Published in final edited form as:

Mamm Genome. 2020 December ; 31(9-12): 263–286. doi:10.1007/s00335-020-09849-x.

Derivation of stable embryonic stem cell-like, but transcriptionally heterogenous, induced pluripotent stem cells from non-permissive mouse strains

Tiffany A. Garbutt[1], **Kranti Konganti**[2,3], **Thomas Konneker**[1], **Andrew Hillhouse**[2,3], **Drake Phelps**[1], **Alexis Jones**[1], **David Aylor**[1], **David Threadgill**[2,3,4,*]

[1]Program in Genetics, Department of Biological Science, North Carolina State University, Raleigh, NC 27695, USA

[2]Texas A&M Institute for Genome Sciences and Society, Texas A&M University, College Station, TX 77843, USA

[3]Department of Molecular and Cellular Medicine, Texas A&M University, College Station, TX 77843, USA

[4]Department of Biochemistry & Biophysics, Texas A&M University, College Station, TX 77843, USA

Abstract

Genetic background is known to play a role in the ability to derive pluripotent, embryonic stem cells (ESC), a trait referred to as permissiveness. Previously we have demonstrated that induced pluripotent stem cells (iPSC) can be readily derived from non-permissive mouse strains by addition of serum-based media supplemented with GSK3B and MEK inhibitors, termed 2iS media, three days into reprogramming. Here we describe the derivation of second type of iPSC colony from non-permissive mouse strains that can be stably maintained independently of 2iS media. The resulting cells display transcriptional heterogeneity similar to that observed in ESC from permissive genetic backgrounds derived in conventional serum containing media supplemented with leukemia inhibitor factor. However, unlike previous studies that report exclusive subpopulations, we observe both exclusive and simultaneous expression of naive and primed cell surface markers. Herein, we explore shifts in pluripotency in the presence of 2iS and characterize heterogenous subpopulations to determine their pluripotent state and role in heterogenous iPSCs derived from the non-permissive NOD/ShiLtJ strain. We conclude that heterogeneity is a naturally occurring, necessary quality of stem cells that allows for the maintenance of pluripotency. This study further demonstrates the efficacy of the 2iS reprogramming technique. It is also the first study to derive stable ESC-like stem cells from the non-permissive NOD/ShiLtJ and WSB/EiJ strains, enabling easier and broader research possibilities into pluripotency for these and similar non-permissive mouse strains and species.

[*]Contact: David Threadgill, Department of Molecular and Cellular Medicine, Texas A&M University, College Station, TX 77843 USA; Tel: 979-436-0850; Fax: 979-847-9481; dwt@tamu.edu.

Conflict of interest The authors declare no competing financial interests.

Introduction

Embryonic stem cells (ESC) are proliferative, pluripotent cells with the ability to differentiate into all cell types of the adult organism. Murine embryonic stem cell colonies were previously thought to be a homogenous population of transcriptionally naïve cells. However, transcriptional and proteomic heterogeneity in stem cell populations is now well established. Indeed, at the single cell level, conventional mouse ESC derived in serum containing media supplemented with leukemia inhibitory factor (LIF) homogenously express *Oct4* and *Sox2* (Zhang et al.), but are heterogeneous for a number of other pluripotency factors including *Nanog*, *Zfp42* (*Rex1*), Brachyury (*T*), and *Dppa3* (stella) (Chambers et al., 2007; Hayashi et al., 2008; Singh et al., 2007; Suzuki et al., 2006a, b; Tanaka, 2009; Toyooka et al., 2008; Vieyra et al., 2009). In contrast, ESCs derived in serum-free LIF conditions supplemented with glycogen synthetase kinase 3 beta (GSK3B) and mitogen-activated protein kinase (MEK) inhibitors, commonly referred to as 2i/LIF conditions are transcriptionally homogenous. However, culture with just one of either of these inhibitors results in transcriptionally distinct pluripotent states (Wray et al., 2010). This suggests a more nuanced view of pluripotency in which multiple *in vitro* pluripotent states with distinct transcriptional characteristics are possible (Hackett et al., 2017). Indeed, research has demonstrated the existence of a number of distinct pluripotent *in vitro* ESC states (Hackett and Surani, 2014)(Wu and Izpisua Belmonte, 2015).

A study by Hackett et. al. found that *in vitro* murine ESC pluripotency exists along a developmental continuum with distinct states resembling the inner cell mass (ICM) stage of the 3.5 days post coitum (dpc) blastocyst and varying intermediate states of pluripotency up to the primed epiblast lineage (Hackett et al., 2017). The existence of multiple pluripotent states along a developmental continuum may also be a naturally occurring characteristic of early development. I*n vivo*, ICM cells of the pre-implantation blastocyst cycle between primitive endoderm and primitive ectoderm states until cell state specification at the implantation blastocyst stage (Abranches et al., 2013). Both ESC and epiblast stem cells (EpiSC) can be derived from pre-implantation ICM cells when grown in culture conditions that do not restrict either developmental state (Najm et al., 2011). Post-implantation epiblast cells of the egg cylinder stage at 5.0 dpc have also been found to be transcriptionally heterogeneous before re-establishing a state of transcriptional homogeneity at 6.0 dpc (Han et al., 2010). Cell state heterogeneity is likely a reoccurring, transient, but a required developmental trait. Single cell heterogeneity may also be an important quality of stem cell populations that allow the capacity to differentiate, while maintaining a pluripotent cell state until subsequent differentiation signals (Galvin-Burgess et al., 2013; Graf and Stadtfeld, 2008; Smith, 2001).

Genetic background and exogenic inhibitors in *in vitro* culture of ESC contributes more to transcriptional variation in pluripotency states than other factors such as serum (Hackett et al., 2017). This is consistent with our previous findings that genetic background influences the ability to derive induced pluripotent stem cells from select strains of mice (Garbutt et al., 2018). Such mice are also non-permissive for embryonic stem cell derivation. The earliest developmental stage of pluripotency captured *in vitro* for these genetic backgrounds is the EpiSC stage derived from the developmentally primed post-implantation epiblast

(Brons et al., 2007; Tesar et al., 2007). In non-mouse species, such as humans, ESC derived from the inner cell mass of the blastocyst are morphologically, functionally, and epigenetically distinct from mouse ESC and are more similar to the developmentally primed post-implantation mouse EpiSC state (Boyer et al., 2005; Brons et al., 2007; Silva et al., 2008; Tesar et al., 2007; Thomson et al., 1998). Mouse EpiSC are characterized by flattened colony formation, high nuclear-to-cytoplasmic ratio, and increased propensity for differentiation. They can readily contribute to all three germ layers in teratoma assays but have somewhat limited pluripotency potential to contribute to the germline in chimeras. They are also intolerant to passage as single cells. On the other hand, ESC are characterized by three-dimensional dome shaped colony formation, pluripotency potential to contribute to chimera formation, including the germline, and tolerance to single cell passaging (Brons et al., 2007; Guo et al., 2009; Tesar et al., 2007).

GSK3B and MEK inhibitors collectively referred to as 2i, have been used to derive ESC from non-permissive mouse strains and other species (Buehr et al., 2008; Hanna et al., 2009; Nichols et al., 2009). 2i has also been used to complete the reprogramming process in partially reprogrammed cells (Silva et al., 2008). However, it is typically used in serum-free conditions referred to as 2i/LIF conditions (Hackett and Surani, 2014). The use of 2i in serum-free conditions early during the reprogramming process results in fibroblast cell death, as MEK is crucial for fibroblast viability (Feng et al., 2009; Roux and Blenis, 2004; Silva et al., 2008). Consequently, 2i is only used in fibroblast-derived iPSC after reprogramming to maintain cells in an ESC-like cell state.

Previously, we showed that homogenous ESC-like iPSC can be derived from the non-permissive NOD/ShiLtJ and WSB/EiJ mouse strains by continual use of serum-based 2i media (2iS) in culture (Garbutt et al., 2018). Here we expand upon our previous findings to derive a second type of iPSC colony from non-permissive mouse strains. We show that culture with 2iS for the first seven days during reprogramming facilitates derivation of 2i independent iPSC colonies from non-permissive strains that heterogeneously express the ESC cell surface marker platelet/endothelial cell adhesion molecule 1 (PECAM1) and the proposed EpiSC cell surface marker CD40 antigen at the single cell level. These colonies represent an additional pluripotent state from a non-permissive background that can be stably maintained in the absence of continual inhibitor culture. These iPSC model transcriptionally heterogeneous pluripotent state of conventionally derived serum/LIF iPSC and ultimately the transcriptionally heterogeneous state of the pre-implantation blastocyst. Contrary to previous reports that heterogeneous populations consist of distinct subpopulations (Chambers et al., 2007; Hayashi et al., 2008; Singh et al., 2007; Suzuki et al., 2006a, b; Tanaka, 2009; Toyooka et al., 2008; Vieyra et al., 2009), we observe exclusive and simultaneous expression of naïve and primed cell surface markers in heterogeneous iPSC colonies. To our knowledge only one other group has documented similarly naïve and EpiSC-like primed cell state transitions in mouse iPSCs (Shakiba et al., 2015). Herein, we characterize these subpopulations to determine their pluripotent state and role in heterogeneous iPSC derived from the NOD/ShiLtJ strain.

Results

Short Culture in 2iS (Short 2iS) reprograms cells to an early heterogeneous ESC-like state

Embryonic stem cell-like iPSC derived from non-permissive strains are metastable and are often dependent on their culture conditions to maintain an ESC-like identity. Indeed, ESC-like iPSC can be derived from the non-permissive NOD background by continual ectopic expression of *Myc* and *Klf4* (Hanna et al., 2009). Ectopic *Myc* and *Klf4* expression could be replaced with GSK3B and MEK inhibition after initial reprogramming to maintain cells in an ESC-like state; but the loss of either ectopic expression or 2i inhibition resulted in the loss of ESC-like cell identity and a reversal to the EpiSC-like state. In contrast, another study found that culture early during reprogramming with the histone deacetylase inhibitor VPA was successful for the derivation of stable iPSC from the non-permissive NOD background (Liu et al., 2011). This suggests that there is a critical time window during early reprogramming that can impact the success and quality of iPSC derivation.

Previously, we showed that 2i can be used early during the reprogramming of fibroblasts if used in combination with serum/LIF conditions, collectively referred to as 2iS (Garbutt et al., 2018). Here, culture in 2iS for only the first seven days of reprogramming facilitated the derivation of stable iPSC from non-permissive strains. We added 2iS three days after viral infection of fibroblasts and maintained infected cells in 2iS media for only the first seven days of reprogramming before switching to standard serum/LIF media (Fig. 1A).

Colonies from the non-permissive NOD/ShiLtJ and WSB/EiJ strains emerged on the 5^{th} day after 2iS addition and were three-dimensional in morphology (Fig. 1B and E). By the 7^{th} day after 2iS addition, there were numerous three-dimensional colonies from both backgrounds. Within one day of removing 2iS media (after 7 days in 2iS), we observed a loss of stability and a deconstruction of morphology for many colonies (Fig. 1C, D, F, and G). Overall, WSB/EiJ colonies appeared to retain colony morphology better than NOD/ShiLtJ after 2iS removal.

Following 2iS removal, four and eight iPSC colonies were derived from the NOD/ShiLtJ and WSB/EiJ backgrounds, respectively. This is a notably lower level of iPSC derivation efficiency compared to that observed with iPSC derivation from permissive mouse strains or non-permissive mouse strains maintained in 2iS culture (Garbutt et. al. 2017). At this stage we also observed intolerance to trypsin (Fig. 1H) and passaged the colonies as cell clumps manually or by using collagenase type IV. Intolerance to passage as single cells using trypsin is a characteristic of the EpiSC state (Brons et al., 2007; Tesar et al., 2007). The cells were also maintained on a mouse embryonic fibroblast (MEF) feeder layer. Of the surviving colonies, we observed a collection of both flat and round colonies that, when either colony morphology was selected and expanded, gave rise to both flat and round colonies (Fig. 1I–L). We observed a gradual transition with increased passage from a population consisting of both flat and round colonies to a population of mostly round three-dimensional colonies on a MEF feeder layer (Fig. 1M and N). This is consistent with the idea that iPSC colonies become more competent with increased culture (Golipour et al., 2012). It is also consistent with the finding that ESC can be derived from post-implantation epiblasts following extended culture on feeder cells and in serum-based LIF media (Bao

et al., 2009). Altogether this suggests that non-permissive iPSC may transition through an EpiSC-like state to acquire ESC-like pluripotency. For the WSB/EiJ background, this transition occurred around passage eight, while for the NOD/ShiLtJ background, it occurred a passage ten.

Colonies from both NOD/ShiLtJ and WSB/EiJ were intolerant to growth on a feeder-free gelatin layer at all stages regardless of increased passage and transitioned to flattened colony morphology, characterized by finger-like projections that were indicative of early differentiation (Fig. 1O and P). This morphology was exacerbated to a more spindle-like appearance when the colonies were passaged with trypsin to a gelatin layer (Fig. 1Q). The three-dimensional dome shaped morphology of NOD/ShiLtJ and WSB/EiJ on MEFs suggests that they may share similarities to the ESC state. Prolonged growth on a feeder-free layer of gelatin is known to predispose ESC to differentiation. iPSC from non-permissive strains show signs of early differentiation within just a day of passage onto gelatin, which suggests that they may be more prone to differentiation than permissive iPSC or true ESC on a gelatin layer.

2iS maintains a metastable homogenous state

To identify cell state, we utilized two documented cell surface markers for the ESC and EpiSC states, PECAM1 and CD40, respectively (Rugg-Gunn et al., 2012). We grew ESCs and EpiSCs from the NOD background (Fig. 2A and B) and confirmed the ability of these markers to denote ESCs and EpiSCs (Fig. 2C and D). NOD/ShiLtJ and WSB/EiJ iPSC that were cultured in 2iS for only the first seven days (Fig. 2E) labeled heterogeneously for PECAM1 and CD40 (Fig. 2F and G).

The GSK3B and MEK inhibitors have been reported to maintain a metastable ESC-like state in non-permissive strains. As such, we investigated if a return to 2iS culture can facilitate a more homogenous ESC-like cell state in these iPSC. Previous studies indicated that epigenetic differences among iPSC is reduced at passage 16 or greater (Polo et al., 2010), and increased passage is associated with increased competency and pluripotent stability (Golipour et al., 2012). However, chromosomal abnormalities are also found in higher passaged stem cells (Sugawara et al., 2006; Ueda et al., 1995). Given the premise that higher passaged cells are more established and epigenetically stable, we passaged NOD/ShiLtJ and WSB/EiJ iPSC for 16 passages in serum/LiF media and then cultured them in 2iS media for eight passages (Fig. 2E). The cells were karyotyped and colonies with chromosomal abnormalities were excluded from subsequent analysis. When these iPSC colonies were again cultured in 2iS, they labeled homogenously for PECAM1 (Fig. 2H and I).

Considering the observed metastability of these iPSC, we sought to investigate the stability of our previously derived NOD/ShiLtJ and WSB/EiJ iPSC that were continually maintained in 2iS media (called iPSC.2iS) (Garbutt et al., 2018). We cultured iPSC.2iS cells for 16 passages in 2iS media, and then cultured them for eight passages in serum/LIF media without out the GSK3B and MEK inhibitors (Fig. 2J). Anecdotal research suggests that maintenance in 2i media may exacerbate the likelihood of chromosomal abnormality with prolonged passage (Czechanski et al., 2014; Hackett and Surani, 2014). To mitigate this, the cells were again karyotyped prior to analysis to remove those lines with

chromosome abnormalities. Similarly, colonies that were continuously cultured in 2iS labeled homogenously for PECAM1 (Fig. 2K and L), but labeled heterogeneously for PECAM1 and CD40 if they were removed and cultured without 2iS (Fig. 2M and N).

