

RESEARCH ARTICLE

# Culture Adaptation Alters Transcriptional Hierarchies among Single Human Embryonic Stem Cells Reflecting Altered Patterns of Differentiation

Paul J. Gokhale<sup>1</sup>\*, Janice K. Au-Young<sup>2</sup>\*, SriVidya Dadi<sup>2</sup>, David N. Keys<sup>2</sup>, Neil J. Harrison<sup>1</sup>, Mark Jones<sup>1</sup>, Shomit Soneji<sup>3</sup>\*, Tariq Enver<sup>3</sup>\*, Jon K. Sherlock<sup>2</sup>, Peter W. Andrews<sup>1</sup>\*

**1** Centre for Stem Cell Biology, Department of Biomedical Science, University of Sheffield, Sheffield, United Kingdom, **2** ThermoFisher, Foster City, California, United States of America, **3** Weatherall Institute of Molecular Medicine, University of Oxford, Oxford, United Kingdom

\* These authors contributed equally to this work.

‡a Current address: Division of Molecular Medicine and Gene Therapy, Lund University, Lund, Sweden

‡b Current address: UCL Cancer Institute, University College London, London, United Kingdom

\* [p.w.andrews@sheffield.ac.uk](mailto:p.w.andrews@sheffield.ac.uk)



**OPEN ACCESS**

**Citation:** Gokhale PJ, Au-Young JK, Dadi S, Keys DN, Harrison NJ, Jones M, et al. (2015) Culture Adaptation Alters Transcriptional Hierarchies among Single Human Embryonic Stem Cells Reflecting Altered Patterns of Differentiation. *PLoS ONE* 10(4): e0123467. doi:10.1371/journal.pone.0123467

**Academic Editor:** Kai Wang, University of Southern California, UNITED STATES

**Received:** July 14, 2014

**Accepted:** March 3, 2015

**Published:** April 14, 2015

**Copyright:** © 2015 Gokhale et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** The work described was funded by: Medical Research Council, UK (<http://www.mrc.ac.uk>), G0300268, PJG, NJH, MJ, PWA; Human Frontiers Science Program, RGP0001/2012, PJG, PWA; and European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 602423, TE, PWA. The above funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The Work

## Abstract

We have used single cell transcriptome analysis to re-examine the substates of early passage, karyotypically Normal, and late passage, karyotypically Abnormal ('Culture Adapted') human embryonic stem cells characterized by differential expression of the cell surface marker antigen, SSEA3. The results confirmed that culture adaptation is associated with alterations to the dynamics of the SSEA3(+) and SSEA3(-) substates of these cells, with SSEA3(-) Adapted cells remaining within the stem cell compartment whereas the SSEA3(-) Normal cells appear to have differentiated. However, the single cell data reveal that these substates are characterized by further heterogeneity that changes on culture adaptation. Notably the Adapted population includes cells with a transcriptome substate suggestive of a shift to a more naïve-like phenotype in contrast to the cells of the Normal population. Further, a subset of the Normal SSEA3(+) cells expresses genes typical of endoderm differentiation, despite also expressing the undifferentiated stem cell genes, *POU5F1* (*OCT4*) and *NANOG*, whereas such apparently lineage-primed cells are absent from the Adapted population. These results suggest that the selective growth advantage gained by genetically variant, culture adapted human embryonic stem cells may derive in part from a changed substate structure that influences their propensity for differentiation.

## Introduction

Populations of stem cells are subject to selection pressure for variants that favor self-renewal over the alternative fates of differentiation or cell death [1] [2]. Indeed, human embryonic stem

was a collaboration with Thermo Fisher (JKA-Y SVD DNK, JS) who provided reagents, helped design the study, performed experiments, collected and analysed data and helped prepare the manuscript.

**Competing Interests:** The authors Janice K. Au-Young, SriVidya Dadi, David N. Keys, and Jon K. Sherlock are employees of Thermo Fisher. This does not alter the authors' adherence to PLOS ONE policies on sharing data and materials.

(ES) cells often undergo culture adaptation after prolonged periods, showing altered growth characteristics, such as increased population doubling times and plating efficiencies [3, 4]. At the same time, they may acquire genetic or epigenetic changes [4–7]. The most commonly observed changes are manifest cytogenetically as non-random gains of chromosome fragments, notably involving chromosome 1, 12, 17 and 20, changes that also characterize embryonal carcinoma (EC) cells, the malignant counterparts of ES cells and the stem cells of teratocarcinomas [1, 7]. The non-random nature of the karyotypic changes, their association with pluripotent stem cells, and the rapid overgrowth of genetically variant cells over several passages [2], all argue for selection based on growth advantages derived from alterations to the genome. The commonality of the cytogenetic changes seen in human ES cells in culture and EC cells from teratocarcinomas suggests that culture adaptation and aneuploidy of ES cells *in vitro* could provide a paradigm for some aspects of tumor progression in cancers that involve a stem cell population.

The growth advantage of the culture adapted cells, which is apparent from their enhanced clonogenic capacity, could arise from a reduced tendency to undergo apoptosis [8], or to differentiate, or from an altered pattern of differentiation [9]. In a previous study comparing early passage, 'normal', and late passage, culture adapted human ES cells, we found that human ES cells within the stem cell compartment can exist in alternative substates defined by differential expression of the ES cell surface marker, SSEA3 [4]: in the early passage cultures, clonogenic cells were present primarily in the SSEA3(+) subset, whereas clonogenic cells were found in both the SSEA3(+) and SSEA3(-) subsets isolated from a culture adapted subline. These results, supported by microarray transcriptome data, suggested that during differentiation human ES cells first lose expression of SSEA3 and subsequently commit to differentiate, while culture adaptation raises a barrier that decreases the probability that cells go on and differentiate after losing SSEA3 expression. As a result, culture adaptation appears to 'trap' stem cells in the SSEA3(-) substate, allowing them to revert to an SSEA3(+) substate. Other studies have similarly suggested that pluripotent stem cells can exist in interconvertible substates, defined by other surface antigens [10] or expression of transcription factors like NANOG [11], STELLA [12], REX1 [13] and HEX [14].

