening ethnically diverse human embryonic stem cells identifies a chromosome 20 minimal amplicon conferring growth advantage. *Nature Biotechnol.* 2011; 29:1132–1144. [PubMed: 22119741]
107. Stevens LC. The development of transplantable teratocarcinomas from intratesticular grafts of pre- and postimplantation mouse embryos. *Dev. Biol.* 1970; 21:364–382. [PubMed: 5436899]

108. Rosenthal MD, Wishnow RM, Sato GH. *In vitro* growth and differentiation of clonal populations of multipotential mouse cells derived from a transplantable testicular teratocarcinoma. *J. Natl Cancer Inst.* 1970; 44:1001–1014. [PubMed: 4936569]
109. Mintz B, Illmensee K. Normal genetically mosaic mice produced from malignant teratocarcinoma cells. *Proc. Natl Acad. Sci.* 1975; 72:3585–3589. [PubMed: 1059147]
110. Martin GR. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc. Natl Acad. Sci.* 1981; 78:7634–7638. [PubMed: 6950406]
111. Thomson JA. Embryonic stem cell lines derived from human blastocysts. *Science.* 1998; 282:1145–1147. [PubMed: 9804556]
112. Takahashi K, et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell.* 2007; 131:861–872. [PubMed: 18035408]
113. Park I-H, et al. Disease-specific induced pluripotent stem cells. *Cell.* 2008; 134:877–886. [PubMed: 18691744]
114. Sneddon JB, Borowiak M, Melton DA. Self-renewal of embryonic-stem-cell-derived progenitors by organ-matched mesenchyme. *Nature.* 2012; 491:765–768. [PubMed: 23041930]
115. Suda Y, Suzuki M, Ikawa Y, Aizawa S. Mouse embryonic stem cells exhibit indefinite proliferative potential. *J. Cell. Physiol.* 1987; 133:197–201. [PubMed: 3667706]
116. Nagy A, et al. Embryonic stem cells alone are able to support fetal development in the mouse. *Development.* 1990; 110:815–821. [PubMed: 2088722]

Box 1 | The rise of pluripotent stem cells in biomedical research

Embryonal carcinoma (EC) cells, derived from teratocarcinomas (formed spontaneously or by implanting 3-day old embryos into adult testis¹⁰⁷), were the first pluripotent stem cells to be isolated and investigated in culture^{1,108}. In these original studies, variability in subcloned EC cell lines had been noted as differing numbers of differentiated cell types detectable in teratomas^{1,2}. Despite the fact that EC cells are pluripotent, they are less-than-ideal systems to study development for several reasons: the intermediate *in vivo* incubation step obscures events that establish pluripotency; their resistance to differentiation may be due to mutational oncogenic events that make them different from embryonic pluripotent cells; and they form chimaeras upon blastocyst injections that only infrequently contribute to the germ line¹⁰⁹. Embryo-derived pluripotent cells were first isolated by explanting the preimplantation blastocyst in cell culture conditions that block differentiation^{47,110}. The isolation of embryonic stem (ES) cells from human embryos¹¹¹ launched a surge in research interest in using pluripotent stem cells as a potential source of material for cellular replacement therapy. The reprogramming of mouse somatic cells in 2006 (REF. 3) and human somatic cells in 2007 (REF. 112) into induced pluripotent stem (iPS) cells has greatly accelerated and broadened the interest in using pluripotent stem cells for drug screening, disease modelling¹¹³ and cell replacement. Induced pluripotency has evolved into a robust platform that is used in thousands of laboratories around the world, and a steadily increasing number of pluripotency-related studies have been published. Since 2010, the number of publications concerned with ES cells has decreased, whereas the number of publications concerning the applications of pluripotent stem cells has increased (see the figure).

Box 1 figure: Timeline of pluripotency research and associated publication volume

Box 2 | Self-renewal and pluripotency

Some cell types that are not pluripotent can nonetheless self-renew under the appropriate conditions. For example, haematopoietic stem cells (HSCs) undergo self-renewing division in the bone marrow, and extended self-renewal can be conferred on multipotent pancreatic progenitors in culture¹¹⁴. Self-renewal also implies immortality, the maintenance of pluripotency after many, presumably indefinite¹¹⁵, passages in culture. Just as some cells self-renew but are not pluripotent, some cells are pluripotent and yet do not extensively self-renew. For example, a zygote can give rise to all cells of the embryo proper and extra-embryonic tissues but undergoes only limited self-renewing cleavage divisions immediately after fertilization. The zygote is called totipotent because of its capacity to generate extraembryonic tissues (such as the trophoctoderm) in addition to the embryonic tissue. Pluripotent tissue arises in the inner cell mass and the epiblast of the embryo, where it manifests little self-renewal due to the rapid onset of gastrulation. Indeed, the state of immortal self-renewal of embryonic stem (ES) cells can be considered an artefact of cell culture. Nonetheless, the behaviour of embryo-derived ES cells remains the gold standard to which cells of undetermined potency are compared.