Interestingly, ESC derived from the 129S1/SvImJ strain and maintained in serum/LIF media also labeled heterogeneously for PECAM1 and CD40 (Fig. 3A). We surmise that NOD/ShiLtJ and WSB/EiJ cells derived in 2iS for seven days are similar to the heterogeneous state of conventionally derived ESC in serum/LIF conditions. Heterogeneous expression is not uncommon in true ESC. Embryonic stem cells derived in serum conditions have been found to be heterogeneous for a number of pluripotency factors (Chambers et al., 2007; Hayashi et al., 2008; Singh et al., 2007; Suzuki et al., 2006a, b; Tanaka, 2009; Toyooka et al., 2008; Vieyra et al., 2009) and have been known to fluctuate between ESC and EpiSC states in culture. Heterogeneity also occurs *in vivo* in the inner cell mass cells of the 3.5 dpc pre-implantation blastocyst stage (Chazaud et al., 2006), the same cell population from which ESC are derived. This heterogeneity is thought to resolve by the 4.5 dpc implantation stage after separation of the primitive endoderm and primitive endoderm (Boroviak et al., 2014). Early heterogeneity is hypothesized to be necessary to impart the ability to respond to differentiation signals while still maintaining a state of undifferentiated pluripotency (Galvin-Burgess et al., 2013; Graf and Stadtfeld, 2008; Smith, 2001). The addition of GSK3B and MEK inhibitors has been found to promote a homogenous transcriptional state and maintain conventional serum/LIF derived ESC in a state of transcriptionally homogenous pluripotency (Hanna et al., 2009; Leitch et al., 2013; Marks et al., 2012; Yamaji et al., 2013; Ying et al., 2008); heterogeneous 129S1/SvImJ ESC derived in serum/LIF conditions transitioned to homogenous PECAM1 expression when grown in 2iS conditions (Fig. 3B). However, reprogramming of permissive strains seem to bypass this early developmental time point as iPSC derived and maintained in serum/LIF conditions from both the 129S1/SvImJ and C57BL/6J strains label homogenously for PECAM1 expression (Fig. 3C and D). PECAM1 continued to be homogenously expressed in 129S1/SvImJ iPSC regardless of maintenance in serum/LIF or 2iS conditions (Fig. 3E).

The addition of GSK3B and MEK inhibitors support single cell homogeneity despite simultaneous serum conditions and is consistent with the idea that that GSK3B and MEK inhibitors stabilize pluripotency by counteracting the differentiation signals found in standard serum-based culture conditions (Hackett and Surani, 2014). The heterogenous state that iPSC from nonpermissive backgrounds revert to, in the absence of 2iS, is similar to the heterogenous state of ESC derived from permissive genetic backgrounds. This suggests that 2iS reprogramming facilitates the derivation of iPSC from non-permissive backgrounds that in the absence of 2iS culture resemble the ESC state of permissive backgrounds.

Present culture in 2iS influences pluripotent state more than history of 2iS culture

To investigate the transcriptional differences between single cells from homogenous and heterogenous cell populations, we randomly captured and analyzed single cells from the four NOD iPSC 2iS culture conditions: (1) iPSCs derived with continual 2iS culture, hereafter referred to as the *Always* 2iS group; (2) iPSC from the *Always* population that were removed from 2iS culture after 16 passages, hereafter referred to as the *Early* 2iS group; (3)

iPSC that were derived and cultured in 2iS for only the first seven days, hereafter referred to as the *Never* 2iS group; and (4) iPSC from the *Never* group that were later cultured in 2iS, hereafter referred to as the *Late* 2iS group (Fig. 4A). Despite differences in the timing of 2iS culture, iPSC from the *Always* and *Late* 2iS groups labeled homogenously from PECAM1. Similarly, despite differences in the timing of 2iS removal, iPSC colonies from the *Early* and *Never* 2iS groups labeled heterogeneously for PECAM1 and CD40. Cells from each group were captured and analyzed using the C1 Fluidigm Single-Cell Fluidics system. A total of 39 single cells (13 single cells from the *Always* group, 12 single cells from the *Late* group, 11 single cells from the *Early* group, 11 single cells from the *Never* group) were analyzed using the Fluidigm Singular Analysis Toolset for R.

Cells from the heterogenous populations (*Early* and *Never* 2iS groups) clustered with cells from the homogenous populations (*Always* and *Late* 2iS groups clustered together (Fig. 4B). This suggests a transcriptional difference between cells from homogeneous populations maintained in 2iS and cells from heterogeneous populations not maintained in 2iS. Surprisingly a number of cells from the heterogenous populations formed a third distinct cluster.

This cluster of heterogeneous cells was defined by an increase in expression of a subset of 33 genes, 13 of which could be mapped to known genes (Fig. 4C). Of the 13 mapped genes, three, *Cthrc1*, *Rnu1-4*, and *Rn7sk*, have relevancy to documented functions in ESCs. The small nuclear RNA *Rnu1-4* functions as part of the spliceosome and plays a role in pre-mRNA processing during lineage differentiation of stem cells (Jeffries et al., 2010; O'Reilly et al., 2013). In addition to its spliceosome activity it regulates the expression of a number of target genes, including those affecting proliferation, through transcriptional elongation (O'Reilly et al., 2013). *Cthrc1* encodes a WNT cofactor protein that selectively activates the WNT/PCP pathway by stabilizing ligand-receptor interactions (Yamamoto et al., 2008) and is also a downstream target of BMP2 (Kimura et al., 2008), which is critical for preventing neuronal differentiation and the maintenance of an ESC identity (Pera et al., 2004; Ying et al., 2003). The majority (85%) of ESC related genes are transcribed by RNA polymerase II and RNA polymerase II is paused at 39% of those genes (Min et al., 2011). RN7SK, a non-coding snRNA, represses transcription of lineage specific genes with bivalent or active chromatin marks through the regulation of transcriptional RNA Pol II pausing, thereby regulating transcriptional poising. RN7SK mediated transcriptional pausing of lineage specific genes in ESCs occurs at a higher frequency in serum conditions than in serum free 2i/LIF conditions (Castelo-Branco et al., 2013). Altogether, the increased expression of these genes suggests that cells in this cluster may exhibit ESC properties of self-renewal and pluripotency, but may be more transcriptionally primed to respond to differentiation.

If the gene group defining this cluster is removed, all cells from the heterogenous populations (*Early* and *Never* 2iS) cluster together and separately from cells isolated from the homogenous populations (*Always* and *Never* 2iS) (Fig. 4D). Notably, this suggests that present culture in 2iS is more influential than individual variation in the timing of 2iS culture. For example, the two homogenous populations (*Always* and *Never* 2iS) were introduced to 2iS at different times. The *Always* population was continually cultured in

2iS, whereas 2iS was only recently introduced to the *Late* population. Similarly, cells that had only received 2iS for the first seven days (*Never*) clustered with cells that were recently removed from 2iS media (*Early*). This is likely due to the ability of 2iS to foster a homogeneous phenotype.

The gene expression associated with heterogenous cells was indicative of their removal from 2iS conditions, and included an increased expression of *Crabp1, S100A6, Tmem237, Lefty1, Plac8, Maghob, Spink3, and Prelid2*. CRABP1 mediates the non-canonical activation of the ERK1/2 differentiation pathway (Collins and Watt, 2008; Persaud et al., 2013), which is downregulated in 2iS conditions. The GSK3B inhibitor in 2iS media decreases the degradation of beta-catenin by GSK3B. Cells not presently maintained in 2iS media may have lower levels of beta-catenin and consequently WNT signaling as evidenced by the increased expression of S100A6, which participates in ubiquitin-mediated degradation of beta-catenin (Ning et al., 2012) and the increased expression of FABP5, which is negatively regulated by WNT/beta-catenin signaling (Collins and Watt, 2008). While WNT signaling in these cells may be reduced compared to cells maintained in 2iS, it is not ablated as evidenced by the increase in expression of the *Tmem237* gene that positively regulates WNT signaling (Huang et al., 2011) In support of this, we also see an increased expression of *Lefty1*. Activation of canonical WNT signaling leads to activation of NODAL signaling and consequently LEFTY1 in human ESCs, of which non-permissive iPSC are arguably similar (Besser, 2004). In the absence of 2i, BMP and NODAL signaling may regulate ESC heterogeneity (Galvin-Burgess et al., 2013). This is consistent with our observation that *Lefty1* has a higher expression in the heterogeneous population of cells, not presently maintained in 2iS. In heterogeneous ESC populations, higher *Lefty1* expression is associated with the subpopulation of cells that express lower levels of the pluripotency gene *Nanog* (Galvin-Burgess et al., 2013). LEFTY1 is also part of the TGFB super family and plays an important role in both human and mouse ESC maintenance (Besser, 2004; Park, 2011).

We also observed increased expression of genes that explain the qualities of these cells in culture. For example, intolerance of iPSCs derived from non-permissive strains that are not presently maintained in 2iS media to trypsin may be explained by an increased expression of the *Spink1* gene that produces a serine protease inhibitor with anti-trypsin activity (Huhtala et al., 1982). *Magohb*, which is upregulated in serum conditions (Zhao et al., 1998), also has increased expression in cells not presently maintained in 2iS, suggesting that 2iS may be influential in overcoming conflicting signals found in serum conditions.

Among cells from the homogenous populations we see increased expression of genes that reflect 2i activity. For example, these cells were associated with increased expression of *Dusp1*, which negatively regulates the ERK1/2 and MAPK differentiation pathways (Geetha et al., 2011; Patterson et al., 2009). This is consistent with the inhibition of the ERK1/2 and MAPK differentiation pathways by the MEK inhibitor. We also observe an increase in *Manba* expression. MANBA becomes activated in response to overexpression of the STAT2 pathway and has a GSK3B inhibitor-like effect on cells (Cinelli et al., 2008; Wang et al., 2010). This supports the inhibition of GSK3B by the CHIR inhibitor in 2iS and leads to an increase WNT signaling. *Il9* is also highly expressed in these cells. It is activated in part by TGFB and results in the activation of the STAT3 pathway that is involved in

ESC self- renewal (Bauer et al., 1998; Chen and Wang, 2014). Most convincingly, cells presently maintained in 2iS highly express the *Tfcp2l1* gene, which is hypothesized to be at the intersection of LIF and 2i mediated self-renewal. It is important for maintaining ESC identity and is a common target of the LIF/STAT3 pathway. It can reprogram post-implantation EpiSCs to naive pluripotent ESCs and its forced expression can even replace LIF or either of the 2i inhibitors to derive and maintain an ESC identity (Martello et al., 2012; Ye et al., 2013). A summary of genes up/down regulated in 2iS media in light of their function within the predicted stem cell maintenance pathway is shown in Figure 4E.

Altogether, these results suggest that the iPSC from both the homogenous and heterogenous iPSC populations exist in varying states of pluripotency as indicated by differences in gene expression. It also indicates the influence that 2i conditions have on regulating iPSC and pluripotency gene expression in general. Despite differences in when 2iS was first added *in vitro* (*Always* or *Late*), cells that are presently maintained in 2iS conditions and that label homogenously for PECAM1 cluster together. Similarly, despite differences in initial 2i removal *in vitro* (*Early* or *Never* 2iS), cells that are not currently maintained in 2iS conditions and label heterogeneously for PECAM1 and CD40 cluster together. Furthermore, genes associated with each cluster pattern are indicative of the role of 2i on shaping gene expression and pluripotency state.

Subpopulation characterization shows PECAM1 positive cells predominate

Next, we investigated the frequency of each subpopulation by sorting cells from the heterogeneous NOD/ShiLtJ population derived from shortened 2iS culture and the heterogeneous NOD/ShiLtJ population that was removed from 2iS after passage 16. The cells sorted into three categories; cells that solely expressed PECAM1, cells that solely expressed CD40, and cells that dually express both PECAM1 and CD40.

Multiple flow sorting runs were performed (Fig. 5A and B). Cells that solely expressed PECAM1 comprised 66-82% of live flow sorted cells, while cells that solely expressed CD40 comprised 3-4% of live flow sorted cells. Co-expression of both markers occurred in 1-3% of live flow sorted cells. Within flow sorting experiments there was little difference between the frequencies of subpopulations between the two heterogeneous samples, indicating similarity between the two heterogeneous samples and their subpopulations despite differences in the timing of 2iS culture (Fig. 5A and B).

CD40+ and dual expressing cells occur at a much lower frequency than PECAM1+ cells, which suggests that single CD40+ cells from heterogenous populations likely eluded random single cell capture. To investigate the state of the randomly captured single cells from the previous single cell experiment, flow sorted subpopulations of PECAM1+, CD40+, and dual expressing cells were RNA sequenced and compared using unbiased hierarchal clustering to the 47 randomly captured single cells.

The two populations of sorted CD40+ cells cluster together to form a separate cluster (Fig. 5C). The sorted PECAM1+ and dual expressing cells cluster together indicating more similarity between the two subpopulations than to the CD40+ subpopulation. All 47 single cells cluster together with the sorted PECAM1+ cells and the dual expressing

cells, confirming that no CD40+ single cells were randomly captured. Because of the close similarity between PECAM1+ cells and the dual expressing cells we cannot determine if any dual expressing single cells were randomly captured.

Transcriptional characterization of subpopulations reveals distinct identities

The cell surface marker PECAM1 has been reproducibly shown to indicate ESC-like pluripotency and the CD40 cell surface marker has been proposed to indicate an EpiSC-like cell identity (Rugg-Gunn et al., 2012). However, cell state is more complex and dynamic than any single marker classification. For this reason, we aimed to better characterize the RNAseq transcriptome profiles of each subpopulation by comparing them to both ESC and EpiSC.

PCA analysis revealed that iPSC from different source populations that labeled positive for PECAM1 were more similar to each other than to iPSC that labeled positive for CD40 from the same source population (Fig. 6A). Similarly, CD40 positive cells from different source populations are more similar to each other than to PECAM1 positive cells from the same source population. PECAM1 positive iPSC and dual expressing iPSC that label positive for both markers separate with ESC, whereas iPSC that label positive for CD40 separate with EpiSC along principle component one.

PECAM1+ iPSC and dual expressing iPSC, when compared to ESC, had fewer differentially expressed genes than when compared to EpiSC (Fig. 6B), indicating that both subpopulations were transcriptionally closer to the ESC state than to the EpiSC state. In contrast, CD40+ iPSC had roughly an equal number of differentially expressed genes when compared to the ESC and EpiSC states, indicating that these cells were equally dissimilar to both cell states, and may represent an alternative cell identity.

Next, we compared PECAM1+ cells to CD40+ cells to identify each subpopulation's unique gene profile through the identification of differentially expressed genes. The predicted functional classes of these genes was assessed using Ingenuity Pathways Analysis (Prasanth et al.). Relative to CD40+ cells, PECAM1+ cells had an increased number of genes related to pluripotency, maintaining stem cells, and ESC self-renewal (Table 1). Among those genes were the PECAM1 marker gene and genes that have been shown to be directly up-regulated in 2i/LIF conditions including *Prdm14* and *Tcl1* (Hackett and Surani, 2014). Overexpression of *Prdm14*, even in standard serum conditions, has been documented to convert heterogeneous subpopulations with low expression of the pluripotency marker *Rex1* to transcriptionally homogeneous populations (Grabole et al., 2013). Relative to CD40+ cells, PECAM1+ cells also had increased number of genes related to DNA methylation, including *Bhmt*, which has been shown to lead to increased DNA methylation in the mouse blastocyst (Zhang et al., 2015). Genes implicated in the development of the blastocyst and embryo implantation were also found predominantly in PECAM1 cells, in addition to genes that regulate cell number.

CD40+ cells relative to PECAM1+ cells had a number of genes that have been functionally documented to increase cell differentiation, including CD40 (Table 2). A number of genes that have been shown to increase cell proliferation and decrease cell death were also

predominately found in CD40+ cells. This suggests that while CD40+ stem cells may be predisposed to differentiation; they may also express mitigating genes that prevent cell death.

Cells that have dual expression of the PECAM1 and CD40 cell surface markers were compared to cells that express either of the two markers. Only a few genes were unique to dual expressing cells were observed. Relative to cells expressing either marker solely, dual expressing cells had decreased expression of *Nrp1* and *Cav1*. Interestingly, loss of NRP1 or CAV1 has been associated with cell quiescence (Cao et al., 2010; Cerezo et al., 2009). Subpopulations of adult stem cells have been known to exist in quiescent states (Li and Clevers, 2010). Quiescence, while not critical for stem cells, is important for maintaining balance in stem cell populations (Orford and Scadden, 2008). This data suggest the possibility that dual expressing cells may play a role in cell quiescence of heterogenous populations.