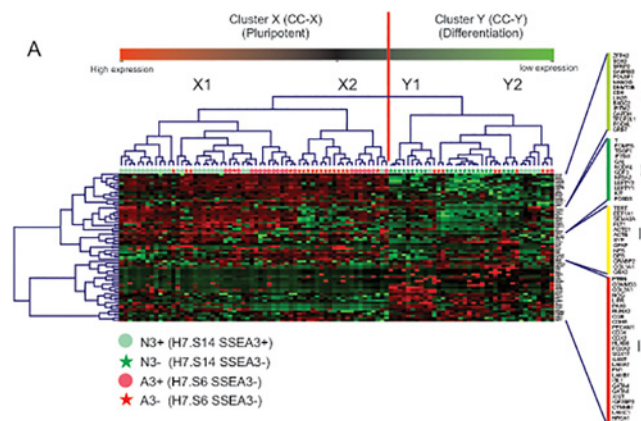
The existence of substates within the stem cell compartment raises the question of whether a stem cell in a particular substate might be biased towards certain lineages before commitment to differentiation [15]. Such biases have been suggested in a promyelocytic leukemic stem cell [16], in neural differentiation of a human EC cell line NTERA2 [17], and during hematopoietic differentiation of human ES cells [18]. The molecular basis for such biases remains largely unknown, although recently WNT signaling has been implicated in biasing human ES cell self-renewal and lineage potential [19]. Nevertheless, the phenomenon of multi-lineage priming whereby transcripts of genes pertinent to specific pathways of differentiation can be detected in individual, undifferentiated cells, has been long known in hematopoietic stem cells [20] [21] while, in naïve pluripotent murine stem cells, there is evidence of transcriptional pausing, indicative of many loci being primed for transcription [22]. Thus, one possibility is that the growth advantages of culture adapted ES cells can arise from alterations to the dynamics of the various lineage primed substates in which they may exist, in turn affecting their patterns of differentiation. To address this we have now re-examined the SSEA3(+) and (-) subsets that we previously identified in the early passage, normal, and late passage, culture adapted, sublines of H7 human ES cells [4], to determine whether at a single cell level there is evidence of lineage priming, and whether this is influenced by culture adaptation in a way that might contribute to a growth advantage for the variant cells.

## Results

SSEA3 positive and negative fractions of low passage karyotypically normal (Normal, N<sup>3+</sup> or N<sup>3-</sup>) and late passage, culture adapted ((47,XX,add(1)(p1)add(1)(q4),der(6)t(6;17)(q27;q11))

(Adapted, A<sup>3+</sup> and A<sup>3-</sup>) H7 human ES cells (sublines H7.s14 and H7.s6, respectively) were isolated by FACS. As before [4], the clonogenic cells in the Normal subline were predominantly in the N<sup>3+</sup> subset, being 30 fold less frequent in the N<sup>3-</sup> subset. In contrast, the cloning efficiency increased for the Adapted cells, and there were substantial numbers of clonogenic cells in both the A<sup>3+</sup> and A<sup>3-</sup> subsets (S1 Fig). Groups of 10, 100 and 1000 cells and 20–25 individual cells from each sorted subset—N<sup>3+</sup>, N<sup>3-</sup>, A<sup>3+</sup> and A<sup>3-</sup>—were then analyzed by qPCR for their expression of 96 genes, using the TaqMan Stem Cell Pluripotency Assays, (see Experimental Procedures). At the population level, the gene expression patterns of each of these four subsets of cells were comparable to those previously observed for these same subsets of H7 ES cells [4], whether based upon direct analyses of cell populations, or aggregation of single cell data (S1 Fig). In addition, as described in our previous study [4] XIST was downregulated, overall, in the H7.s6 adapted cells (S2 Fig) However, among the individual cells there was substantial heterogeneity in gene expression patterns, both between and within these SSEA3-defined subsets (S1 and S2 Data), indicating that the single cell analysis revealed features that are not evident from population level analyses.

To explore this further, we subjected the single cell data to unsupervised two-way clustering of cells and gene expression (Fig 1). Two broad cell clusters (CC) were apparent, labeled CC-X and CC-Y, each of which could be broken into two subclusters, X1 and X2, Y1 and Y2, respectively, while five distinct gene clusters (GC) were identified, GC-I, GC-II, GC-III, GC-IV and GC-V. GC-V was excluded from the heatmap displayed in Fig 1 and further analysis due to the very low to effectively zero expression of the genes in this cluster. In the Normal cells, CC-X and CC-Y approximated to the subsets defined by the presence or absence of SSEA3 expression, respectively, with 80% of the N<sup>3+</sup> cells in CC-X and 100% of the N<sup>3-</sup> cells in CC-Y (Fig 1). Inspection of the gene expression patterns associated with these cell clusters shows that CC-X was mostly linked to relatively high expression of genes in GC-I and GC-II, which include many of the genes characteristically associated with undifferentiated ES cells (e.g. *NANOG*, *POU5F1*, *SOX2*), but low expression of many genes in clusters GC-III, and particularly GC-IV, which include genes associated with various pathways of differentiation. These GC-III and -IV genes tended to be more highly expressed in CC-Y cells, which therefore most likely represent differentiated cells or cell committed to differentiation. There were, however, a small group of N<sup>3+</sup> cells who falls into the CC-Y cluster and who appeared to have differential expression of



**Fig 1. Unsupervised two-way cluster analysis of single cells analyzed by Taqman stem cell pluripotency assays.** Samples are plotted horizontally and Taqman probe sets used to interrogate the single cells are displayed vertically. The vertical red line separates the two major clusters identified in the dataset. Green on the heat map indicates high Delta Ct values i.e. low gene expression.