Box 3 | Assessing pluripotency

Assays to characterize pluripotent stem cells vary in time and cost and in the fidelity with which they measure the pluripotent state and its functional consequences. Pluripotent stem cells are assessed by morphology, cell surface phenotype, genome-wide expression profile, *in vitro* differentiation, teratoma formation, chimerism and tertraploid complementation.

Morphology

Mouse pluripotent stem cells are domed, refractive and tightly compacted such that individual cells are not discernible under low magnification. Human pluripotent stem cells cultured on mitotically inactivated mouse embryonic feeder cells have a flat, cobblestone morphology. During reprogramming, partially reprogrammed colonies can arise early, and these are unlikely to give rise to fully pluripotent stem cells. This indicates that morphology alone is insufficient to identify pluripotent stem cells.

Marker expression

Cell surface antigens used to identify pluripotent stem cells were originally identified in embryonal carcinoma cells and have been subsequently refined to the adhesion molecule SSEA1 (stage-specific embryonic antigen 1; also known as CD15) in mouse pluripotent stem cells and an antigen that is recognized by the TRA-1–60 antibody in human pluripotent stem cells.

Genome-wide expression profiling

Gene expression microarrays, and now RNA sequencing, are powerful technologies to identify genes that are differentially expressed in pluripotent stem cells compared with other cell types. Unsupervised analysis can be used to compare pluripotent stem cells at a global level to other pluripotent stem cells (for example, to identify similar subtypes) or to other cell types.

***In vitro* differentiation**

Upon withdrawal of leukaemia inhibitory factor (LIF) in mouse cell cultures and fibroblast growth factor (FGF) in human cell cultures, pluripotent stem cells spontaneously differentiate. Different techniques induce the formation of cystic structures that are known as embryoid bodies. These structures are evaluated by their immunological phenotype and the expression of cell type-specific marker for each germ layer.

Teratoma formation

Pluripotent stem cells injected into immune-compromised mice form teratoma-like structures that contain cell types from all three germ layers. Currently, this is the most stringent test of pluripotency available for human pluripotent stem cells.

Chimerism

Pluripotent stem cells injected into mouse blastocysts (diploid complementation) can contribute to all lineages of the embryo, including the germ line, with subsequent generation of viable offspring (see the figure, part a).

Tertraploid complementation

A stringent assay for determining cell potency whereby mouse pluripotent stem cells are introduced into tetraploid embryos; as tetraploid cells do not contribute extensively to embryonic development, the birth of mouse pups only occurs when pluripotent stem cells

are fully competent to contribute to all embryonic cell lineages^{7,116} (see the figure, part b).

Figure 1. Model of the mouse pluripotency network

A model representing our current understanding of the maintenance of the pluripotent state in mouse pluripotent stem cells. The transcription factors OCT4, SOX2 and NANOG (commonly referred to as OSN) form an autoregulated core that upregulates other pluripotency factors, as well as signalling pathways of ectoderm or mesendoderm fates and the pro-differentiation MAPK1 pathway. At the same time, OSN factors repress lineage-specific programmes (in concert with Polycomb group (PcG) chromatin regulators). When cultured in the presence of LIF (leukaemia inhibitory factor; which acts through STAT3 (signal transducer and activator of transcription 3)) and BMP4 (bone morphogenetic protein 4; which acts through ID (inhibitor of differentiation) and SMAD1), MAPK1-enabled differentiation is blocked in mouse embryonic stem (ES) cells. Perturbations to the pluripotency network are likely to have distinct functional consequences for *in vitro* differentiation depending on which part of the network is affected. Disruptions to the differentiation signalling cascade infrastructures (red boxes) are more likely to be apparent at the expression level in the pluripotent state, whereas disruptions to either lineage commitment and differentiation programmes or genetic programmes responsible for rapidly repressing the pluripotency network upon differentiation (blue boxes) are more likely to remain silent until directed to differentiate. FGF, fibroblast growth factor; MEP, mesendodermal progenitor; miR, micro RNA; EP, ectodermal progenitor; OLIG2, oligodendrocyte transcription factor 2; OTX1, orthodenticle homolog 1; RBL2, retinoblastoma-like 2; TBX3, T box 3; ZIC2, zinc-finger protein of the cerebellum 2.

Table 1
Sources of variability and heterogeneity in pluripotent stem cells

Variations in pluripotent stem cell lines

Type of variability	Occurrence			
	Among ESC	Among iPSC	Between ESC, iPSC	Heterogeneity within PSC line
Functional • in vitro differentiation	+	+	+	+
RNA • mRNA	-	+	+	+
Epigenetic • DNA methylation	?	+	+	?
Genetic • Germline (background) • Result of derivation • selection • induced	+	+	+	?
	?	+	na	?

ES cell, embryonic stem cell; iPSC cell, induced pluripotent stem cell; NA, not applicable.