We predict that stem cell colonies need all three subpopulations to maintain equilibrium in the absence of 2iS. PECAM1 expressing cells are critical for producing pluripotency factors. CD40 positive cells are important for producing factors that prevent cell death. Dual expressing cells exist in a quiescent state that, while expressing PECAM1, may fluctuate between either the PECAM1 or CD40 positive state to maintain colony balance. If cells from either the PECAM1 positive subpopulation or the CD40 positive subpopulation are grown separately in culture, we see a re-emergence of the heterogeneous phenotype among single cells (Fig. 6C and D). This indicates that cell heterogeneity is not a result of contamination between different cultures, but rather a cell autonomous phenotype dependent on both the background genotype and culture environment.

FOXO1 and MYC are significantly different between PECAM1 and CD40 positive cells

Transcription factors are essential to stem cell transcriptional regulation, self-renewal, and pluripotency. For this reason, we assessed the binding activity of 16 common stem cell transcription factors in the PECAM1 positive, CD40 positive, and dual expressing subpopulations (Fig. 7A). Nuclear proteins were isolated from each sample and hybridized to the DNA probes corresponding to the specific transcription-factor binding sites. The number of bound transcription-factor-DNA complexes per sample was then quantified as an estimate of individual transcription factor binding activity. Consistent with previous findings, there is little difference in transcription factor binding activity between replicas of PECAM1+ and CD40+ positive cells taken from two different heterogeneous source populations with two different courses of 2iS culture (Fig. 7B and C). However, differences exist in transcription factor binding activity based on the source population for dual expressing cells. Dual expressing cells from the source population, that were removed from 2iS after 16 passages, had a higher level of transcription factor activity for all 16 tested stem cell transcription factors than dual expressing cells from the source population that were removed from 2iS only seven days after reprogramming (Fig. 7D).

There is a significant difference in FOXO1 transcription factor binding activity between PECAM1 and CD40 positive cells (Fig. 7A). FOXO1 is essential to the maintenance of ESC pluripotency in humans and mice (Zhang et al., 2011). Loss of FOXO1 activity has been

documented to result in a loss of alkaline phosphatase activity, suppression of pluripotency gene expression, and up-regulation of developmental genes such as the mesoderm specific brachyury.

There is also a significant difference in MYC transcription factor binding activity between PECAM1 and CD40 positive cells and between PECAM1 positive cells and dual expressing cells (Fig. 7A). In addition to being a core pluripotency factor often used in reprogramming, MYC plays an important role in preventing differentiation signals and maintaining an overall active transcriptional state for self-renewal and proliferation (Polo et al., 2012; Rahl et al., 2010; Sridharan et al., 2009). While *Myc* is dispensable to the reprogramming cocktail, its removal results in decreased reprogramming efficiency (Nakagawa et al., 2008; Wernig et al., 2008). MYC activity has been shown to create an open chromatin structure that allows for the binding of other critical pluripotency factors (Kim et al., 2008; Soufi et al., 2012). Higher transcription-factor binding activity of FOXO1 and MYC in PECAM1 positive cells is consistent with the transcriptome data indicating higher expression of pluripotency and self-renewal genes in PECAM positive expressing cells compared to CD40 positive cells.

Discussion

Previously we showed that culture in the presence of 2iS promotes the derivation of ESC-like iPSC from non-permissive backgrounds if used early during the reprogramming process and continually throughout colony maintenance (Garbutt et. al. 2017). This study expands upon that finding by showing that even truncated culture of 2iS for only the first seven days of reprogramming prior to colony selection and expansion is successful in deriving iPSC colonies from non-permissive strains, albeit at a reduced efficiency compared to continual 2iS culture. The ability to derive iPSC from non-permissive backgrounds using 2iS for seven days speaks to the importance of using the GSK3B and MEK inhibitors early during the reprogramming process. However, it may be prudent to use continual 2iS culture throughout colony derivation and expansion to derive more iPSC from non-permissive backgrounds and increase reprogramming efficiency. These iPSC can then be transitioned between homogenous or heterogenous states as desired by culturing the cells in or out of 2iS media.

For iPSC derived from non-permissive genetic backgrounds culture in 2iS-free conditions promotes a heterogenous cell state characterized by intolerance to passage as single cells and heterogeneous expression of the ESC cell surface marker PECAM1 and the EpiSC cell surface maker CD40. Transcriptional heterogeneity is posited to be is a natural, even necessary occurrence of the developing embryo that confers the ability to differentiate but remain pluripotent (Galvin-Burgess et al., 2013; Smith, 2001). ESC maintained in serum/LIF conditions are transcriptionally heterogeneous (Chambers et al., 2007; Hayashi et al., 2008; Singh et al., 2007; Suzuki et al., 2006a, b; Tanaka, 2009; Toyooka et al., 2008; Vieyra et al., 2009); whereas ESC derived and maintained in serum-free 2i/LIF conditions are transcriptionally homogenous (Wray et al., 2010). Heterogeneous subpopulations of ESC are functionally distinct from EpiSC, indicating that they represent different levels of pluripotency as opposed to different cellular states (Tesar et al., 2007). This is consistent with our findings that CD40 positive cells are transcriptionally equally distant from ESC

as they are to EpiSC as judged by the number of differentially expressed genes in each comparison. We conclude CD40 expression is an indication of different features of pluripotency.. Individual cells found in our heterogeneous populations are likely not fluctuating between ESC and EpiSC states, but rather between differing levels of cellular naivety. (Fig. 8). When these subpopulations are separated and allowed to grow in culture, they reestablish heterogeneity, which is consistent with previous findings (Chambers et al., 2007; Galvin-Burgess et al., 2013; Hayashi et al., 2008), and suggests that heterogeneity at the single cell level is needed to maintain colony equilibrium in the absence of 2iS media.

The transcriptionally homogenous state may be artificial as ESC maintained in serum-free 2i/LIF conditions closely resemble cells of diapause embryos (Abranches et al., 2014; Murayama et al., 2015; Papatsenko et al., 2015; Russell et al., 2015; Sugimoto et al., 2015). All iPSC in this study were cultured in serum regardless of inhibitor use. For this reason, we believe that the observed heterogeneity is not the result of serum, but the result of the absence of 2iS. These findings are consistent with the finding by Hackett et. al. 2017 that genetic background and exogenic inhibitors contribute more to transcriptional variation than serum (Hackett et al., 2017). The GSK3B and MEK inhibitors regulate homogeneity by down-regulating differentiation signals and selectively eliminating primed cells from heterogeneous populations (Hackett and Surani, 2014; Silva et al., 2008; Ying et al., 2008). Thus, we can conclude that in 2iS media, CD40 positive cell are selected against, placing 2iS at the helm of maintaining transcriptional homogeneity.

Our research shows that early culture with 2iS during reprogramming, regardless if only for the first seven days or continually, allows for the derivation of iPSC that can be removed from 2iS culture and maintained in a transcriptionally heterogeneous ESC-like state. They can be readily transitioned into a transcriptionally homogenous state by the addition of 2iS media. To our knowledge this is the first study to document the derivation of stem cells from non-permissive strains that can survive without continual addition of inhibitors.

Materials and methods

iPSC derivation

Primary mouse fibroblasts at passage three were seeded at a density of 1×10^5 cells onto a 0.1% gelatin coated six-well plate and allowed to expand in MEF expansion media for 24 hours in a 37ºC, 5% CO_2 incubator. A total of $6{\times}10^5$ fibroblasts ($1{\times}10^5$ cells per well on a six well plate) were reprogrammed for each strain. Fibroblasts were reprogrammed using the STEMCCA Cre-Excisable Constitutive Polycistronic (OKSM) Lentivirus Reprogramming Kit (Millipore) as per manufacturer’s instructions.

Upon initial iPSC emergence, individual iPSC colonies were selected and expanded, as previously described (Garbutt et al., 2018). Individual iPSC colonies were derived from a single reprogrammed and expanded cell, representing biological replicates from the same reprogramming experiment. This was further confirmed by karyotyping each expanded colony to reveal colony sex (XY vs. XX) as well as chromosomal normality. Reprogramming was repeated three times, and in both cases the same observations were made. To prevent inherent differences caused by comparing samples from two

different reprogramming experiments, individually derived and confirmed iPSC colonies, representing biological replicates from the same reprogramming experiment were assessed and reported upon for this study.

Media

ESC/LIF media: ESC/LIF media was comprised of 100% DMEM GlutaMax (Life Technologies), 1% Penicillin / Streptomycin Solution (100X) (Life Technologies), 15% ES Qualified Fetal Bovine Serum (Atlanta Biologicals), beta-mercaptoethanol (1000x) (Life Technologies), 1% Non-essential Amino Acid (NEAA) solution (100x) (Life Technologies), LIF (10 million units/1ml) (Millipore).

2iS media: 2iS media was comprised of ESC/LIF media supplemented with 3 μm of GSK3B inhibitor CHIR99201 and 1 μm of MEK inhibitor PD0325901.

NOD ESC media: NOD/ShiLtJ ESC for immunofluorescence labeling were provided by Jacob Hanna's lab (Hanna et al., 2009) and were grown in NOD ESC media: 50% DMEM/F12 (Life Technologies), 50% Neurobasal (Life Technologies), 50mM b-mercaptoethanol (1000x) (Life Technologies), 12.5 ug/ml insulin, 100 ug/ml apo-transferrin, 0.02 ug/ml progesterone (Sigma), 16 ug/ml putrescine (Sigma), 10 ug LIF (Peprotech), 3 μm Chir99021 (StemGent), 1 μm PD0325901 (StemGent), 1% Penicillin / Streptomycin Solution (100X) (Life Technologies), 1% L-glutamine (Life Technologies), 1% Non-essential Amino Acid (NEAA) solution (100x) (Life Technologies), 1% B27 supplement, sodium selenite (Sigma), BSA fraction V (Life Technologies).

NOD EpiSC media: NOD EpiSC were provided by Ludovic Vallier's lab and were grown in chemically defined media as previously published (Brons et al., 2007).

Culture conditions

Culture of permissive strains: ESC and iPSC from 129S1/SvImJ and iPSC from C57BL/6J were maintained in standard ESC/LIF media. Cells were passaged using 0.05% Trypsin-EDTA (1X) with phenol red every two to three days.

Continual 2iS conditions: One day after the reprogramming virus was added, the media was removed and fresh ESC/LIF media was added to the wells. Three days after virus addition, the ESC/LIF media was removed and 2iS media was added. Cells grown under these conditions were selected, expanded, and continually maintained in 2iS media for at least 16 passages.

2iS removed conditions: iPSC cells that were continually maintained in 2iS media for at least 16 passages were removed from 2iS media and maintained in standard ESC/LIF media for an additional eight passages.

Shortened 2iS conditions: One day after the reprogramming virus was added, the media was removed and fresh ESC/LIF media was added to the wells. Three days after virus addition, the ESC/LIF media was removed and 2iS media was added. 2iS was added

each day for only seven days. After seven days, 2iS media was removed and cells were maintained in standard ESC/LIF media throughout colony selection and expansion for a total of at least 16 passages.

2iS returned conditions.—iPSC cells that were maintained in ESC/LIF media for 16 passages were removed from ESC/LIF media and maintained in 2iS media for an additional eight passages.

ESC and EpiSC samples

Mouse embryonic fibroblasts from 129S1/SvImJ, C57BL/6J, NOD/ShiLtJ, and WSB/EiJ for iPSC derivation were derived and maintained as previously described (Garbutt et al., 2018). NOD/ShiLtJ, WSB/EiJ, and 129S1/SvImJ ESC used for RNA sequencing were provided by Dr. Laura Reinholdt. The cells were maintained in ESC/LIF serum conditions with sodium pyruvate and supplemented with GSK3B and MEK inhibitors. The NOD/ShiLtJ EpiSC used for RNA sequencing was supplied by Ludovic Vallier (Brons et al., 2007). The 129 EpiSC (FT129/4: 129/Sv background) used for RNA sequencing was provided by Alice Jouneau (Jouneau et al., 2012). Both EpiSCs were maintained as a previously described above (Brons et al., 2007; Jouneau et al., 2012).

MEF depletion

Cells grown on a feeder free layer were MEF depleted as previously described (Garbutt et al., 2018). Cells that were grown on a supportive feeder layer were MEF depleted according to a modified protocol (Chenoweth and Tesar, 2010). Spent medium was aspirated from the six-well tissue culture dishes and replaced with 1 ml of 1mg/ml collagenase type IV. The cells were incubated at 37ºC for 12 – 22 minutes. After a gentle tap, the majority of the stem cells will detach from the feeder layer within this incubation period. Detached cells were gently pipetted and transferred to a 15 ml conical tube. The six-well plate was washed gently with ESC/LIF media two times to remove any remaining cells. Media from these washes were also added to the same 15 ml conical tube. These steps were repeated for each well of the same sample and passage and combined into one 15 ml conical tube to increase the number of recovered stem cells. The cells were pelleted at 200 rcf for 15 – 30 sec. This centrifugation speed separates the denser stem cells from any remaining fibroblast cells. The stem cells will collect at the bottom of the tube and residual fibroblasts will remain in the media. The supernatant was pipetted (not aspirated) and discarded and the pellet was re-suspended in 5 ml of media and spun again. This step was repeated two times for a total of three spins. After the 3rd spin, the supernatant was gently removed through pipetting. The pellet should now contain only the denser stem cells.

iPSC visualization and immunofluorescence labeling

Brightfield images of iPSC were visualized on a traditional inverted microscope. Immunofluorescent antibody labeling was performed as previously reported (Garbutt et. al. 2017). Primary antibodies were as follows: PECAM1 monoclonal primary antibody (BD Biosciences 553370) at 1:400 and CD40 polyclonal primary antibody (R&D Systems AF440) diluted to 1:500. Secondary antibodies were as follows: polyclonal antibody Alexa Flour 488-labeled donkey anti-rat IgG, Jackson Antibodies 712-545- 153 for PECAM1 and

Alexa Flour 546-labeled donkey anti-goat IgG, Life Tech A11056 for CD40). Secondary antibodies were used at a concentration of 1:400. Cells were visualized using a Lieca DM 5500B fluorescence microscope.

Single cell capture

Single cells were imaged and captured using the Fluidigm® C1TM system and C1 integrated fluidic circuits (IFCs). Prior to capture, cells were MEF depleted as previously described and labeled for viability using the LIVE/DEAD Viability/Cytotoxicity Kit (Life Technologies). Cells with a compromised membrane labeled positive for the red-fluorescent ethidium homodimer- 1. Cells with an intact membrane at the time of capture labeled positive the green-fluorescent calcein AM. Fluorescent images of captured cells were used to determine which cells were alive at the time of capture. Dead cells are known to have a transcriptionally different profile as they begin to express genes associated with apoptosis. A total of 47 live NOD/ShiLtJ cells were sequenced: 13 cells from the population that had continually received 2iS (Always); 12 cells from the population that received 2iS after 16 passages (Late); 11 cell from the population that was removed from 2iS after 16 passages (Early); and 11 cells from the population that had received 2iS only during the first seven days of reprogramming (Never).

Flow cytometry

MEF depleted pelleted cells were re-suspended in 2 ml (1 ml per cm^2 of surface area) of Accutase. The cells were incubated in a 15 ml tube containing Accutase at 37ºC for seven minutes. Following incubation, the cells were pipetted to facilitate final dissociation and were checked under the microscope for a single cell suspension. ESC/LIF media was added to tube and the cells were pelleted. The cells were re-suspended (washed) in 5 ml of PBS, pelleted, and re-suspended in 1 ml of fresh PBS containing 3% FBS. The monoclonal conjugated CD40 antibody (PE rat anti-mouse CD40; BD Biosciences) was used at a concentration of 1:50. The monoclonal conjugated PECAM1 (CD31) antibody (APC rat anti-mouse CD31; BD Biosciences) was used at concentration of 1:50. Antibody dilutions were made in PBS containing 3% FBS. These two antibodies were added to the cells, gently flicked to mix, and incubated in the dark on ice for 30 minutes. At 15 minutes the sample was gently flicked again to mix and continued to incubate for the remaining 15 minutes. All subsequent steps were performed in a darkened room. Following incubation, the cells were pelleted, washed with 500 ul PBS, and re-suspended in 500 ul of PBS containing 1 ug/ml of DAPI. DAPI labels cells whose membranes are not intact (*i.e.* dead cells) and was included to identify and exclude dead cells from further analysis and collection. The cells were sorted at the NCSU College of Veterinary Medicine Flow Cytometry and Cell Sorting Laboratory using the Dako Cytomation Inc. MoFlo (Modular Flow Cytometer). Flow sorted data was analyzed using the FCS Express version 5 for flow cytometry software.