doi:10.1371/journal.pone.0123467.g001

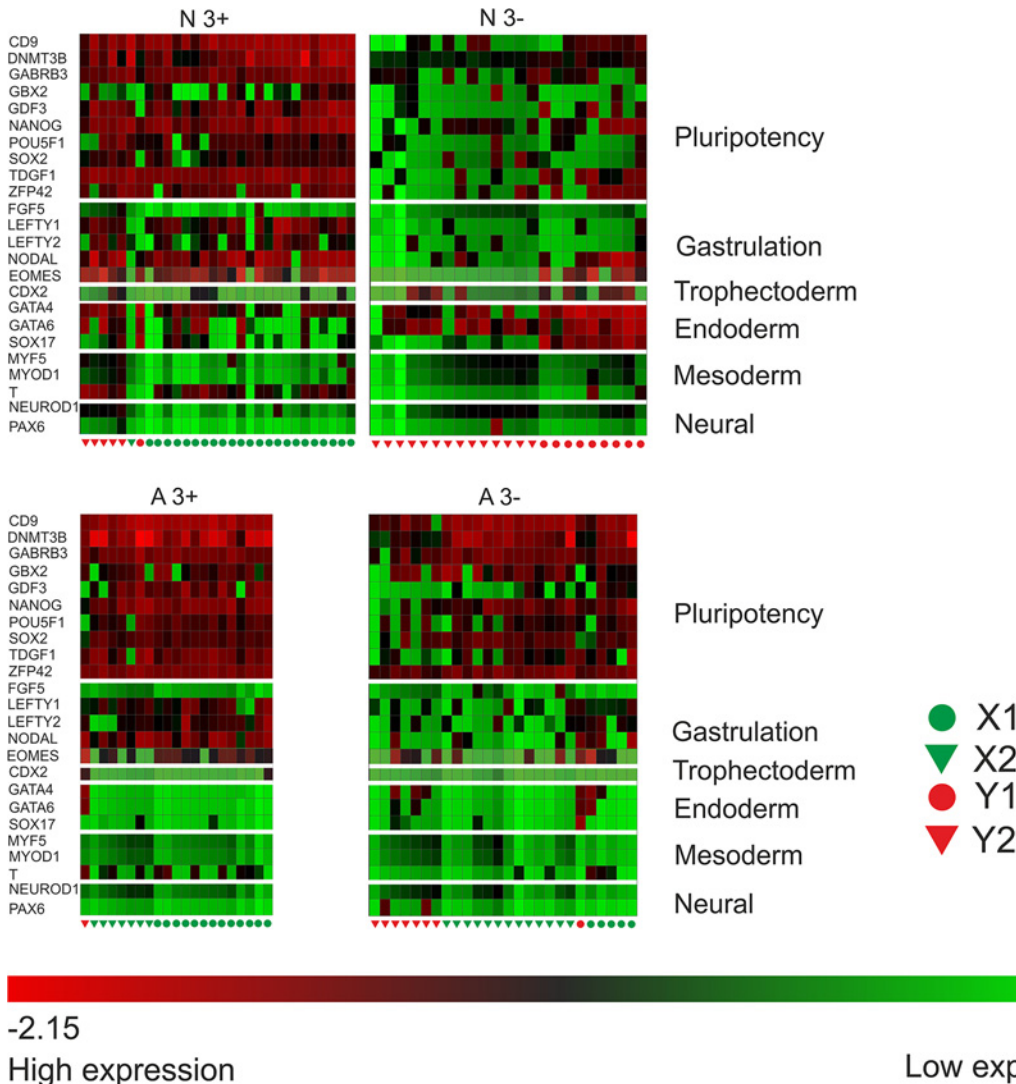
GC-II genes such as *T*, *EOMES* etc. and may represent  $N3^+$  cells closer to commitment than their counterparts in CC-X. The distinction between CC-X and CC-Y cells is also supported if genes well recognized to be associated with undifferentiated stem cells or various differentiated derivatives are considered (S3 Fig). On the other hand, the Adapted cells behaved differently, with CC-X being associated with cells either expressing or lacking SSEA3: 95% of the  $A^{3+}$  cells were in CC-X, while 69% of the  $A^{3-}$  subset also comprised CC-X cells (Fig 1). These results are consistent with our previous observation that, although loss of SSEA3 is associated with differentiation in Normal human ES cells, this relationship may be altered upon culture adaptation, with a significant proportion of Adapted cells losing SSEA3 expression but remaining in the undifferentiated stem cell compartment. Further, although the Normal CC-Y subset included many CC-Y1 and CC-Y2 cells, the Adapted CC-Y cells were almost all from the CC-Y2 sub-cluster, suggesting that culture adaptation influences not only the propensity of ES cells to differentiate, but also their pattern of differentiation.

Considering the two principal CC-X subclusters, CC-X2 was almost exclusively associated with the Adapted cells, which comprised 20 out of 21 CC-X2 cells analyzed (Fig 1). Among these Adapted cells, CC-X2 included both  $A^{3+}$  and  $A^{3-}$  cells, but was more associated with cells lacking SSEA3—62%  $A^{3-}$  versus 33%  $A^{3+}$ . The one remaining CC-X2 cell was from the  $N^{3+}$  fraction. By contrast, CC-X1 was more associated with the Normal cells and cells expressing SSEA3—56%  $N^{3+}$ , 32%  $A^{3+}$  but only 12%  $A^{3-}$  (Fig 1). Thus the single cell gene expression data identified novel substates that are related to, but not congruent with those identified by expression of SSEA3 and population level gene expression studies. These novel substates were also altered upon culture adaptation. In particular, CC-X2 tended to be associated with the SSEA3(-) cells that are found within the stem cell compartment and are much more common in cultures of Adapted cells.

CC-X1 and CC-X2 were mainly defined by differential expression of genes in GC-I and GC-II. In particular, GC-I genes tended to be expressed across both sub-clusters, whereas GC-II genes tended to be expressed at higher levels in CC-X1 and lower levels in CC-X2, though there remained substantial heterogeneity between individual cells (Fig 1). It is notable that GC-II includes several genes that are commonly associated with late epiblast and gastrulation in mammals, such as *BRACHYURY*, *EOMES*, *LEFTY1*, *LEFTY2*, *NR5A2* and *NODAL*, suggesting that CC-X2 cells may be less ‘epiblast-like’, or ‘primed’, than CC-X1 cells. A corollary is that culture adaptation has tended to push the population towards a more naïve (CC-X2) state.

Although this unsupervised analysis of the single cell gene expression data identifies obvious clusters of cells that relate to those defined by other criteria (e.g. surface antigen expression, clonogenicity and culture adaptation), implying some degree of co-ordination at the population level, there is marked heterogeneity of gene expression in individual cells suggesting that the co-ordination is far from complete. For example, among the trio of pluripotency-associated transcription factors, *POU5F1*, *OCT4*, *NANOG* and *SOX2*, cells expressing high levels of *POU5F1* but low levels of *NANOG* and/or *SOX2*, or vice versa can be found. The same observation holds when other genes are also considered, whether those typically associated with pluripotent stem cells, or those associated with different stages of embryonic development or different germ layers (Fig 2). Most notably genes typically associated with endoderm, mesoderm and ectoderm lineages were often expressed in the undifferentiated Normal cells—i.e. the CC-X  $N^{3+}$ —irrespective of their expression of individual genes associated with pluripotency or gastrulation, suggestive of multi-lineage priming. By contrast, this phenomenon was much less evident in the Adapted cells, whether  $A^{3+}$  or  $A^{3-}$ , again suggesting that adaptation tends to push the cells back from a more primed to a more naïve state.

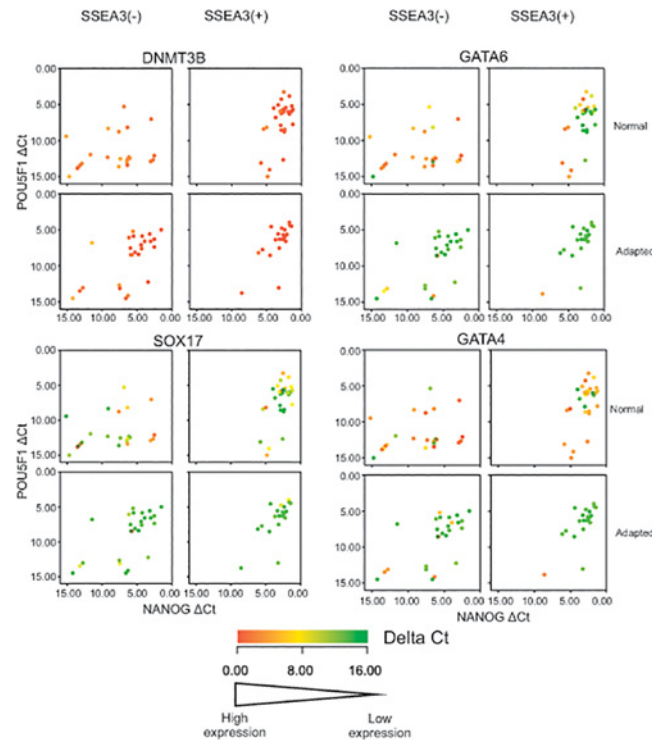
The contrast between the Normal and Adapted cells with respect to lineage associated genes was most striking in the case of genes associated with endoderm differentiation, notably *GATA4*, *GATA6* and *SOX17*. When the single cell data is represented in a scatter plot of



**Fig 2. Simultaneous gene expression in normal and abnormal SSEA3 + and - fractions.** Visualization of simultaneous gene expression in single Normal H7.s14 cells from the SSEA3 positive and negative fractions (N3<sup>+</sup>, N3<sup>-</sup> respectively) and Adapted H7.s6 cells from the SSEA3 positive and negative fractions (A3<sup>+</sup> A3<sup>-</sup> respectively) for selected genes from the Human Stem Cell Pluripotency assays associated with (i) the undifferentiated pluripotent state (*CD9*, *DNMT3B*, *GABRB3*, *GBX2*, *GDF3*, *NANOG*, *POU5F1*, *SOX2*, *TDGF1*, *ZFP42*), (ii) gastrulation (*FGF5*, *LEFTY1*, *LEFTY2*, *NODAL*, *EOMES*), (iii) trophoectoderm (*CDX2*), (iv) Endoderm (*GATA4*, *GATA6*, *SOX17*) (v) Mesoderm (*MYF5*, *MYOD1*, *BRACHYURY*) and (vi) Neural (*NEUROD1*, *PAX6*). The color of each square represents a single genes' expression (Delta Ct) in a particular cell. Green indicates high Delta Ct values i.e. low gene expression.

doi:10.1371/journal.pone.0123467.g002

expression of the two characteristic stem cell marker genes, *POU5F1* and *NANOG*, a clear cluster of cells expressing these genes is evident in the N3<sup>+</sup> and A3<sup>+</sup> populations, and persists in the A3<sup>-</sup> populations, but is lost in the N3<sup>-</sup> population (Fig 3). If it is posited that these *POU5F1*/*NANOG* expressing clusters represent the undifferentiated stem cells, the observations are consistent with the conclusions above and in our earlier study [4] that many A3<sup>-</sup> cells remain within the stem cell compartment whereas the N3<sup>-</sup> cells have initiated differentiation. However, when the expression of other genes are examined, it is clear that many of the *POU5F1*/*NANOG* expressing N3<sup>+</sup> cells also express one or more of the endoderm-related genes, *GATA4*, *GATA6* and *SOX17*. Again there was no evident co-ordinate expression of these genes. On the other hand they are almost not expressed in the A3<sup>+</sup> cells or the A3<sup>-</sup> cells, although many cells in the



**Fig 3. Two-way scatter plots of *POU5F1* (*OCT4*) expression versus *NANOG* expression in the SSEA3 positive and negative assayed fractions from the Normal and Adapted H7 human ES cells.** The genes plotted are *DNMT3B*, *GATA4*, *GATA6* and *SOX17*. Each cell is represented by an individual dot and colored according to the Delta Ct value of the genes indicated in the legend appended to each graph. Increasing Delta Ct indicates decreasing expression. Note that, in contrast to the relatively uniform expression of *DNMT3B* in all cells in each sorted population, the expression of *GATA4*, *GATA6* and *SOX17* is quite distinct in most cells of the SSEA3(+) Normal population but in almost no cells in the SSEA3(+) Adapted population.

doi:10.1371/journal.pone.0123467.g003

$N^3$  population express them. As we have previously suggested, one of the selective advantages that could drive adaptation is the loss of an ability to generate endoderm, which is a potential source of factors such as BMPs that may drive further differentiation of persisting undifferentiated stem cells [9]. The present results suggest that the loss of a propensity for endoderm reflects changes to the dynamics of substates within the stem cell compartment and the elimination of an endoderm primed substate.

## Discussion

The results of our present single cell transcriptome study are consistent with our earlier population based study of early passage, Normal, and late passage, Adapted, human ES cells, and confirm that the surface antigen SSEA3 defines substates, the dynamics of which are altered during culture adaptation [4]. Unlike the expression of many markers that have been used to identify substates of pluripotent stem cells, the SSEA3 epitope is not the product of a single gene, as it comprises a set of carbohydrate moieties of the glycosphingolipid, sialyl $\alpha$ 2-3gal $\beta$ 1-3globoside [23–25]. Further, whether cells are scored as SSEA3(+) or (-) appears to depend upon how that glycolipid is displayed rather than only its presence or absence [25]. Consequently, although we do not have a clear understanding of the molecular mechanism of SSEA3 positivity, we suggest that it provides a signal that integrates a wide array of transcriptomic and metabolomic information about cell state. Certainly various studies, including the current

results, indicate that the scoring of cells as SSEA3(+) and (-) does have functional correlates [26] [25] [27] [4] [28].

Broadly, the single cell clustering analysis indicated that all of the cells can be grouped into two clusters corresponding to undifferentiated stem cells, CC-X, which express genes classically associated with pluripotent stem cells, such as *ZFP42*, *LIN28*, *CD9*, *POU5F1*, *SOX2* and *NANOG*, and differentiated cells, CC-Y, which do not express significant levels of these genes but did express various genes associated with differentiation (S3 Fig). Whether from the cluster analysis (Fig 1), or by consideration of two key markers of undifferentiated pluripotent stem cells, *POU5F1* and *NANOG*, (Fig 3), the transcriptomes of the  $N^{3-}$  cells, all of which belong to Cell Cluster CC-Y, are consistent with them having left, or being in the process of leaving the stem cell compartment, as we argued before based on population level transcriptome data and on functional, clonogenicity assays [4]. By contrast, all the other subsets of Normal and Adapted cells contained undifferentiated, CC-X cells, though in a decreasing proportion from  $A^{3+}$  to  $N^{3+}$  to  $A^{3-}$ . However, these data also exposed marked heterogeneity within the cell clusters from which two observations stand out.

First, a subset of the cell cluster CC-X, CC-X2, was almost exclusively associated with the culture adapted cells. It has been widely proposed that human ES cells correspond to an epiblast like state, resembling murine EpiStem cells, which has also been called a primed state [29, 30] [31, 32], although several papers have reported culture conditions whereby human ES cells adopt some characteristics of more naïve, murine ES cells [33–37]. However, the heterogeneity between individual human ES cells in our current study suggests that populations of these cells may exist as distributions along a naïve-primed axis. Moreover, since the CC-X2 cells are primarily defined by a tendency to low expression of genes typically associated with the epiblast embryonic state, such as *BRACHYURY*, *EOMES*, *LEFTY1*, *LEFTY2*, *FOXD3*, this result is consistent with the notion that culture adaptation leads to the cells adopting a more naïve-like state. In addition, the adapted cells as in our previous study [4] had very low expression of *XIST*. Lack of X-inactivation has been described as a feature of naïve pluripotent stem cells (reviewed in [38]) and therefore would be anticipated in any naïve-like cells present in human PSC cultures. However, gains of an X chromosome are commonly reported in hES cell lines (see review in [1]; [7]) and may represent an independent change brought about by the cells being cultured under strongly selective conditions. Recently, the phenomenon of ‘erosion’ of X-inactivation in human pluripotent stem cells has been reported [39], again suggesting that changes in the inactivation of the X-chromosome may be due to selection rather than a shift to a naïve-like state.