Culture of flow sorted cells

Following flow sorting, the cells were pelleted, and flow-sorting media was removed. The cells were re-suspended in standard ESC/LIF media and approximately 5,000 cells from CD40 positive sorted cells and PECAM1 positive sorted cells were plated in separate wells of six-well plates containing a coverslip and coated with an inactivated feeder layer.

The media was changed the following day and the cells were allowed to grow as usual until the emergence of colonies large enough to withstand immunofluorescence labeling (approximately 3 – 4 days).

RNA sequencing

Single cells: Single cells were randomly captured, and cDNA amplified using the Fluidigm® C1TM system and C1 IFCs. cDNA libraries were generated using the NexteraXT library prep kit and the Fluidigm® C1TM system and C1 IFCs. Microfluidic technology performs reverse transcription, cDNA amplification, and library preparation in nanoliter volumes. Single cell capture and library prep were performed at the Texas A&M Institute for Genome Sciences & Society facility. Shallow single cell mRNA sequencing of 50,000 reads per cell is sufficient for unbiased cell type classification of single cells captured using microfluidics from a heterogeneous source population. Here we exceed that minimum by generating 2 million reads per cell, which has also been found to provide sensitive but accurate coverage.

Flow sorted cells: Flow sorted cells were pelleted, washed with PBS, and flash frozen in liquid nitrogen as cell pellets. RNA was extracted from the cell pellets using the Promega Maxwell 16 LEV simplyRNA cells kit. cDNA, Library prep, and RNA sequencing was performed by the Texas A&M Institute for Genome Sciences & Society facility. Extracted RNA was run on a bioanalyzer to confirm sample integrity. cDNA libraries were created using the TruSeg Stranded mRNA library prep kit from Illumina. Samples were pooled and run on a 150 × 150 high output NextSeq, generating 800 million paired end reads (approximately 32 million reads/sample).

RNAseq Analysis

Single cell data analysis: Trimmomatic was used to remove sequencing adapters. Each strain was aligned to their reference genome available online from the UNC system Genetics website (csbio.unc.edu). Following alignment, alignment positions were transformed into C57BL/6J coordinate space using the Pylapels program. FPKM values were generated using the cufflink software and analyzed using the Fluidigm Singular Analysis Toolset 3.0 package for R (www.fluidigm.com). Outliers were identified from each of the four 2iS culture conditions based on the identifyOutliers function in the Singular Analysis toolset package. This function considers a set of genes detected in at least half of the samples and samples with a median expression value below the 15th percentile for those genes were considered outliers and were removed. Two outliers were identified and removed from each of the four 2iS culture conditions leaving a total of 39 remaining cells for subsequent analysis. Hierarchical clustering of the top 200 genes, identified according to one-way ANOVA p-values, across the 39 single cells was performed using the Fluidigm Singular package that clusters genes based on Pearson correlation and samples based on the Euclidean method. For genes and sample clusters the complete linkage method is used to find similar clusters.

ESC controls, EpiSC controls, and flow sorted sample data analysis: 150 base-pair reads for each sample were generated using Illumina NextSeq 2500 instrument on 8 sequencing lanes. A total of 453 million reads were checked for any adapter sequences and

to trim any low-quality bases using Trimmomatic resulting in approximately 285 million filtered reads (63%) out of which a total of 229 million filtered reads (approximately 80%) mapped to the GRcm38/mm10 assembly. Read mapping was performed using TopHat version 2.1.1. HTSeq was used to generate raw read counts per gene using intersection-nonempty parameter to account for ambiguous read mappings. Differential gene expression tests were then performed used DESeq2 following recommended guidelines. Plots were generated using R programming language. The resulting gene expression values were uploaded to the Ingenuity Pathways Analysis (Prasanth et al.) program (QIAGEN, Venlo, Netherlands; www.ingenuity.com Application Build 261899, Content Version 18030641) for gene function and biological pathway analysis. Additionally, the compare feature of IPA was used to identify the number of unique and overlapping differentially expressed genes between the PECAM1 positive iPSC and ESC, CD40 positive iPSC and EpiSC, and dual expressing iPSC and ESC and EpiSC.

Embryonic stem cells from the 129S1/SvImJ, NOD/ShiLtJ, and WSB/EiJ backgrounds were used as replicates to predict the general ESC transcriptome to compare to flow sorted iPSC populations. Likewise, epiblast stem cells from the 129 and NOD backgrounds were used as replicates to predict the general EpiSC transcriptome to compare to flow sorted iPSC subpopulations.

Similarly, labeled flow sorted cells from the two heterogeneous source populations were used as replicates to compare to the differently labeled flow sorted cells and to compare to the general ESC and EpiSC profiles. For example, flow sorted CD40+ cells from the source population that had been maintained in 2iS media for only the first seven days and flow sorted CD40+ cells from the source population that had been removed from 2iS media after continual culture in 2iS for >16 passages were used as replicates.

Our premise for using these samples as replicates is that they have the same phenotype (positive labeling for a specific marker), and they are both not currently maintained in 2iS media. Based on our single cell data we saw no difference between cells under similar culture conditions, *i.e.* current presence in or out of 2iS media. Furthermore, we were investigating the difference between subgroups (CD40+ cells vs. PECAM1+ cells), and not the difference between culture conditions.

To identify the unique profile of dual expressing cells, dual expressing cells were compared simultaneously to both CD40+ cells and PECAM1+ cells using the additive function (Dual vs. PECAM1 + CD40) in DeSeq2. The reason for this is because dual expressing cells express both markers and individual one to one comparison of the transcriptome may yield differentially expressed genes characteristic of the opposing label. For example, Dual-expressing cells compared to PECAM1+ cells may yield differentially expressed genes that are also characteristic of the CD40+ profile because dual-expressing cells also express CD40+. Thus, both PECAM1+ and tCD40+ profiles were considered when identifying a unique profile of dual-expressing cells.

Gene datasets and functional association analysis

Gene datasets identified by DeSeq2 sample comparisons were imported into the Ingenuity Pathway Analysis Program to detect functional classifications and associations. The p-value listed for each functional group is a measure of the likelihood that the association between the genes in our dataset and a given function is due to random chance. The p-value is calculated by considering the total number of genes in our dataset involved with a given function and the total number of genes that are known based on the literature to be associated with that function. The p-value was generated using the Fisher exact test and was corrected using the Benjamini-Hochberg method.

IPA uses a z score to generate predictions about the activation state of a given function by first considering the documented activation state in the literature for a given gene and then by considering the observed differential up/down regulation of the gene in the provided data set. Considering that there may be conflicting reports for a given gene activity in the literature, the weight of the findings for each gene is tabulated. A prediction of a biological function is given the label, “bias” when the gene regulation (up/down) in the data set and the regulation of the gene in the literature are skewed. If there is an unequal amount of up/down regulated genes in a dataset there is a higher probability that the prediction will be biased because the algorithm takes into account how the gene is being regulated in the data. In our data, we compare PECAM1+ to CD40+ cells. Genes with increased expression in this data set have increased expression in PECAM1+ cells, whereas genes in this dataset with decreased expression have increased expression in CD40+ cells. Our predictions have a bias label because the individual gene lists were imported separately into IPA, so as to define a unique transcriptional profile from each strain.

Transcription factor profiling

Whole cell lyses from flash frozen flow-sorted samples were processed by Signosis to isolate nuclear proteins. Transcription factor DNA-binding activity was assessed by Signosis using their Stem Cell TF Activation Profiling Plate Array I FA-1003 for 16 common stem cell transcription factors: EGR1, OCT4, FOXD3, FOXO, NANOG, SOX2, SOX18, ETS, GLI, KLF4, MEF2, MYC, RUNX1, PAX6, TCF/LEF and GATA. Biotin-labeled transcription factor probes that bind to the consensus sequence of specific transcription factor DNA-binding sites were incubated with the nuclear protein extracts of each sample to allow for the formation of a transcription factor-DNA binding complex. A filter plate was then used to retain bound probes and remove free DNA probes. The number of bound probe complexes, representing transcription factor binding activity within the sample, were then quantified by hybridizing the probe to a corresponding well on a 96 well hybridization plate. Probes were detected with streptavidin-HPR and luminescence was reported as relative light units (RLUs) on a microplate luminometer.

Acknowledgements

We would like to thank Jacob Hanna (Weizmann Institute of Science) for the NOD/ShiLtJ ESC, Ludovic Vallier (Welcome Sanger Institute) for the NOD/ShiLtJ EpiSC, Laura Reinholdt (The Jackson Laboratory) for the NOD/ShiLtJ, WSB/EiJ, and 129S1/SvImJ ESC, and Alice Jouneau (French National Institute for Agricultural Research) for the 129 EpiSC.

Funding

This work was supported by National Institutes of Health grants P50 MH090338 and RM1 HG008529 to DWT and U01 ES026717 to DLA, and a National Science Foundation Graduate Fellowship to TG.

References

Abranches E, Bekman E, and Henrique D (2013). Generation and characterization of a novel mouse embryonic stem cell line with a dynamic reporter of Nanog expression. PloS one 8, e59928. [PubMed: 23527287]

Abranches E, Guedes AM, Moravec M, Maamar H, Svoboda P, Raj A, and Henrique D (2014). Stochastic NANOG fluctuations allow mouse embryonic stem cells to explore pluripotency. Development (Cambridge, England) 141, 2770–2779.

Bao S, Tang F, Li X, Hayashi K, Gillich A, Lao K, and Surani MA (2009). Epigenetic reversion of post-implantation epiblast to pluripotent embryonic stem cells. Nature 461, 1292–1295. [PubMed: 19816418]

Bauer JH, Liu KD, You Y, Lai SY, and Goldsmith MA (1998). Heteromerization of the gammac chain with the interleukin-9 receptor alpha subunit leads to STAT activation and prevention of apoptosis. The Journal of biological chemistry 273, 9255–9260. [PubMed: 9535918]

Besser D (2004). Expression of nodal, lefty-a, and lefty-B in undifferentiated human embryonic stem cells requires activation of Smad2/3. The Journal of biological chemistry 279, 45076–45084. [PubMed: 15308665]

Boroviak T, Loos R, Bertone P, Smith A, and Nichols J (2014). The ability of inner-cell-mass cells to self-renew as embryonic stem cells is acquired following epiblast specification. Nature cell biology 16, 516–528. [PubMed: 24859004]

Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS, Zucker JP, Guenther MG, Kumar RM, Murray HL, Jenner RG, et al. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. Cell 122, 947–956. [PubMed: 16153702]

Brons IG, Smithers LE, Trotter MW, Rugg-Gunn P, Sun B, Chuva de Sousa Lopes SM, Howlett SK, Clarkson A, Ahrlund-Richter L, Pedersen RA, et al. (2007). Derivation of pluripotent epiblast stem cells from mammalian embryos. Nature 448, 191–195. [PubMed: 17597762]

Buehr M, Meek S, Blair K, Yang J, Ure J, Silva J, McLay R, Hall J, Ying QL, and Smith A (2008). Capture of authentic embryonic stem cells from rat blastocysts. Cell 135, 1287–1298. [PubMed: 19109897]

Cao Y, Szabolcs A, Dutta SK, Yaqoob U, Jagavelu K, Wang L, Leof EB, Urrutia RA, Shah VH, and Mukhopadhyay D (2010). Neuropilin-1 mediates divergent R-Smad signaling and the myofibroblast phenotype. The Journal of biological chemistry 285, 31840–31848. [PubMed: 20675371]

Castelo-Branco G, Amaral PP, Engstrom PG, Robson SC, Marques SC, Bertone P, and Kouzarides T (2013). The non-coding snRNA 7SK controls transcriptional termination, poising, and bidirectionality in embryonic stem cells. Genome biology 14, R98. [PubMed: 24044525]

Cerezo A, Guadamillas MC, Goetz JG, Sanchez-Perales S, Klein E, Assoian RK, and del Pozo MA (2009). The absence of caveolin-1 increases proliferation and anchorage- independent growth by a Rac-dependent, Erk-independent mechanism. Molecular and cellular biology 29, 5046–5059. [PubMed: 19620284]

Chambers I, Silva J, Colby D, Nichols J, Nijmeijer B, Robertson M, Vrana J, Jones K, Grotewold L, and Smith A (2007). Nanog safeguards pluripotency and mediates germline development. Nature 450, 1230–1234. [PubMed: 18097409]

Chazaud C, Yamanaka Y, Pawson T, and Rossant J (2006). Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the Grb2-MAPK pathway. Developmental cell 10, 615–624. [PubMed: 16678776]

Chen N, and Wang X (2014). Role of IL-9 and STATs in hematological malignancies (Review). Oncology letters 7, 602–610. [PubMed: 24520283]

Chenoweth JG, and Tesar PJ (2010). Isolation and maintenance of mouse epiblast stem cells. Methods in molecular biology (Clifton, NJ) 636, 25–44.

Cinelli P, Casanova EA, Uhlig S, Lochmatter P, Matsuda T, Yokota T, Rulicke T, Ledermann B, and Burki K (2008). Expression profiling in transgenic FVB/N embryonic stem cells overexpressing STAT3. BMC developmental biology 8, 57. [PubMed: 18500982]

Collins CA, and Watt FM (2008). Dynamic regulation of retinoic acid-binding proteins in developing, adult and neoplastic skin reveals roles for beta-catenin and Notch signalling. Developmental biology 324, 55–67. [PubMed: 18805411]

Czechanski A, Byers C, Greenstein I, Schrode N, Donahue LR, Hadjantonakis AK, and Reinholdt LG (2014). Derivation and characterization of mouse embryonic stem cells from permissive and nonpermissive strains. Nature protocols 9, 559–574. [PubMed: 24504480]

Feng B, Ng JH, Heng JC, and Ng HH (2009). Molecules that promote or enhance reprogramming of somatic cells to induced pluripotent stem cells. Cell stem cell 4, 301–312. [PubMed: 19341620]

Galvin-Burgess KE, Travis ED, Pierson KE, and Vivian JL (2013). TGF-beta-superfamily signaling regulates embryonic stem cell heterogeneity: self-renewal as a dynamic and regulated equilibrium. Stem cells (Dayton, Ohio) 31, 48–58.

Garbutt TA, Konneker TI, Konganti K, Hillhouse AE, Swift-Haire F, Jones A, Phelps D, Aylor DL, and Threadgill DW (2018). Permissiveness to form pluripotent stem cells may be an evolutionarily derived characteristic in Mus musculus. Sci Rep 8, 14706. [PubMed: 30279419]

Geetha N, Mihaly J, Stockenhuber A, Blasi F, Uhrin P, Binder BR, Freissmuth M, and Breuss JM (2011). Signal integration and coincidence detection in the mitogen-activated protein kinase/extracellular signal-regulated kinase (ERK) cascade: concomitant activation of receptor tyrosine kinases and of LRP-1 leads to sustained ERK phosphorylation via down-regulation of dual specificity phosphatases (DUSP1 and −6). The Journal of biological chemistry 286, 25663–25674. [PubMed: 21610072]

Golipour A, David L, Liu Y, Jayakumaran G, Hirsch CL, Trcka D, and Wrana JL (2012). A late transition in somatic cell reprogramming requires regulators distinct from the pluripotency network. Cell stem cell 11, 769–782. [PubMed: 23217423]

Grabole N, Tischler J, Hackett JA, Kim S, Tang F, Leitch HG, Magnusdottir E, and Surani MA (2013). Prdm14 promotes germline fate and naive pluripotency by repressing FGF signalling and DNA methylation. EMBO reports 14, 629–637. [PubMed: 23670199]

Graf T, and Stadtfeld M (2008). Heterogeneity of embryonic and adult stem cells. Cell stem cell 3, 480–483. [PubMed: 18983963]

Guo G, Yang J, Nichols J, Hall JS, Eyres I, Mansfield W, and Smith A (2009). Klf4 reverts developmentally programmed restriction of ground state pluripotency. Development (Cambridge, England) 136, 1063–1069.