Counterintuitively, despite the simple view that associates loss of SSEA3 with priming to differentiate [4], the CC-X2 cells tended to be SSEA3(-), although not exclusively so. On the other hand it is notable that mouse ICM cells, which correspond to the naïve condition of murine ES cells are also SSEA3(-). The observation that hPSCs exist in a range of states within the stem cells compartment, with some cells in more ‘naïve-like’ state concurs with recent single-cell data that utilized a different set of cell surface markers as discriminants and concluded that heterogeneity may be an inherent feature of hPSCs in culture [40].

Second, several of the  $N^{3+}$  cells expressed genes such as *GATA4*, *GATA6* and *SOX17*, suggesting that these cells are primed for endoderm differentiation, whereas cells expressing these genes were notably absent from the  $A^{3+}$  subset. Mutations in variant stem cells could provide growth advantages not only by affecting cell division or cell survival but also by altering their pattern of differentiation since derivative cells may form a niche that can influence the behavior of persisting stem cells. For example, Bendall et al [41] have reported that fibroblasts derived from human ES cells drive the further proliferation of the remaining human ES cells by producing IGF2 in response to exogenous FGF. We have previously reported that

culture adapted H7 human ES cells have a much reduced propensity for endoderm differentiation [9]. This may be a source of selective growth advantage since endoderm is known to produce BMP, a factor that induces human ES cell differentiation so that spontaneous endoderm differentiation can lead to a positive feedback loop that rapidly depletes the pool of undifferentiated cells [42]. Our present results suggest that the reduced tendency for endoderm differentiation of the culture adapted human ES cells arises from altered dynamics of substates of cells within the stem cell compartment prior to commitment to differentiation. Likewise we previously concluded in another system, neuronal differentiation of NTERA2 EC cells induced to differentiate with retinoic acid, that the eventual fate of the differentiated cells depended upon the behavior of lineage biased substates within the stem cell compartment prior to induction of differentiation [17].

This study highlights the increasing importance of analyzing cells at a single cell level rather than relying of average measures from population analyses. The results are consistent with the view that within the stem cell compartment human ES cells may transiently occupy different substates. Further, since cells in these substates express different patterns of gene expression, it seems likely that, when subjected to differentiation signals, they may exit the pluripotent state on different differentiation trajectories thus leading to heterogeneous differentiation [15]. Although the heterogeneity in expression of genes such as GATA4, GATA6 and SOX17 could be due in part to differential expression through the cell cycle [43], the lack of coordinate gene expression between lineage markers in single cells (Fig 2) suggests that the substates we have inferred reflect other attributes of cellular regulation. Understanding the mechanisms that control the dynamics of these stem cell substates therefore has implications for seeking culture conditions to minimize the appearance of genetically variant cells, by reducing the selective advantages associated with some biased differentiation potentials. Likewise it has implications for approaches to controlling the differentiation of ES cells to specific cell types for particular applications.

## Materials and Methods

### Human ES cell culture

Two sublines of the human ES cell line, H7 [44], were used in the current study [3] [4]. One subline, designated H7.s14, has a normal female karyotype (46, XX) and was used at a passage level less than 40. The second subline, designated H7.s6, has an abnormal karyotype (47,XX, add(1)(p1)add(1)(q4),der(6)t(6;17)(q27;q1)), and was used at a passage level in excess of p130. The cells were cultured on mitomycin C inactivated mouse embryonic fibroblasts derived from e12.5 MF-1 mouse embryos, in Knockout DMEM (Invitrogen), supplemented with 20% Knockout Serum Replacement (Invitrogen) and 4 ng/ml bFGF (PeproTech), as previously described [3]. The abnormal line, H7.s6, shows characteristics of culture adaptation, such as increased population growth rate and cloning efficiency, compared to the normal, H7.s14 line.

### Cell sorting and single cell isolation

Cells were harvested with trypsin:EDTA, stained by indirect immunofluorescence for the expression of the ES cell surface antigen, SSEA3, using pre-titered supernatant from hybridoma MC-631 [26] and FITC-conjugated goat anti-mouse IgG+IgM (Caltag), as previously described [4, 27]. Following staining, the cells were analyzed by flow cytometry, and sorted using a DakoCytometry Flow Cytometer. Specified numbers of cells from antigen-positive and antigen-negative subpopulations were deposited into the wells of a 96-well thermal cycling plate, in 0.5 µl of sorting buffer. Pre-sort and post-sort buffer samples were collected into the plates as negative controls.



## Cell Lysis and reverse transcription

Following sorting, the human ES cells were lysed using the TaqMan gene expression Cells-to-CT™ kit (Ambion Cat No. AM1728). Lysis solution (4.1μl) with DNase was added to single cells in 96-well plates and incubated for 5 min at room temperature. The reaction was terminated with 0.5μL of stop solution. Cell lysates were frozen at -80°C. Then reverse transcription reagents (RT) were added to the 96-well plate containing the single cell lysates. The RT reaction contained 40% lysate for the single cell samples and was performed at 37°C for 60 min followed by a 95°C heat inactivation step for 5 min. (See [S6 Fig](#) for a summary of the lysis, reverse transcription, pre-amplification and qPCR protocol)

## Multiplex preamplification

The resulting cDNA in a volume of 12.5μl was used for multiplex preamplification with a TaqMan PreAmp Pool for the Human Stem Cell Pluripotency assay panel, with 95 specific cDNA targets [[45](#)], Life Technologies Cat. No. 4414077). The preamplification was performed by incubating at 95°C for 10 min followed by 14 cycles at [95°C for 15 sec, 60°C for 4 min] in a Applied Biosystems GeneAmp PCR System 9700.

## QC of pre-amplified product

A preliminary Real-Time PCR experiment was performed to distinguish wells containing single cells from empty wells. The preamplified product was diluted (1:128) and mixed with TaqMan Gene Expression Master Mix (Life Technologies PN 4370074) to measure expression of 3 targets (Life Technologies *NANOG* assay ID Hs02387400\_g1, *ACTB* assay ID Hs99999903\_m1, and *GAPDH* assay ID Hs99999905\_m1). The Ct measurement was used to confirm that the well contained a cell. Samples with Ct values of *ACTB* and *GAPDH* that showed positive expression and good correlation were further tested.

## Quantitative Real-Time PCR

Preamplified product was diluted (1:128), and quantitative real-time PCR of 95 targets in single-plex using duplicates was performed. The targets used were described previously [[45](#)]. qPCR was performed with Human TaqMan Stem Cell Pluripotency assays (Life Technologies Cat. No. 4385344) in 384-well plates amplified in Applied Biosystems 7900HT Fast Real-Time PCR System. Raw cycle threshold values from each sample preparation for *GAPDH* and beta-ACTIN (*ACTB*) were used to identify data points lying greater than one standard deviation from the fraction population mean expression value. In sample preparations from single cells, those with higher or lower Ct values than one standard deviation of the population mean generally had correspondingly higher or lower Ct values for two other control genes *CTNNB1* and *RAF1*. These samples were excluded from further analysis as likely to contain damaged/ dying cells or multiple cells respectively. All subsequent gene expression data were expressed as the difference between Ct values for individual genes and the corresponding Ct value for *ACTB* obtained on the same cells ( $\Delta Ct_{ACTB}$ ).

Hierarchical clustering by genes and cells and heat mapping was created using Genesis [[46](#)] by Pearson correlation distance with complete linkage.

To test the feasibility of examining the gene expression of individual cells, single cells, and multiples of 10, 100, 1000 and 10,000 cells from the N3(+), N3(-), A3(+) and A3(-) subsets were deposited directly following FACS into the wells of a 96-well plate and then examined by TaqMan qPCR for *ACTB*, *GAPDH* and *NANOG*. The expression of *ACTB*, *GAPDH* and *NANOG* in preparations of 10,000, 1,000, 100, 10 and single cells from each subset showed a stepwise decrease

(i.e. increase in Ct value), with Ct values that were inversely proportional to the log cell number in each preparation (S4 Fig), indicating that the assay was sufficiently sensitive to read out changes in expression from single cells. In the case of the single cell preparations, samples were excluded from further analysis if their ACTB/GAPDH expression values fell outside one standard deviation of the mean (S5 Fig). Most of these outlier samples also expressed correspondingly high or low levels of the additional two control genes,  $\beta$ -catenin (*CTNNB1*) and *RAF1*, suggesting that they either contained multiple cells or cells undergoing apoptosis or necrotic death.

The overall gene expression patterns observed in the samples prepared from the different subsets of H7 cells were comparable to those found in our earlier transcriptome analysis of SSEA3-sorted subsets of the same sublines of H7 ES cells, analyzed by Affymetrix microarray [4]. When compared to our previous study, similar relative levels of gene expression were noted in the present study, e.g. for the pluripotent signature genes *NANOG*, *POU5F1*, *TGDF1*, *GABRB3*, *DNMT3B* and *GDF3* (S1 Fig). Taken together, these observations suggest that the expression data derived from the TLDA analysis provide a reliable picture of the transcriptome of the different subsets of H7 cells isolated in these experiments, broadly consistent with the results of previous population level based studies.

## Supporting Information

### S1 Data. Delta Ct values, normalized to beta-actin (ACTB).

(XLS)

### S2 Data. Raw Ct values.

(XLS)

**S1 Fig. Bulk comparison of Normal and Adapted SSEA3 sorted fractions. Panel A:** Single cell re-plating efficiencies of H7.s14 and H7.s6 SSEA3 positive and negative cells. **Panel B,** Comparison of selected pluripotency associated genes from Affymetrix gene microarray derived expression values from Enver et. al. [4] and the Delta Ct (normalized to beta-actin (ACTB)) expression values from the analysis of 1000 sorted cells run on TaqMan assays in this study. All values are expressed as normalized values to the H7.s14 SSEA3+ sample.

(EPS)

**S2 Fig. Two-way scatter plots of *POU5F1* (*OCT4*) expression versus *NANOG* expression against *XIST* in the SSEA3 positive and negative two assayed fractions from the Normal and Adapted H7 human ES cells.** Each cell is represented by an individual dot and colored according to the Delta Ct value of *XIST*. Increasing Delta Ct indicates decreasing expression.

(EPS)

**S3 Fig. Average gene expression between clusters X and Y.** The average gene expression (Delta Ct) of genes known to be generally associated with pluripotency of human ES cells (*CD9*, *DNMT3B*, *FGF4*, *GABRB3*, *GAL*, *GBX2*, *GRB7*, *LEFTY1*, *LEFTY2*, *LIFR*, *LIN28*, *NANOG*, *NODAL*, *NOGGIN*, *PODXL*, *POU5F1*, *SOX2*, *TGDF1*, *TERT*, *ZFP42*) compared to differentiation associated genes (*CDX2*, *DDX4*, *EOMES*, *FGF5*, *FOXA2*, *GATA4*, *GATA6*, *HLXB9*, *NEUROD1*, *PAX4*, *PAX6*, *SOX17*, *T*) between the two major clusters (X and Y) identified in Fig 1.

(EPS)

**S4 Fig. Evaluation of single cell TaqMan real-time PCR for sensitivity and outliers.** Gene expression values for ACTB (beta-actin), GAPDH on single cells, 100, 1000 and 10000 cells sorted for H7.s14 SSEA3+, H7.s14 SSEA3-, H7.s6 SSEA3+, H7.s6 SSEA3-. 'NTC' indicates RNA extracted from a given fraction but not transcribed to cDNA.

(EPS)

**S5 Fig. Two-way plots of GAPDH expression versus ACTB expression in the four assayed fractions compared to the control genes *CTNNB1* and *RAFI* from the human pluripotency low density array [45].** Each cell is represented by an individual dot and colored according to the Delta Ct value of the genes indicated in the legend appended to each graph. Increasing Delta Ct indicates decreasing expression. The shape of each point indicates whether a data point falls within one standard deviation of the population mean for each fraction.

(EPS)

**S6 Fig. Summary of the protocol used to analyze sorted single cells using TaqMan assays.**

(EPS)

**S1 Table. List of genes contained in each cluster identified in Fig 1A.** Genes colored green indicate an average expression value of < 10 Delta Ct(ACTB) in undifferentiated samples in the ISCI project dataset [45].

(DOCX)

## Acknowledgments

The authors thank members of the Centre for Stem Cell Biology, University of Sheffield for helpful discussions.

## Author Contributions

Conceived and designed the experiments: PJG JKA-Y DNK TE JS PWA. Performed the experiments: PJG JKA-Y SVD DNK NJH MJ. Analyzed the data: PJG JKA-Y DNK SS TE PWA. Wrote the paper: PJG JKA-Y DNK TE PWA.