Hackett JA, Kobayashi T, Dietmann S, and Surani MA (2017). Activation of Lineage Regulators and Transposable Elements across a Pluripotent Spectrum. Stem cell reports 8, 1645–1658. [PubMed: 28591649]

Hackett JA, and Surani MA (2014). Regulatory principles of pluripotency: from the ground state up. Cell stem cell 15, 416–430. [PubMed: 25280218]

Han DW, Tapia N, Joo JY, Greber B, Arauzo-Bravo MJ, Bernemann C, Ko K, Wu G, Stehling M, Do JT, et al. (2010). Epiblast stem cell subpopulations represent mouse embryos of distinct pregastrulation stages. Cell 143, 617–627. [PubMed: 21056461]

Hanna J, Markoulaki S, Mitalipova M, Cheng AW, Cassady JP, Staerk J, Carey BW, Lengner CJ, Foreman R, Love J, et al. (2009). Metastable pluripotent states in NOD-mouse-derived ESCs. Cell stem cell 4, 513–524. [PubMed: 19427283]

Hayashi K, de Sousa Lopes SMC, Tang F, Lao K, and Surani MA (2008). Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states. Cell stem cell 3, 391–401. [PubMed: 18940731]

Huang L, Szymanska K, Jensen VL, Janecke AR, Innes AM, Davis EE, Frosk P, Li C, Willer JR, Chodirker BN, et al. (2011). TMEM237 is mutated in individuals with a Joubert syndrome related disorder and expands the role of the TMEM family at the ciliary transition zone. American journal of human genetics 89, 713–730. [PubMed: 22152675]

Huhtala ML, Pesonen K, Kalkkinen N, and Stenman UH (1982). Purification and characterization of a tumor-associated trypsin inhibitor from the urine of a patient with ovarian cancer. The Journal of biological chemistry 257, 13713–13716. [PubMed: 7142173]

Jeffries CD, Perkins DO, and Guan X (2010). Gene processing control loops suggested by sequencing, splicing, and RNA folding. BMC bioinformatics 11, 602. [PubMed: 21167075]

Jouneau A, Ciaudo C, Sismeiro O, Brochard V, Jouneau L, Vandormael-Pournin S, Coppee JY, Zhou Q, Heard E, Antoniewski C, et al. (2012). Naive and primed murine pluripotent stem cells have distinct miRNA expression profiles. RNA (New York, NY) 18, 253–264.

Kim J, Chu J, Shen X, Wang J, and Orkin SH (2008). An extended transcriptional network for pluripotency of embryonic stem cells. Cell 132, 1049–1061. [PubMed: 18358816]

Kimura H, Kwan KM, Zhang Z, Deng JM, Darnay BG, Behringer RR, Nakamura T, de Crombrugghe B, and Akiyama H (2008). Cthrc1 is a positive regulator of osteoblastic bone formation. PloS one 3, e3174. [PubMed: 18779865]

Leitch HG, McEwen KR, Turp A, Encheva V, Carroll T, Grabole N, Mansfield W, Nashun B, Knezovich JG, Smith A, et al. (2013). Naive pluripotency is associated with global DNA hypomethylation. Nature structural & molecular biology 20, 311–316.

Li L, and Clevers H (2010). Coexistence of quiescent and active adult stem cells in mammals. Science (New York, NY) 327, 542–545.

Liu J, Ashton MP, Sumer H, O'Bryan MK, Brodnicki TC, and Verma PJ (2011). Generation of stable pluripotent stem cells from NOD mouse tail-tip fibroblasts. Diabetes 60, 1393–1398. [PubMed: 21464439]

Marks H, Kalkan T, Menafra R, Denissov S, Jones K, Hofemeister H, Nichols J, Kranz A, Stewart AF, Smith A, et al. (2012). The transcriptional and epigenomic foundations of ground state pluripotency. Cell 149, 590–604. [PubMed: 22541430]

Martello G, Sugimoto T, Diamanti E, Joshi A, Hannah R, Ohtsuka S, Gottgens B, Niwa H, and Smith A (2012). Esrrb is a pivotal target of the Gsk3/Tcf3 axis regulating embryonic stem cell self-renewal. Cell stem cell 11, 491–504. [PubMed: 23040478]

Min IM, Waterfall JJ, Core LJ, Munroe RJ, Schimenti J, and Lis JT (2011). Regulating RNA polymerase pausing and transcription elongation in embryonic stem cells. Genes & development 25, 742–754. [PubMed: 21460038]

Murayama H, Masaki H, Sato H, Hayama T, Yamaguchi T, and Nakauchi H (2015). Successful reprogramming of epiblast stem cells by blocking nuclear localization of beta-catenin. Stem cell reports 4, 103–113. [PubMed: 25556568]

Najm FJ, Chenoweth JG, Anderson PD, Nadeau JH, Redline RW, McKay RD, and Tesar PJ (2011). Isolation of epiblast stem cells from preimplantation mouse embryos. Cell stem cell 8, 318–325. [PubMed: 21362571]

Nakagawa M, Koyanagi M, Tanabe K, Takahashi K, Ichisaka T, Aoi T, Okita K, Mochiduki Y, Takizawa N, and Yamanaka S (2008). Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. Nature biotechnology 26, 101–106.

Nichols J, Jones K, Phillips JM, Newland SA, Roode M, Mansfield W, Smith A, and Cooke A (2009). Validated germline-competent embryonic stem cell lines from nonobese diabetic mice. Nature medicine 15, 814–818.

Ning X, Sun S, Zhang K, Liang J, Chuai Y, Li Y, and Wang X (2012). S100A6 protein negatively regulates CacyBP/SIP-mediated inhibition of gastric cancer cell proliferation and tumorigenesis. PloS one 7, e30185. [PubMed: 22295074]

O'Reilly D, Dienstbier M, Cowley SA, Vazquez P, Drozdz M, Taylor S, James WS, and Murphy S (2013). Differentially expressed, variant U1 snRNAs regulate gene expression in human cells. Genome research 23, 281–291. [PubMed: 23070852]

Orford KW, and Scadden DT (2008). Deconstructing stem cell self-renewal: genetic insights into cell-cycle regulation. Nature reviews Genetics 9, 115–128.

Papatsenko D, Darr H, Kulakovskiy IV, Waghray A, Makeev VJ, MacArthur BD, and Lemischka IR (2015). Single-Cell Analyses of ESCs Reveal Alternative Pluripotent Cell States and Molecular Mechanisms that Control Self-Renewal. Stem cell reports 5, 207–220. [PubMed: 26267829]

Park KS (2011). Tgf-Beta family signaling in embryonic stem cells. International journal of stem cells 4, 18–23. [PubMed: 24298330]

Patterson KI, Brummer T, O'Brien PM, and Daly RJ (2009). Dual-specificity phosphatases: critical regulators with diverse cellular targets. The Biochemical journal 418, 475–489. [PubMed: 19228121]

Pera MF, Andrade J, Houssami S, Reubinoff B, Trounson A, Stanley EG, Ward-van Oostwaard D, and Mummery C (2004). Regulation of human embryonic stem cell differentiation by BMP-2 and its antagonist noggin. Journal of cell science 117, 1269–1280. [PubMed: 14996946]

Persaud SD, Lin YW, Wu CY, Kagechika H, and Wei LN (2013). Cellular retinoic acid binding protein I mediates rapid non-canonical activation of ERK1/2 by all-trans retinoic acid. Cellular signalling 25, 19–25. [PubMed: 22982089]

Pollen AA, Nowakowski TJ, Shuga J, Wang X, Leyrat AA, Lui JH, Li N, Szpankowski L, Fowler B, Chen P, et al. (2014). Low-coverage single-cell mRNA sequencing reveals cellular heterogeneity and activated signaling pathways in developing cerebral cortex. Nature biotechnology 32, 1053–1058.

Polo JM, Anderssen E, Walsh RM, Schwarz BA, Nefzger CM, Lim SM, Borkent M, Apostolou E, Alaei S, Cloutier J, et al. (2012). A molecular roadmap of reprogramming somatic cells into iPS cells. Cell 151, 1617–1632. [PubMed: 23260147]

Polo JM, Liu S, Figueroa ME, Kulalert W, Eminli S, Tan KY, Apostolou E, Stadtfeld M, Li Y, Shioda T, et al. (2010). Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. Nature biotechnology 28, 848–855.

Prasanth KV, Camiolo M, Chan G, Tripathi V, Denis L, Nakamura T, Hubner MR, and Spector DL (2010). Nuclear organization and dynamics of 7SK RNA in regulating gene expression. Molecular biology of the cell 21, 4184–4196. [PubMed: 20881057]

Rahl PB, Lin CY, Seila AC, Flynn RA, McCuine S, Burge CB, Sharp PA, and Young RA (2010). c-Myc regulates transcriptional pause release. Cell 141, 432–445. [PubMed: 20434984]

Roux PP, and Blenis J (2004). ERK and p38 MAPK-activated protein kinases: a family of protein kinases with diverse biological functions. Microbiology and molecular biology reviews : MMBR 68, 320–344. [PubMed: 15187187]

Rugg-Gunn PJ, Cox BJ, Lanner F, Sharma P, Ignatchenko V, McDonald AC, Garner J, Gramolini AO, Rossant J, and Kislinger T (2012). Cell-surface proteomics identifies lineage-specific markers of embryo-derived stem cells. Developmental cell 22, 887–901. [PubMed: 22424930]

Russell R, Ilg M, Lin Q, Wu G, Lechel A, Bergmann W, Eiseler T, Linta L, Kumar PP, Klingenstein M, et al. (2015). A Dynamic Role of TBX3 in the Pluripotency Circuitry. Stem cell reports 5, 1155–1170. [PubMed: 26651606]

Shakiba N, White CA, Lipsitz YY, Yachie-Kinoshita A, Tonge PD, Hussein SMI, Puri MC, Elbaz J, Morrissey-Scoot J, Li M, et al. (2015). CD24 tracks divergent pluripotent states in mouse and human cells. Nature Communications 6, 7329.

Silva J, Barrandon O, Nichols J, Kawaguchi J, Theunissen TW, and Smith A (2008). Promotion of reprogramming to ground state pluripotency by signal inhibition. PLoS biology 6, e253. [PubMed: 18942890]

Singh AM, Hamazaki T, Hankowski KE, and Terada N (2007). A heterogeneous expression pattern for Nanog in embryonic stem cells. Stem cells (Dayton, Ohio) 25, 2534–2542.

Smith AG (2001). Embryo-derived stem cells: of mice and men. Annual review of cell and developmental biology 17, 435–462.

Soufi A, Donahue G, and Zaret KS (2012). Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. Cell 151, 994–1004. [PubMed: 23159369]

Sridharan R, Tchieu J, Mason MJ, Yachechko R, Kuoy E, Horvath S, Zhou Q, and Plath K (2009). Role of the murine reprogramming factors in the induction of pluripotency. Cell 136, 364–377. [PubMed: 19167336]

Sugawara A, Goto K, Sotomaru Y, Sofuni T, and Ito T (2006). Current status of chromosomal abnormalities in mouse embryonic stem cell lines used in Japan. Comparative medicine 56, 31–34. [PubMed: 16521857]

Sugimoto M, Kondo M, Koga Y, Shiura H, Ikeda R, Hirose M, Ogura A, Murakami A, Yoshiki A, Chuva de Sousa Lopes SM, et al. (2015). A simple and robust method for establishing homogeneous mouse epiblast stem cell lines by wnt inhibition. Stem cell reports 4, 744–757. [PubMed: 25818811]

Suzuki A, Raya A, Kawakami Y, Morita M, Matsui T, Nakashima K, Gage FH, Rodriguez-Esteban C, and Izpisua Belmonte JC (2006a). Maintenance of embryonic stem cell pluripotency by Nanog-mediated reversal of mesoderm specification. Nature clinical practice Cardiovascular medicine 3 Suppl 1, S114–122.

Suzuki A, Raya A, Kawakami Y, Morita M, Matsui T, Nakashima K, Gage FH, Rodriguez-Esteban C, and Izpisua Belmonte JC (2006b). Nanog binds to Smad1 and blocks bone morphogenetic protein-induced differentiation of embryonic stem cells. Proceedings of the National Academy of Sciences of the United States of America 103, 10294–10299. [PubMed: 16801560]

Tanaka TS (2009). Transcriptional heterogeneity in mouse embryonic stem cells. Reproduction, fertility, and development 21, 67–75. [PubMed: 19152747]

Tesar PJ, Chenoweth JG, Brook FA, Davies TJ, Evans EP, Mack DL, Gardner RL, and McKay RD (2007). New cell lines from mouse epiblast share defining features with human embryonic stem cells. Nature 448, 196–199. [PubMed: 17597760]

Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, Marshall VS, and Jones JM (1998). Embryonic stem cell lines derived from human blastocysts. Science (New York, NY) 282, 1145–1147.

Toyooka Y, Shimosato D, Murakami K, Takahashi K, and Niwa H (2008). Identification and characterization of subpopulations in undifferentiated ES cell culture. Development (Cambridge, England) 135, 909–918.

Ueda O, Jishage K, Kamada N, Uchida S, and Suzuki H (1995). Production of mice entirely derived from embryonic stem (ES) cell with many passages by coculture of ES cells with cytochalasin B induced tetraploid embryos. Experimental animals 44, 205–210. [PubMed: 7556421]

Vieyra DS, Rosen A, and Goodell MA (2009). Identification and characterization of side population cells in embryonic stem cell cultures. Stem cells and development 18, 1155–1166. [PubMed: 19113897]

Wang Z, Iwasaki M, Ficara F, Lin C, Matheny C, Wong SH, Smith KS, and Cleary ML (2010). GSK-3 promotes conditional association of CREB and its coactivators with MEIS1 to facilitate HOX-mediated transcription and oncogenesis. Cancer cell 17, 597–608. [PubMed: 20541704]

Wernig M, Meissner A, Cassady JP, and Jaenisch R (2008). c-Myc is dispensable for direct reprogramming of mouse fibroblasts. Cell stem cell 2, 10–12. [PubMed: 18371415]

Wray J, Kalkan T, and Smith AG (2010). The ground state of pluripotency. Biochemical Society transactions 38, 1027–1032. [PubMed: 20658998]

Wu J, and Izpisua Belmonte, Juan C. (2015). Dynamic Pluripotent Stem Cell States and Their Applications. Cell stem cell 17, 509–525. [PubMed: 26544113]

Yamaji M, Ueda J, Hayashi K, Ohta H, Yabuta Y, Kurimoto K, Nakato R, Yamada Y, Shirahige K, and Saitou M (2013). PRDM14 ensures naive pluripotency through dual regulation of signaling and epigenetic pathways in mouse embryonic stem cells. Cell stem cell 12, 368–382. [PubMed: 23333148]

Yamamoto S, Nishimura O, Misaki K, Nishita M, Minami Y, Yonemura S, Tarui H, and Sasaki H (2008). Cthrc1 selectively activates the planar cell polarity pathway of Wnt signaling by stabilizing the Wnt-receptor complex. Developmental cell 15, 23–36. [PubMed: 18606138]

Ye S, Li P, Tong C, and Ying QL (2013). Embryonic stem cell self-renewal pathways converge on the transcription factor Tfcp2l1. The EMBO journal 32, 2548–2560. [PubMed: 23942238]

Ying QL, Nichols J, Chambers I, and Smith A (2003). BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. Cell 115, 281–292. [PubMed: 14636556]

Ying QL, Wray J, Nichols J, Batlle-Morera L, Doble B, Woodgett J, Cohen P, and Smith A (2008). The ground state of embryonic stem cell self-renewal. Nature 453, 519–523. [PubMed: 18497825]

Zhang B, Denomme MM, White CR, Leung KY, Lee MB, Greene ND, Mann MR, Trasler JM, and Baltz JM (2015). Both the folate cycle and betaine-homocysteine methyltransferase contribute

methyl groups for DNA methylation in mouse blastocysts. FASEB journal : official publication of the Federation of American Societies for Experimental Biology 29, 1069–1079. [PubMed: 25466894]

Zhang J, Klos M, Wilson GF, Herman AM, Lian X, Raval KK, Barron MR, Hou L, Soerens AG, Yu J, et al. (2012). Extracellular matrix promotes highly efficient cardiac differentiation of human pluripotent stem cells: the matrix sandwich method. Circulation research 111, 1125–1136. [PubMed: 22912385]

Zhang X, Yalcin S, Lee DF, Yeh TY, Lee SM, Su J, Mungamuri SK, Rimmele P, Kennedy M, Sellers R, et al. (2011). FOXO1 is an essential regulator of pluripotency in human embryonic stem cells. Nature cell biology 13, 1092–1099. [PubMed: 21804543]

Zhao XF, Colaizzo-Anas T, Nowak NJ, Shows TB, Elliott RW, and Aplan PD (1998). The mammalian homologue of mago nashi encodes a serum-inducible protein. Genomics 47, 319–322. [PubMed: 9479507]

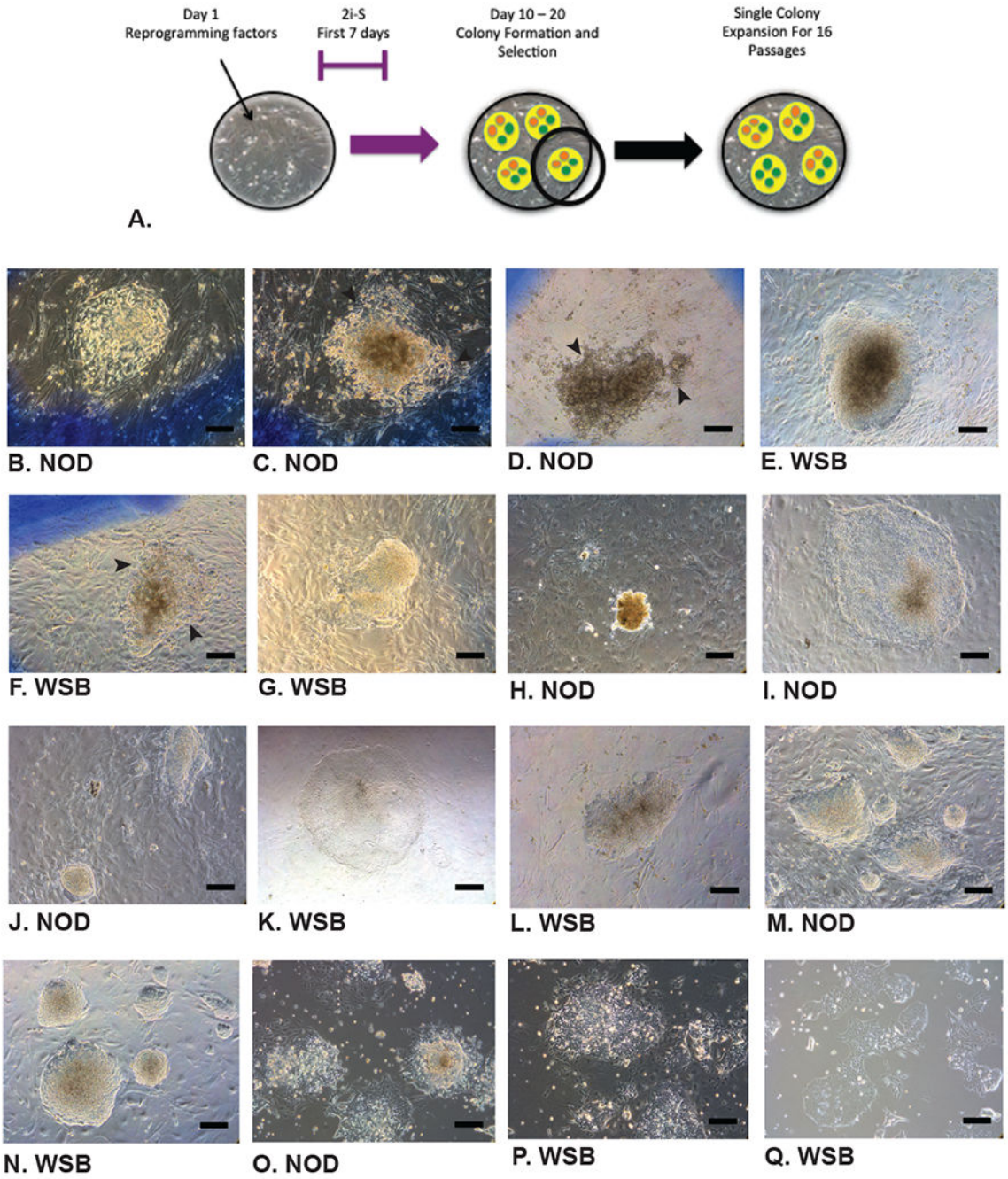

Fig. 1.

Variation in iPSC colony morphology is determined by culture conditions. **A** 2iS was added three days after the lentivirus and maintained for only the first seven days. Colonies were selected and expanded for 16 passages in standard media. **B, E** NOD/ShiLtJ and WSB/EiJ iPSC colonies before 2iS removal. **C, F** NOD/ShiLtJ and WSB/EiJ iPSC colonies after 2iS removal show a loss of colony stability. **D, G** WSB/EiJ colonies retain morphology better than NOD/ShiLtJ colonies with the loss of 2iS. **H** Colonies passaged as single cells differentiated and died. **I-L** Colonies passaged as cell clumps gave rise to both flat colonies

and round colonies within the same population when maintained on a mitotically inactive MEF feeder layers. **M, N** Fewer flattened colonies persisted with each passage, giving rise to a population of mostly three-dimensional dome shaped colonies on a MEF feeder layer for both NOD/ShiLtJ at passage 10 and WSB/EiJ at passage 8. **O, P** iPSC colonies from NOD/ShiLtJ and WSB/EiJ were intolerant to growth on a feeder-free gelatin layer even when passaged as cell clumps. Colonies became flattened and began to differentiate. **Q** The phenotype was exacerbated to a spindle like appearance when trypsin was used. Scale bars 200 μm.

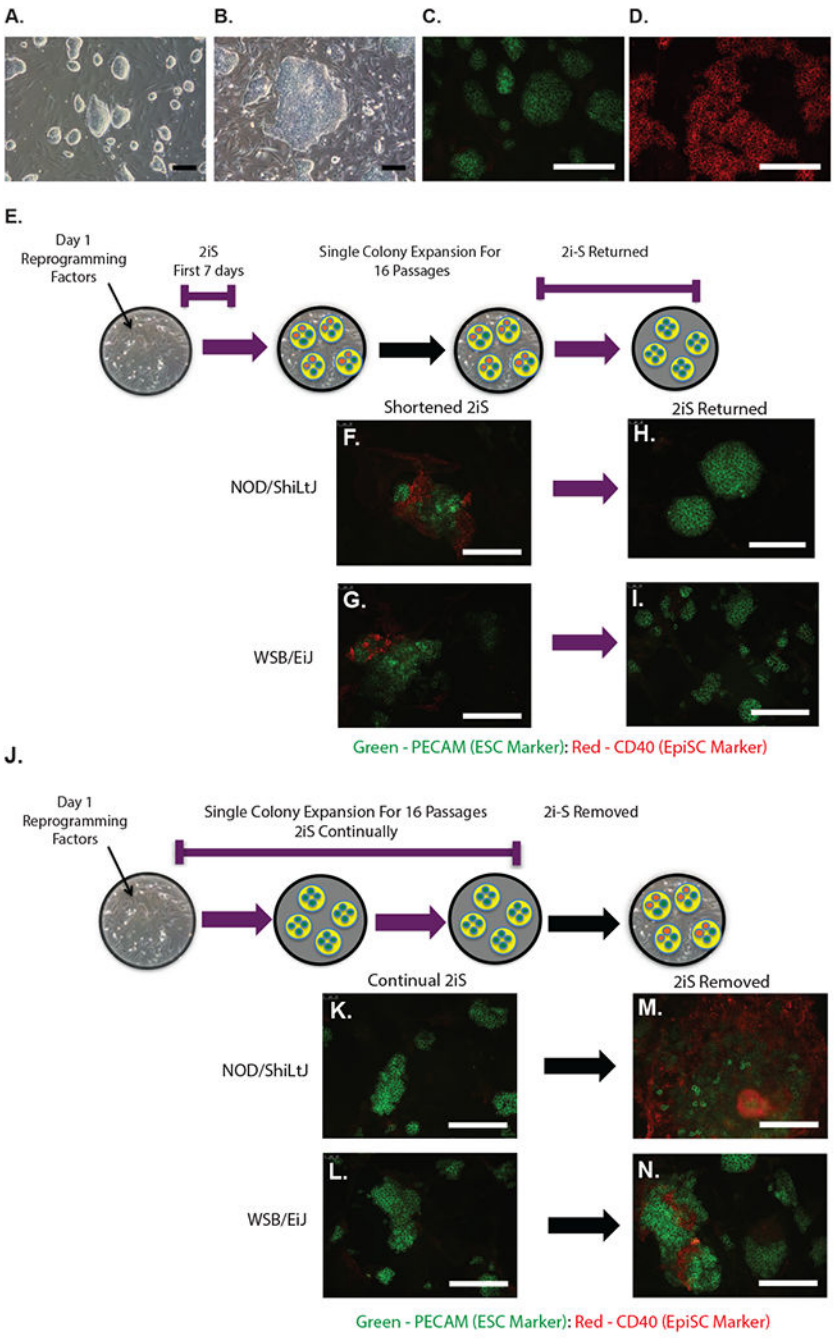

Fig. 2.

Cell surface marker expression in under different culture conditions. **A, B** Control NOD/ShiLtJ ESC and NOD/EiJ EpiSC grown in serum-free media on a mitotically inactive MEF feeder layer. NOD/ShiLtJ ESC form three-dimensional dome shaped colonies and NOD/ShiLtJ EpiSC formed broad, flattened colonies. **C** NOD/ShiLtJ ESC label positive for the ESC surface marker PECAM1 and do not label for the EpiSC surface marker CD40. **D** NOD/ShiLtJ EpiSC label positive for the EpiSC surface marker CD40 and do not label for the PECAM1 embryonic stem cell surface marker. **E** NOD/ShiLtJ and WSB/EiJ iPSC

derived with shortened 2iS culture and returned to 2iS culture after 16 passages. **F,G** In 2iS free conditions, these cells label heterogeneously for both PECAM1 and CD40. **H,I** When these colonies are grown in 2iS media again after 16 passages they label homogenously for PECAM1 expression. **J** NOD/ShiLtJ and WSB/EiJ iPSC.2iS colonies that were derived and maintained in 2iS for 16 passages, were then removed from 2iS culture. **K,L** In 2iS conditions cells label homogenously for PECAM1. **M,N** In 2iS free conditions the cells transition to a heterogenous expression pf PECAM1 and CD40.

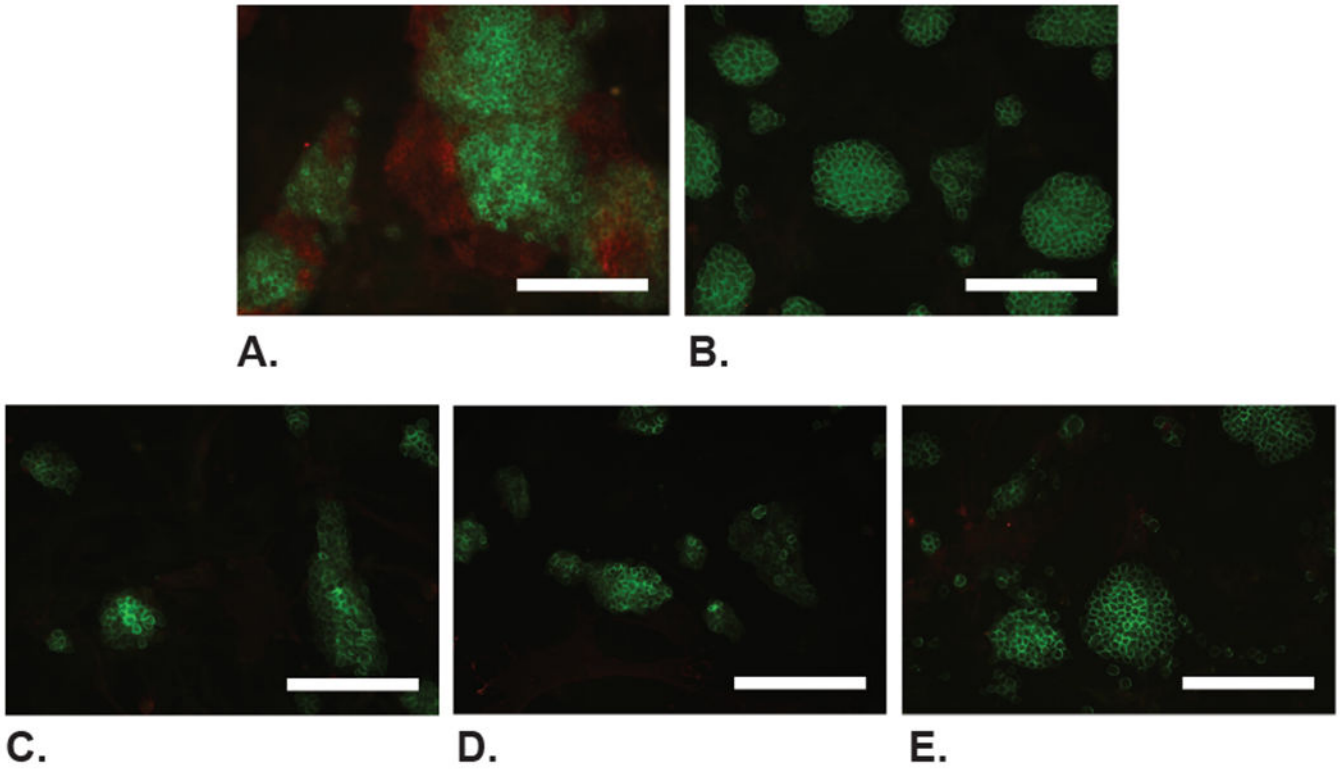

Fig. 3.

Control iPSC cell surface marker expression under different culture conditions. **A** 129S1/SvImJ ESC derived in serum/LIF conditions label heterogeneously for PECAM1 and CD40. **B** 2iS transitions colonies from heterogeneous to homogenous PECAM1 expression. **C,D** iPSC derived in serum/LIF conditions from permissive C57BL/6J and 129S1/SvImJ strains label homogenously for the PCEAM1. **E** iPSC from 129S1/SvImJ continue to label homogenously for PECAM1 after being grown in 2iS conditions for eight passages.