## References

1. Baker DE, Harrison NJ, Maltby E, Smith K, Moore HD, Shaw PJ, et al. (2007) Adaptation to culture of human embryonic stem cells and oncogenesis in vivo. *Nat Biotechnol* 25: 207–215. PMID: [17287758](#)
2. Olariu V, Harrison NJ, Coca D, Gokhale PJ, Baker D, Billings S, et al. (2010) Modeling the evolution of culture-adapted human embryonic stem cells. *Stem Cell Res* 4: 50–56. doi: [10.1016/j.scr.2009.09.001](#) PMID: [19837641](#)
3. Draper JS, Moore HD, Ruban LN, Gokhale PJ, Andrews PW (2004) Culture and characterization of human embryonic stem cells. *Stem Cells Dev* 13: 325–336. PMID: [15345125](#)
4. Enver T, Soneji S, Joshi C, Brown J, Iborra F, Orntoft T, et al. (2005) Cellular differentiation hierarchies in normal and culture-adapted human embryonic stem cells. *Hum Mol Genet* 14: 3129–3140. PMID: [16159889](#)
5. Draper JS, Smith K, Gokhale P, Moore HD, Maltby E, Johnson J, et al. (2004) Recurrent gain of chromosomes 17q and 12 in cultured human embryonic stem cells. *Nat Biotechnol* 22: 53–54. PMID: [14661028](#)
6. Allegrucci C, Wu YZ, Thurston A, Denning CN, Priddle H, Mummery CL, et al. (2007) Restriction landmark genome scanning identifies culture-induced DNA methylation instability in the human embryonic stem cell epigenome. *Hum Mol Genet* 16: 1253–1268. PMID: [17409196](#)
7. Amps K, Andrews PW, Anyfantis G, Armstrong L, Avery S, Baharvand H, et al. (2011) Screening ethnically diverse human embryonic stem cells identifies a chromosome 20 minimal amplicon conferring growth advantage. *Nat Biotechnol* 29: 1132–1144. doi: [10.1038/nbt.2051](#) PMID: [22119741](#)
8. Avery S, Hirst AJ, Baker D, Lim CY, Alagaratnam S, Skotheim RI, et al. (2013) BCL-XL Mediates the Strong Selective Advantage of a 20q11.21 Amplification Commonly Found in Human Embryonic Stem Cell Cultures. *Stem cell reports* 1: 379–386. doi: [10.1016/j.stemcr.2013.10.005](#) PMID: [24286026](#)
9. Fazeli A, Liew CG, Matin MM, Elliott S, Jeanmeure LF, Wright PC, et al. (2011) Altered patterns of differentiation in karyotypically abnormal human embryonic stem cells. *Int J Dev Biol* 55: 175–180. doi: [10.1387/ijdb.103177af](#) PMID: [21671221](#)

10. Laslett AL, Grimmond S, Gardiner B, Stamp L, Lin A, Hawes SM, et al. (2007) Transcriptional analysis of early lineage commitment in human embryonic stem cells. *BMC developmental biology* 7: 12. PMID: [17335568](#)
11. Chambers I, Silva J, Colby D, Nichols J, Nijmeijer B, Robertson M, et al. (2007) Nanog safeguards pluripotency and mediates germline development. *Nature* 450: 1230–1234. PMID: [18097409](#)
12. Hayashi K, Lopes SM, Tang F, Surani MA (2008) Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states. *Cell stem cell* 3: 391–401. doi: [10.1016/j.stem.2008.07.027](#) PMID: [18940731](#)
13. Toyooka Y, Shimosato D, Murakami K, Takahashi K, Niwa H (2008) Identification and characterization of subpopulations in undifferentiated ES cell culture. *Development* 135: 909–918. doi: [10.1242/dev.017400](#) PMID: [18263842](#)
14. Canham MA, Sharov AA, Ko MS, Brickman JM (2010) Functional heterogeneity of embryonic stem cells revealed through translational amplification of an early endodermal transcript. *PLoS Biol* 8: e1000379. doi: [10.1371/journal.pbio.1000379](#) PMID: [20520791](#)
15. Enver T, Pera M, Peterson C, Andrews PW (2009) Stem cell states, fates, and the rules of attraction. *Cell stem cell* 4: 387–397. doi: [10.1016/j.stem.2009.04.011](#) PMID: [19427289](#)
16. Chang HH, Hemberg M, Barahona M, Ingber DE, Huang S (2008) Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. *Nature* 453: 544–547. doi: [10.1038/nature06965](#) PMID: [18497826](#)
17. Tonge PD, Olariu V, Coca D, Kadirkamanathan V, Burrell KE, Billings SA, et al. (2010) Prepatterning in the stem cell compartment. *PloS one* 5: e10901. doi: [10.1371/journal.pone.0010901](#) PMID: [20531938](#)
18. Hong SH, Rampalli S, Lee JB, McNicol J, Collins T, Draper JS, et al. (2011) Cell fate potential of human pluripotent stem cells is encoded by histone modifications. *Cell stem cell* 9: 24–36. doi: [10.1016/j.stem.2011.06.002](#) PMID: [21726831](#)
19. Blauwkamp TA, Nigam S, Ardehali R, Weissman IL, Nusse R (2012) Endogenous Wnt signalling in human embryonic stem cells generates an equilibrium of distinct lineage-specified progenitors. *Nat Commun* 3: 1070. doi: [10.1038/ncomms2064](#) PMID: [22990866](#)
20. Hu M, Krause D, Greaves M, Sharkis S, Dexter M, Heyworth C, et al. (1997) Multilineage gene expression precedes commitment in the hemopoietic system. *Genes Dev* 11: 774–785. PMID: [9087431](#)
21. Cheng T, Shen H, Giokas D, Gere J, Tenen DG, Scadden DT (1996) Temporal mapping of gene expression levels during the differentiation of individual primary hematopoietic cells. *Proceedings of the National Academy of Sciences of the United States of America* 93: 13158–13163. PMID: [8917561](#)
22. Marks H, Kalkan T, Menafra R, Denisov S, Jones K, Hofemeister H, et al. (2012) The transcriptional and epigenomic foundations of ground state pluripotency. *Cell* 149: 590–604. doi: [10.1016/j.cell.2012.03.026](#) PMID: [22541430](#)
23. Kannagi R, Levery SB, Ishigami F, Hakomori S, Shevinsky LH, Knowles BB, et al. (1983) New globo-series glycosphingolipids in human teratocarcinoma reactive with the monoclonal antibody directed to a developmentally regulated antigen, stage-specific embryonic antigen 3. *The Journal of biological chemistry* 258: 8934–8942. PMID: [6863318](#)
24. Kannagi R, Cochran NA, Ishigami F, Hakomori S, Andrews PW, Knowles BB, et al. (1983) Stage-specific embryonic antigens (SSEA-3 and -4) are epitopes of a unique globo-series ganglioside isolated from human teratocarcinoma cells. *Embo J* 2: 2355–2361. PMID: [6141938](#)
25. Fenderson BA, Andrews PW, Nudelman E, Clausen H, Hakomori S (1987) Glycolipid core structure switching from globo- to lacto- and ganglio-series during retinoic acid-induced differentiation of TERA-2-derived human embryonal carcinoma cells. *Dev Biol* 122: 21–34. PMID: [3297853](#)
26. Shevinsky LH, Knowles BB, Damjanov I, Solter D (1982) Monoclonal antibody to murine embryos defines a stage-specific embryonic antigen expressed on mouse embryos and human teratocarcinoma cells. *Cell* 30: 697–705. PMID: [6183004](#)
27. Draper JS, Pigott C, Thomson JA, Andrews PW (2002) Surface antigens of human embryonic stem cells: changes upon differentiation in culture. *J Anat* 200: 249–258. PMID: [12033729](#)
28. Tonge PD, Shigeta M, Schroeder T, Andrews PW (2011) Functionally defined substates within the human embryonic stem cell compartment. *Stem Cell Res* 7: 145–153. doi: [10.1016/j.scr.2011.04.006](#) PMID: [21763622](#)
29. Brons IG, Smithers LE, Trotter MW, Rugg-Gunn P, Sun B, Chuva de Sousa Lopes SM, et al. (2007) Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* 448: 191–195. PMID: [17597762](#)
30. Tesar PJ, Chenoweth JG, Brook FA, Davies TJ, Evans EP, Mack DL, et al. (2007) New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* 448: 196–199. PMID: [17597760](#)