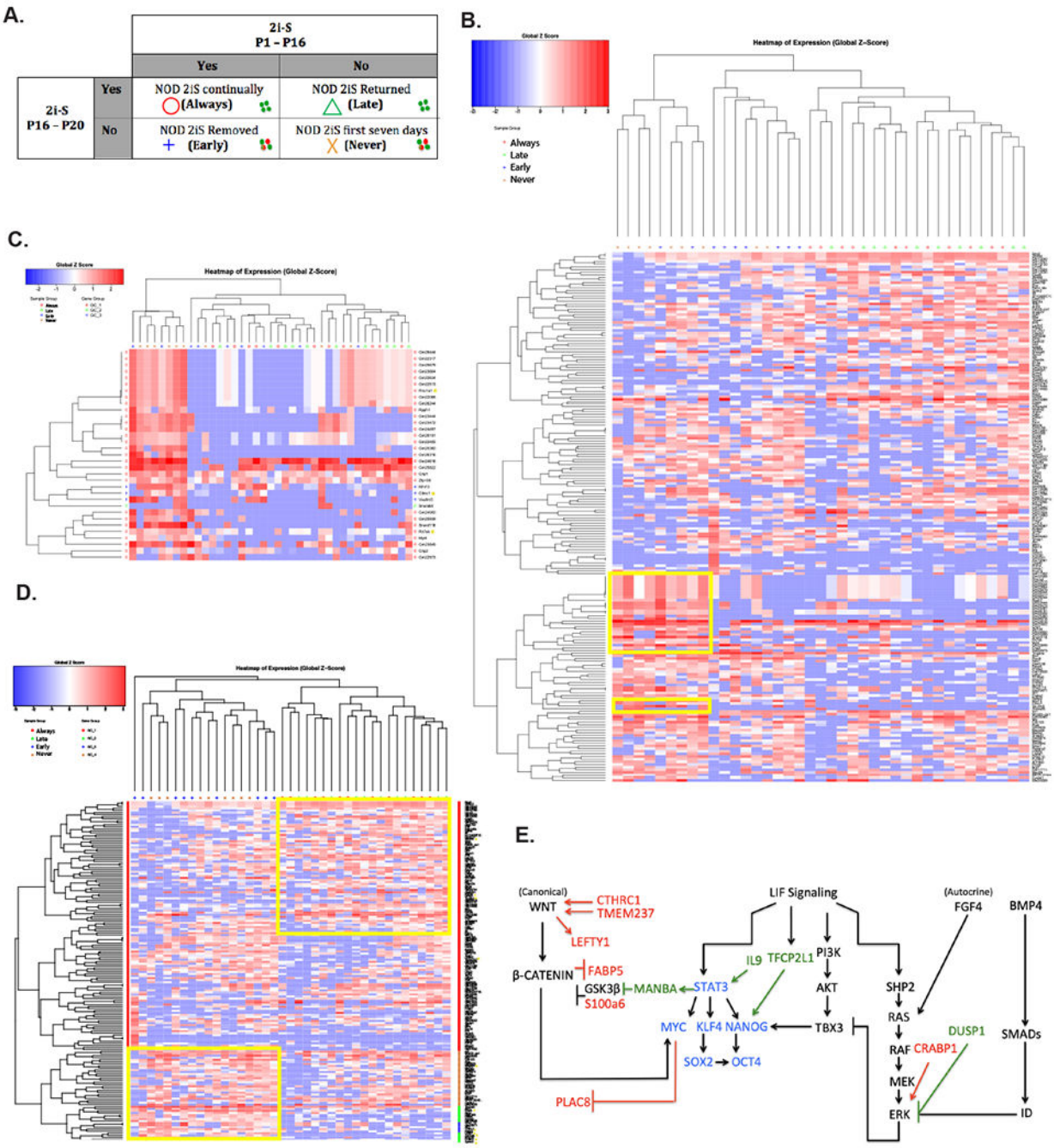

Fig. 4.

Analysis of single cells from each 2iS culture condition **A** To investigate single cell differences we randomly captured cells from iPSC colonies cultured in each 2iS condition: cells from the *Alway*, group were continually cultured in 2iS for 16 passages, cells from the *Late* group were cultured in 2iS after 16 passages, cells from the *Early* group were cultured in 2iS media for 16 passages and was then cultured in serum/LIF conditions lacking 2i, cells from the *Never* group were cultured in 2iS for only the first seven days prior to colony selection and expansion. It is important to note that NOD iPSCs cannot

be derived without some form 2i culture. The *Never* groups represents the closest that we can get at deriving iPSCs without continual supplemental conditions. **B** Single cells from the two homogenous populations cluster together regardless of when 2iS was added (*i.e. Always/Late*). The two heterogenous populations also cluster together (*i.e. Never/Early*). However, a subset of heterogenous cells cluster apart from all the other iPSC samples. The gene group characterizing this cluster is highlighted in yellow. **C** This is a heatmap of the large highlighted cluster in panel B, that separate a subset of heterogenous samples. **D** If we remove this subset separating the two heterogenous samples, single cells separate according to current culture in/out of 2iS media, regardless of individual variations in the timing of 2iS addition. **E** Predicted gene activity of LIF pathways when cells are grown in 2iS. Genes with increased expression in 2iS conditions are indicated in green. Genes with increased expression in 2iS free standard conditions are indicated in red. Genes with increased expression in 2iS media function to further activate the core pluripotency network indicated in blue. TCFP2l1 is a target of the LIF signaling pathway and activates NANOG expression. Homogenous NANOG expression is a key characteristic of populations maintained in 2i media. A side effect of IL9 activation is the activation of the STAT3 pathway, which is central to the core network. MANBA is activated by STAT overexpression and may function like a GSK3β inhibitor. Additionally, in 2iS conditions we see a down-regulation of the ERK differentiation pathway by Dusp1. In contrast, in standard 2iS free conditions CRABP1 may activate the ERK differentiation pathway. WNT signaling is decreased as evidenced by the increased expression of FABP5, which is negatively regulated, by beta-catenin and S100A6, which degrades beta-catenin. Decreased WNT signaling may lead to decreased MYC expression, which is supported by the increased expression of PLAC8 that is negatively regulated by MYC. However, WNT signaling is not ablated as evidenced by LEFTY1 expression, which is upregulated by WNT and TMEM237 and CTHRC1, which activate WNT signaling

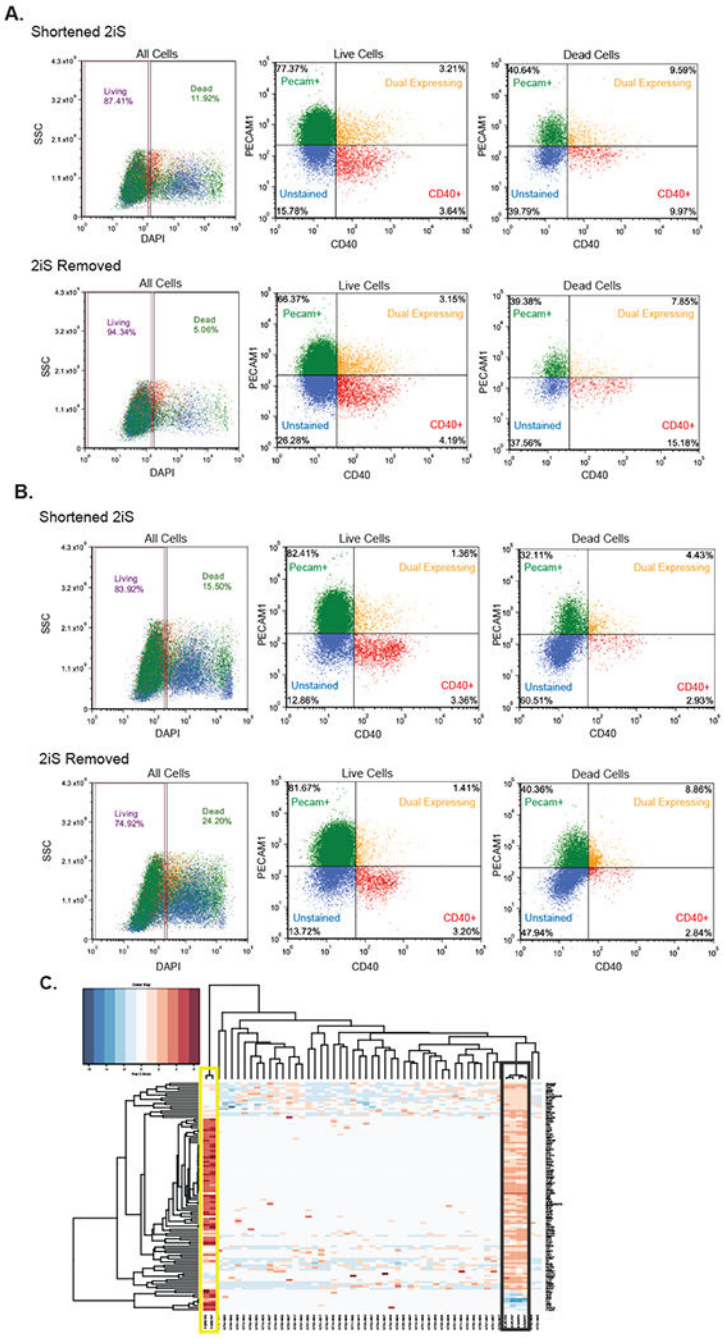

Fig. 5.

Flow cytometry of cells grown under different conditions. **A** The occurrence of each subpopulation in heterogeneous iPSC populations is about the same regardless of the historical length of 2iS exposure. The majority of live cells are PECAM1 positive. **B** Repeat of flow sorting experiment. The majority of live cells are positive for PECAM1. There are slightly less dual expressing cells in this second sort. **C** Flow sorted CD40+ cells cluster separately (highlighted in yellow) from flow sorted PECAM+ cells and flow sorted dual expressing cells. The PECAM1+ cells and dual expressing cells cluster together

with captured single cells. The separate cluster of flow sorted CD40+ cells indicate that no CDO40+ cells were captured in our signal cells analysis.

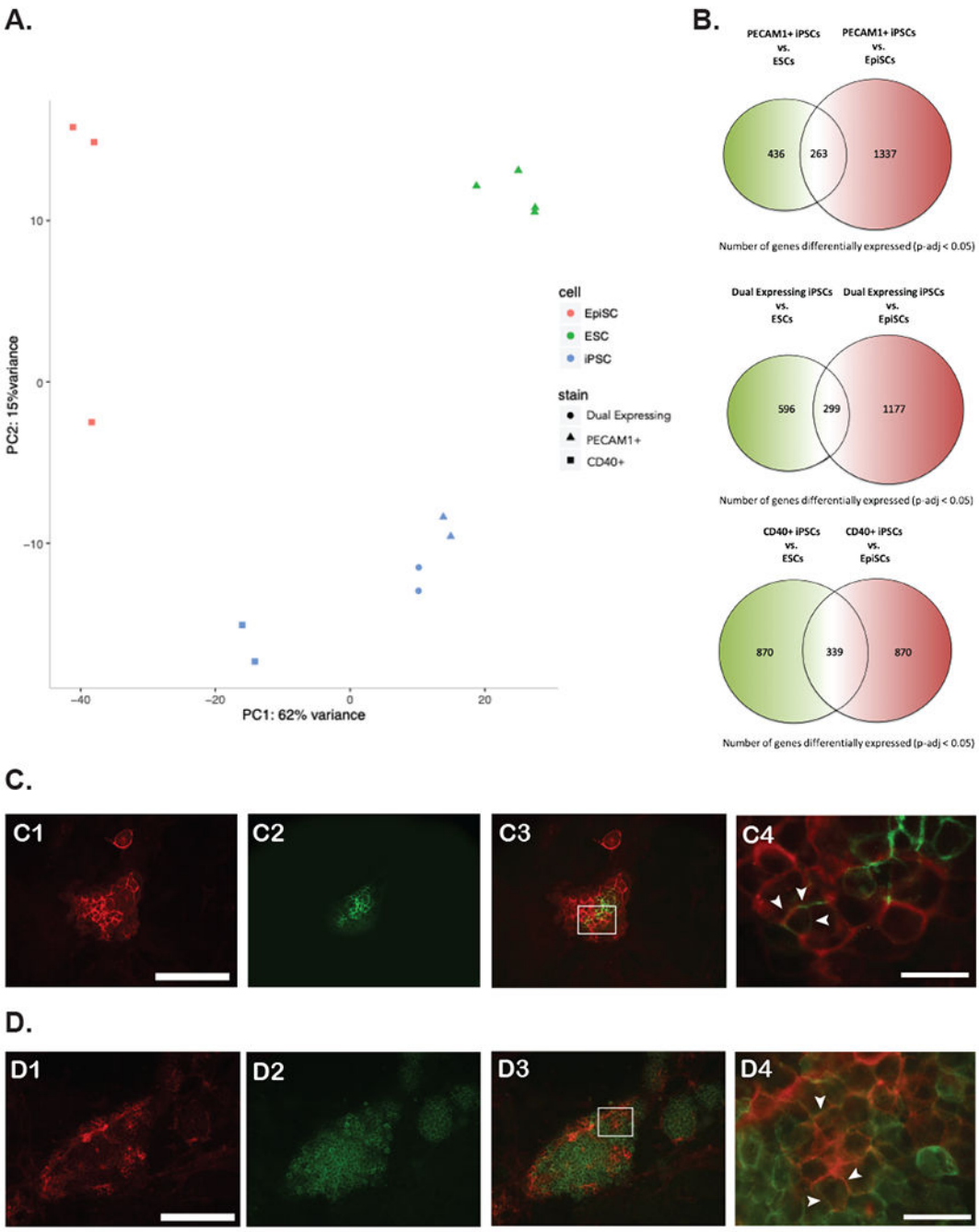

Fig. 6.

Differential gene expression in cells grown under different condition. **A** PCA plot of subpopulation relationships. PECAM1+ cells from both the flow sorted heterogeneous populations cluster together separately from flow sorted CD40+ or dual expressing cells from the same source population. There is less variance between PECAM1+ cells and dual expressing cells than between either and CD40+ cells. **B** There are fewer differentially expressed genes between PECAM1+ cells and the ESC state than PECAM1+ cells and the EpiSC state. The same is true for flow-sorted dual expressing cells. However, CD40+ cells

are equally dissimilar to both the ESC state and the EpiSC state as judged by a roughly equal number of differentially expressed genes in each comparison. **C1-C4** Culture of PECAM1+ cells flow sorted cells separately yields a re-emergence of all three subpopulations. C1-C2 single channel images of each marker. C3 Merged image of both markers. C1-C3 scale bar 200 μm. C4 enlarged region of C3 panel, scale bar 30 μm. **D1-D4** Culture of CD40+ flow sorted cells separately yields a re-emergence of all three subpopulations. These images are representative of the population three to four days after separate growth. The cells have not been passaged.

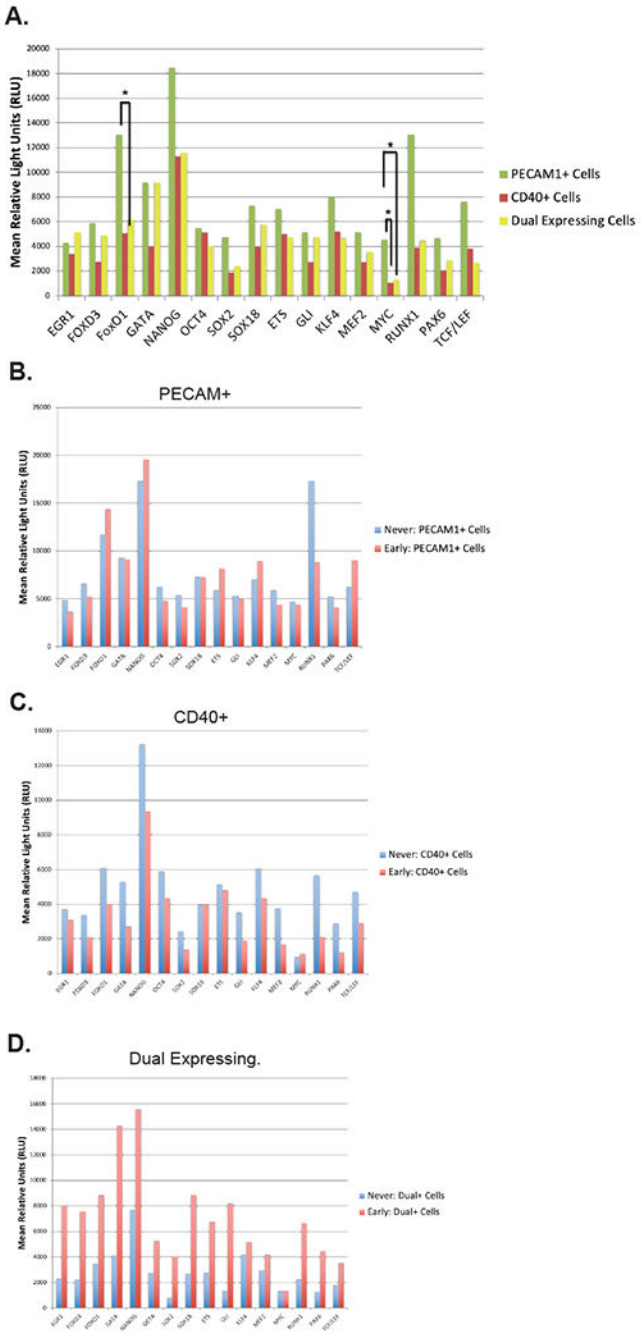

Fig. 7.

Transcription factor activity in cells expressing different markers. **A** FOXO1 production is significantly higher in PECAM1+ cells than in CD40+ cells. MYC production is significantly higher in PECAM1+ cell than in CD40+ cells. It is also significantly higher in PECAM1+ cells than in dual expressing cells. **B, C** There appears to be little difference in the level of each stem cell transcription factor between PECAM1 and CD40+ cells. **D** Dual expressing cells from the 2iS culture group have higher production of all stem cell transcription factors than cells not maintained in 2iS.

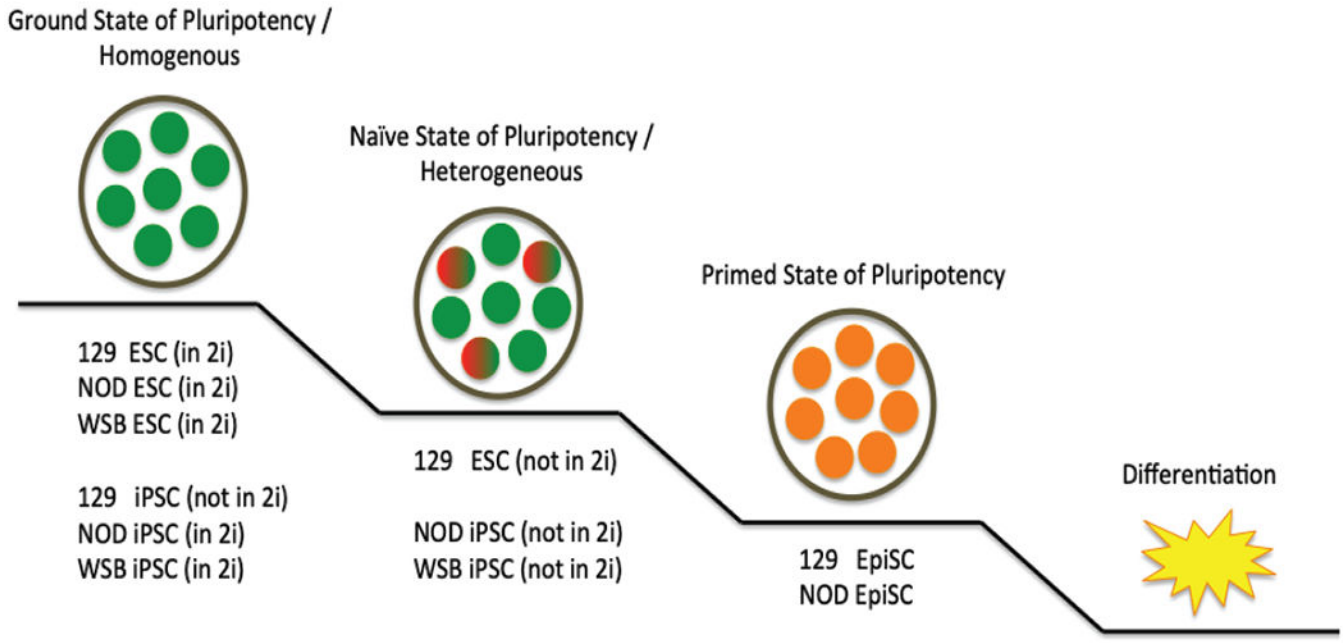

Fig. 8.

Pluripotency potential of stem cells used in this study. At the top of the hill are cells with the most potential. These are transcriptionally homogeneous iPSC colonies. The 2i inhibitors foster a transcriptionally homogeneous PECAM1 positive state, with the exception of iPSC derived from permissive strains such as the 129S1/SvImJ strain that can achieve transcriptional homogeneity in the absence of 2i. ESC from permissive strains and iPSC from non-permissive strains not maintained in 2iS conditions exist at an intermediate level of pluripotency that is both pluripotent and poised for differentiation. This intermediate stage is defined by heterogeneous PECAM1 and CD40 expression. EpiSC are the most poised for differentiation and have the least pluripotency potential of the stem cells studied here. They are the last state of stem cell identity before terminal differentiation. Although CD40 is expressed in EpiSC, it is not unique to the EpiSC state as intermediate ESC also express CD40. A complete loss of pluripotency occurs when cells terminally differentiate.

Table 1

Functional classifications and genes differentially expressed in flow sorted PECAM1+ cells

Function	Genes
Quantity of Cells: p-value 0.000355: Predicted to be increased (bias)	
Increase	Myc, Rasgrp1, Sohlh2, Dmrtc2, Cngb1, Essrb, Pla2g5, Tex19.1, Rlbpl, Tcl1, Zap70, Hormad2, Cardll, Ctcfl, Zp3, Hnfla, Spn, Mybl2, Klf4, Mapt, Klkl, Fgf4, Sox2, Nrobl, Daz1, Tnn, Bhmt, Zfp42, Foxn4
Decrease	Uspl7lb, Piwil2, Dlgap3, Itgal, Ubash3a, Sox15, Cd3e, Gstp2, Nefh, Aire, Pecam1, Inpp5d, Spic, Tet2
Affect	Myolg, Vwf, Mlana, Jamz, Kat6b
Generation of Cells: p-value 3.72E-08: Predicted to be increased (bias)	
Increased	Myc, Rasgrp1, Xlr3c, D1Pas1, Zap70, Itgal, Ctcfl, Zp3, Ace, Hnf1a, Klf4, Uba1y, Tbx3, Ephx2, Dnmt3l, Mapt, Fgf4, Sox2, Fkbp6
Decreased	Cd3e, Klk1, Nefh, Pecam1, Adap1
Affected	Zfp42, Cabp4, Tet2, Tex14, Hormad1, Klf17, Prss42, Morc1, Dcdc2a, Tnn, Stra8, Tex11, Ak7, Dazl, Nlrp14, Inpp5d, Nr0b1, Vwf, Hsf2bp, Trpm 1, Aire, Piwil1, Mcf2, Mov10l1, Smc1b, Spn, Zbtb32, Card11, Hormad2, Aurkc, Hydin, Prdm14, Dlgap3, Piwil2, Tie1, Tex19.1, Rfx2, Dmrtc2, Sohlh2, Tdrd12
Maintenance of Stem Cells: p-value 0.000124: No prediction by IPA	
Affects	Prame17, Tcl1, Prdm1, Prdm14, Esrrb, Tbx3, Fgf4, Klf4, Sox2
Pluripotency of cells: p-value 0.00026: Predicted to be increased (bias)	
Increases	Myc, Tcl1, Hlkf4, Fgf4, Sox2, Fbxo15, Zfp42
Development of the Blastocyst: p-value 0.00026: No prediction by IPA	
Affect	Zfp42, Hormad1, Fgf4, Hnfla, Ubtfl1, Spic, Mybl2, Bhmt, Zp3
Implantation of the Embryo: p-value 0.0384: No Prediction by IPA	
Increases	Ooep
Affects	Prdm14, Nr5a2, Zscan4c, Ubtfl1
DNA Methylation: p-value 0.00737: No prediction by IPA	
Increases	Bhmt, Myc, Dnmt3I
Decreases	Tcl1
Affects	Movlol1, Ctcf1

Table 2

Functional classifications and genes differentially expressed in flow sorted CD40+ cells

Function	Genes
Differentiation: p-value 6.25E-33: Predicted to be increased (bias)	
Increases	Proxl, Hs6st2, Enpp2, Nefm, Lrrc8c, Pde5a, Tgfb3, Cclll, Nsgl, Efnb2, Slpr3, Ednra, Ar, Cd40, Fgf5, Itgb8, Fos, Cnrl, Binl, Scx, Plat, Sparcll, Plekhol, Rorc, Sparc, Ill 8, Col4al, Ebfl, Sema6a, Adcy6, Ifngr2, Isgl 5, Fgf8, Col25al, Sl 00a4, Pgr, Slcl a2, Erbb4, Foxa2, Meoxl, Kctdl l, Adoral, Itga2, Bcl2, Ghr, Mmd, Serpinfl, Cavl, Has2, Pdgfra, Angptl, Zeb2, Mapkl2, Igfbp5, Nkx6-l, Cysltrl, Cd274, Hoxb2, Jun, Nuakl, Bgn, Ptn, Camk4, Dab2, Dbnl, Pam, Camklg, Ltbr, Wnt5a, Mmp9, Lgalsl, Edn3, Enl, Fbn2, Loxl2, Nfia, Efna5, Apln, Pcsk9, F2r, Slcla3, Cdhll, Hoxbl, Pdlim7, Sox5, Rgs2, Irsl, Gab2, Shh, Rorl, Frzb, Tmemll9, Wls, Gdpd2, Sh3pxd2b, Igf2, Soxl7, Ebf2, Rspo3, Ssbp2, Pkdcc, Tox, Eomes, Tpbg, Amot
Decreases	Gdfll, Wnt3, Tbx2, Otx2, Meisl, Hmga2, Mfap2, Chrd, St8sia2, Mafb, Pax3, Sox9, Etsl, Vim, Zhx2, Dkkl, Hoxa9, Nog, Dll3, Abcg2, Dcx, Rgs4, Atf3, Tgfblil, App, Hdac7, Cntn2, Sema3a, Pthlr, Gsn, Twist2, Cd44, Ubash3b, Cd24a, Cxcll0, Nrcam, Irf2, Duspl0, Six2, St8sia4, Sema3f, Ill7rd, Lfng, Zfp52l, Lbh, Robo2, Cyp26al
Affects	Rnasel, Aplnr, Gpm6a, Shc4, Herc6, Lrrk2, Ssbp3, Cadps2, Fzd8, Sema4c, Gpc3, Lztsl, Nrep, Dok5, Pmp22, Cdh2, Fabp7, Psmb8, Dkk2, Fst, Hoxb3, Hoxc8, Ctnna2, Mrc2, Bmpl, Pcsk2, Nrpl, Col6al, Gasl, Creb5, Lpp, Ptpn9, Tagln, Hivep2, Ill5ra, Stk39, Itpkb, Onecut2, Fzd2, Nes, Mixll
Proliferation: p-value: 4.26E-20: Predicted to be increased (bias)	
Increase	Serpinhl, Hs6st2, Enpp2, Capn2, Fat4, Tpm2, Tbx2, Otx2, Meisl, Hmga2, Gbp2, Foxp2, Nedd9, Pde5a, Pls3, Chrd, Tgfb3, Cclll, Stk39, Nsgl, Ill5ra, Efnb2, Nnat, Calcrl, Slpr3, Pbx3, Ednra, Ar, Pax3, Fgf5, Alx4, Fos, Cnrl, Ptpn9, T, Fosb, Scx, Plat, Plekhol, Esml, Tnfaip6, Gasl, Sox9, Etsl, St6gall, Ephal, Pim2, Ill8, Ebfl, Klf6, Sema6a, Vim, Isgl5, Fgf8, Car3, Hoxa9, Sl00a4, Abcg2, Pgr, Slcla2, Erbb4, Itga2, Bcl2, Ghr, Ccnd2, Has2, Pdgfra, App, Angptl, Slc8al, Hoxb3, Etvl, Mapkl 2, Igfbp5, Nkx6-l, Pthlr, Cysltrl, Jun, Cd44, Ptn, Cdkl4, Camk4, Sh3kbpl, Ltbr, Wnt5a, Mmp9, Hsdllb2, Lgalsl, Cd24a, Dclkl, Edn3, Nrcam, Irf2, Cdh2, Nfia, Duspl0, Apln, Six2, Dok5, Erapl, Gria3, Nrep, F2r, Setbpl, Slcla3, Cdhll, Lztsl, Hoxbl, Sema4c, Rgs2, Antxrl, Irsl, Gab2, Shh, Rorl, l8l00llOl0Rik, Asb4, Tmemll9, Zfp52l, Collal, Ssbp3, Zfp703, Wls, Zfp503, Igf2, Shc4, Rabl3, Aplnr, Tox, Eomes, Amot
Decreases	Unc5c, Proxl, Gdfl l, Wnt3, Srgn, St8sia2, Cd40, Tagln, Itgb8, Binl, Dachl, Peal 5a, Sparcll, Cnnl, Taxl bp3, Rorc, Myl9, Sparc, Nrpl, Col4al, Pcsk2, Adcy6, Nog, Igfbp4, Rprm, Dcx, Setd7, Rgs4, Foxa2, Atf3, Kctdl l, Pex2, Tgfblil, Gal, Serpinfl, Cavl, Col4a2, Hdac7, Zeb2, Adamtsl, Rgl3, Sema3a, Emp3, Fst, Gsn, Cd274, Bgn, Emilinl, Dab2, Cxcll0, Fbn2, Efna5, Pmp22, St8sia4, Sema3f, Ill7rd, Gpc3, Frzb, Ccdc8, Cerl, Soxl7, Eif4e3, Rnasel, Gbp7
Affects	Fbxl7, Gpm6a, Robo2, Lrrk2, Dchsl, Gprc5c, Fzd8, Sox5, Kcnj5, Pxdn, Cacnalg, Loxl2, Cldnl, Fabp7, Lrp2, Dbnl, Nuakl, Twist2, Cntn2, Adoral, Pcdhl, Bmpl, Dkkl, Maml3, Col6al, Ms4a4d, Ndnf, Vcaml, Mafb, Cav2, Nefm, Psme2, Add2, Zyx
Cell Death: p-value 1.05E-18. Predicted to be decreased	
Decreases	Zyx, Serpinhl, Gdfll, Eya4, Wnt3, Fat4, Aphlb, Otx2, Nefm, Meisl, St8sia2, Scg5, Ill5ra, Mafb, Efnb2, Calcrl, Slpr3, Pbx3, Ednra, Pax3, Fgf5, Alx4, Ptpn9, T, Scx, Plat, Dachl, Peal5a, Plekhol, Ndnf, Foxbl, Rorc, Sox9, Ptprt, St6gall, Edil3, Pim2, Sparc, Col6al, Nrpl, Col4al, Ebfl, Klf6, Sema6a, Adcy6, Ifngr2, Cdh6, Fgf8, Car3, Hoxa9, Col25al, Igfbp4, Sl00a4, Abcg2, Pgr, Slcla2, Erbb4, Rgs4, Foxa2, Parva, Adoral, Itga2, Bcl2, Ghr, Pde4b, Gal, Serpinfl, Cavl, Has2, Pdgfra, Gpm6b, Angptl, Zeb2, Slc8al, Pthlr, Fst, Gsn, Twist2, Nuakl, Bgn, Cd44, Lrp2, Wnt5a, Hsdllb2, Psmb8, Irf2, Anxa5, Cdh2, DusplO, Apln, Gria3, St8sia4, Slcla3, Cdhll, Lztsl, Actcl, Sema3f, Pdlim7, Gpc3, Sox5, Irsl, Gab2, Shh, Atpla2, Rorl, l8l00llOl0Rik, Cadps2, Collal, Sepnl, Igf2, Lrrk2, Astnl, Tox3, Adgrl2, Tox, Eomes, Itprip, Amot
Increases	Peg3, Unc5c, Gulpl, Capn2, Srgn, Nedd9, Pde5a, Gpr37, Chrd, Tgfb3, Cclll, Itpkb, Ifit2, Nnat, Ar, Cd40, Fos, Cnrl, Binl, Fosb, Vcaml, Arap3, Etsl, Pcdhgc3, Ill8, Isgl5, Dkkl, Bmpl, Tgfbi, Nog, Rprm, Atf3, Ifihl, Tgfblil, Ccnd2, Col4a2, App, Adamtsl, Mapkl2, Igfbp5, Sema3a, Emp3, Cd274, Jun, Sall2, Ptn, Stra6, Camk4, Cacnalh, Dab2, Cacnb2, Sh3kbpl, Ltbr, Mmp9, Lgalsl, Cd24a, CxcllO, Enl, Cacnalg, Pmp22, Pcsk9, F2r, Ill7rd, Lhxl, Antxrl, Frzb, Ccdc8, Isml, Pcdhgc4, Pcdhgc5, Gpm6a, Rnasel, Antxr2
Affects	Shc4, Soxl7, Car4, Hoxbl, Snph, Efna5, Loxl2, Fbn2, Nrcam, Lspl, Ctnna2, Zhx2, Vim, Pcsk2, Gasl, Ms4a4d, Fzd2, Nes, Tbx2, Psme2, Add2