31. Nichols J, Smith A (2009) Naive and primed pluripotent states. *Cell stem cell* 4: 487–492. doi: [10.1016/j.stem.2009.05.015](https://doi.org/10.1016/j.stem.2009.05.015) PMID: [19497275](https://pubmed.ncbi.nlm.nih.gov/19497275/)
32. Pera MF, Tam PP (2010) Extrinsic regulation of pluripotent stem cells. *Nature* 465: 713–720. doi: [10.1038/nature09228](https://doi.org/10.1038/nature09228) PMID: [20535200](https://pubmed.ncbi.nlm.nih.gov/20535200/)
33. Xu Y, Zhu X, Hahm HS, Wei W, Hao E, Hayek A, et al. (2010) Revealing a core signaling regulatory mechanism for pluripotent stem cell survival and self-renewal by small molecules. *Proceedings of the National Academy of Sciences of the United States of America*.
34. Buecker C, Chen HH, Polo JM, Daheron L, Bu L, Barakat TS, et al. (2010) A murine ESC-like state facilitates transgenesis and homologous recombination in human pluripotent stem cells. *Cell stem cell* 6: 535–546. doi: [10.1016/j.stem.2010.05.003](https://doi.org/10.1016/j.stem.2010.05.003) PMID: [20569691](https://pubmed.ncbi.nlm.nih.gov/20569691/)
35. Hanna J, Cheng AW, Saha K, Kim J, Lengner CJ, Soldner F, et al. (2010) Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. *Proceedings of the National Academy of Sciences of the United States of America* 107: 9222–9227. doi: [10.1073/pnas.1004584107](https://doi.org/10.1073/pnas.1004584107) PMID: [20442331](https://pubmed.ncbi.nlm.nih.gov/20442331/)
36. Gafni O, Weinberger L, Mansour AA, Manor YS, Chomsky E, Ben-Yosef D, et al. (2013) Derivation of novel human ground state naive pluripotent stem cells. *Nature* 504: 282–286. doi: [10.1038/nature12745](https://doi.org/10.1038/nature12745) PMID: [24172903](https://pubmed.ncbi.nlm.nih.gov/24172903/)
37. Ware CB, Nelson AM, Mecham B, Hesson J, Zhou W, Jonlin EC, et al. (2014) Derivation of naive human embryonic stem cells. *Proceedings of the National Academy of Sciences of the United States of America* 111: 4484–4489. doi: [10.1073/pnas.1319738111](https://doi.org/10.1073/pnas.1319738111) PMID: [24623855](https://pubmed.ncbi.nlm.nih.gov/24623855/)
38. Welling M, Geijsen N (2013) Uncovering the true identity of naive pluripotent stem cells. *Trends in cell biology* 23: 442–448. doi: [10.1016/j.tcb.2013.04.004](https://doi.org/10.1016/j.tcb.2013.04.004) PMID: [23685019](https://pubmed.ncbi.nlm.nih.gov/23685019/)
39. Mekhoubad S, Bock C, de Boer AS, Kiskinis E, Meissner A, Eggan K (2012) Erosion of dosage compensation impacts human iPSC disease modeling. *Cell stem cell* 10: 595–609. doi: [10.1016/j.stem.2012.02.014](https://doi.org/10.1016/j.stem.2012.02.014) PMID: [22560080](https://pubmed.ncbi.nlm.nih.gov/22560080/)
40. Hough SR, Thornton M, Mason E, Mar JC, Wells CA, Pera MF (2014) Single-cell gene expression profiles define self-renewing, pluripotent, and lineage primed States of human pluripotent stem cells. *Stem cell reports* 2: 881–895. doi: [10.1016/j.stemcr.2014.04.014](https://doi.org/10.1016/j.stemcr.2014.04.014) PMID: [24936473](https://pubmed.ncbi.nlm.nih.gov/24936473/)
41. Bendall SC, Stewart MH, Menendez P, George D, Vijayaragavan K, Werbowetski-Ogilvie T, et al. (2007) IGF and FGF cooperatively establish the regulatory stem cell niche of pluripotent human cells in vitro. *Nature* 448: 1015–1021. PMID: [17625568](https://pubmed.ncbi.nlm.nih.gov/17625568/)
42. Pera MF, Andrade J, Houssami S, Reubinoff B, Trounson A, Stanley EG, et al. (2004) Regulation of human embryonic stem cell differentiation by BMP-2 and its antagonist noggin. *J Cell Sci* 117: 1269–1280. PMID: [14996946](https://pubmed.ncbi.nlm.nih.gov/14996946/)
43. Singh AM, Chappell J, Trost R, Lin L, Wang T, Tang J, et al. (2013) Cell-cycle control of developmentally regulated transcription factors accounts for heterogeneity in human pluripotent cells. *Stem cell reports* 1: 532–544. doi: [10.1016/j.stemcr.2013.10.009](https://doi.org/10.1016/j.stemcr.2013.10.009) PMID: [24371808](https://pubmed.ncbi.nlm.nih.gov/24371808/)
44. Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, Marshall VS, et al. (1998) Embryonic stem cell lines derived from human blastocysts. *Science* 282: 1145–1147. PMID: [9804556](https://pubmed.ncbi.nlm.nih.gov/9804556/)
45. Adewumi O, Aflatoonian B, Ahrlund-Richter L, Amit M, Andrews PW, Beighton G, et al. (2007) Characterization of human embryonic stem cell lines by the International Stem Cell Initiative. *Nat Biotechnol* 25: 803–816. PMID: [17572666](https://pubmed.ncbi.nlm.nih.gov/17572666/)
46. Sturm A, Quackenbush J, Trajanoski Z (2002) Genesis: cluster analysis of microarray data. *Bioinformatics* 18: 207–208. PMID: [11836235](https://pubmed.ncbi.nlm.nih.gov/11836235/)