To: *bioRxiv*

January 29, 2024

Resistance to Naïve and Formative Pluripotency Conversion in RSeT Human Embryonic Stem Cells

Kevin G. Chen[1]*, Kory R. Johnson[2], Kyeyoon Park[1], Dragan Maric[3], Forest Yang[1], Wen Fang Liu[1], Yang C. Fann[2], Barbara S. Mallon[1], and Pamela G. Robey[4]*

[1]NIH Stem Cell Unit, [2]Intramural IT and Bioinformatics Program, [3]Flow and Imaging Cytometry Core Facility, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, Maryland 20892, USA

[4]Skeletal Biology Section, National Institute of Dental and Craniofacial Research, National Institutes of Health, Bethesda, Maryland 20892, USA

Keywords: human embryonic stem cells, RSeT medium, transcriptome, naïve pluripotency, cell proliferation

***Correspondence:** cheng@mail.nih.gov (K.G.C.), probey@dir.nidcr.nih.gov (P.G.R)

Words: excluding Figure Legends and References (6635)

ABSTRCT

One of the most important properties of human embryonic stem cells (hESCs) is related to their primed and naïve pluripotent states. Our previous meta-analysis indicates the existence of heterogeneous pluripotent states derived from diverse naïve protocols. In this study, we have characterized a commercial medium (RSeT)-based pluripotent state under various growth conditions. Notably, RSeT hESCs can circumvent hypoxic growth conditions as required by naïve hESCs, in which some RSeT cells (e.g., H1 cells) exhibit much lower single cell plating efficiency, having altered or much retarded cell growth under both normoxia and hypoxia. Evidently, hPSCs lack many transcriptomic hallmarks of naïve and formative pluripotency (a phase between naive and primed states). Integrative transcriptome analysis suggests our primed and RSeT hESCs are close to the early stage of post-implantation embryos, similar to the previously reported primary hESCs and early hESC cultures. Moreover, RSeT hESCs did not express naïve surface markers such as CD75, SUSD2, and CD130 at a significant level. Biochemically, RSeT hESCs exhibit a differential dependency of FGF2 and co-independency of both Janus kinase (JAK) and TGFβ signaling in a cell-line-specific manner. Thus, RSeT hESCs represent a previously unrecognized pluripotent state downstream of formative pluripotency. Our data suggest that human naïve pluripotent potentials may be restricted in RSeT medium. Hence, this study provides new insights into pluripotent state transitions *in vitro*.

INTRODUCTION

Naïve or ground pluripotent states were initially proposed by Smith and colleagues based on studies in mouse embryonic stem cells (mESCs) from pre-implantation mouse embryos (Nichols and Smith, 2009; Ying et al., 2008) and on studies of lineage-primed epiblast stem cells (EpiSCs) derived from the post-implantation mouse epiblast (Brons et al., 2007; Tesar et al., 2007). The establishment of human naïve pluripotency is believed to have significant impacts on understanding early human embryonic development, facilitating single human pluripotent stem cell (hPSC) growth *in vitro*, genetic manipulation, disease-modeling, optimizing interspecies chimerism, and drug discovery [reviewed in (Dodsworth et al., 2015; Dong et al., 2019; Hanna et al., 2010; Weinberger et al., 2016; Wray et al., 2010; Wu et al., 2017; Zimmerlin et al., 2017)].

In the past decade, numerous groups reported the conversion of primed hPSCs to naïve hPSCs or direct derivation of them from the human inner cell mass (ICM) (Chan et al., 2013; Collier et al., 2017; Gafni et al., 2013; Guo et al., 2016; Kilens et al., 2018; Liu et al., 2017; Takashima et al., 2014; Theunissen et al., 2016; Ware et al., 2014; Zimmerlin et al., 2016). Despite the enthusiasm of deriving naïve hPSCs with diverse growth conditions, there are increasingly intense efforts to derive naïve hPSCs based on the use of the commercial medium RSeT (www.stemcell.com) (Collier et al., 2017; Kilens et al., 2018; Liu et al., 2017; Szczerbinska et al., 2019). Our meta-analysis indicated that RSeT-based hPSCs have a closer transcriptomic similarity to primed and formative hPSCs than other “naïve” counterparts derived from different laboratories (Johnson et al., 2021). Likely, RSeT hPSCs provide an important clue to unraveling human naïve and formative pluripotency in development and in cell culture platforms. However, the cellular properties and pluripotent states in RSeT hPSCs under different growth conditions are not well defined or characterized.

In this study, we report the conversion and more detailed characterization of RSeT hESCs, including transcriptomic analysis of primed and naïve gene expression, diagnostic cell surface marker expression, and the response to core signaling molecules of both primed and naïve pluripotency. We aim to provide a systemic evaluation of RSeT conditions and resolve some experimental and conceptual inconsistencies concerning human naïve and formative pluripotency (i.e., a transition phase between the naïve and primed states of epiblast) (Smith 2017). Our data reveal that RSeT medium does not effectively support the conversion of primed hESCs to a naïve state. The converted RSeT hESCs have a distinct and previously unrecognized pluripotent state between the formative and primed states, displaying insensitivity to oxygenic tension alterations, altered response to core signaling molecules, and significantly different cell growth potentials.

RESULTS

Conversion of primed hESC lines to RSeT states without the hypoxic growth requirements for naïve hESCs

RSeT medium is based on naïve human stem cell medium (NHSM) (Gafni et al., 2013), but with a different formulation (www.stemcell.com). Our investigation revealed that RSeT-converted hPSCs exhibit a transcriptomic profile akin to primed hPSCs (Johnson et al., 2021), although it was not previously recognized (refer to Figure 1A). To fully characterize RSeT hPSCs, we converted three primed hESC lines (H1, H7, and H9) under RSeT growth conditions (Figure 1B), with the initial use of the ROCK inhibitor Y-27632 (ROCKi) to enhance single-cell plating efficiency (SCPE).

The converted RSeT hESCs displayed a slow growth pattern with morphologically small domed-shaped colonies as described for naïve hPSCs (Figure 1C). However, we found that all three hESC lines behaved differently under the RSeT conditions. In general, H9 cells were the easiest to be induced into and maintained, as domed colonies (Figure 1C, bottom panel), followed by H7 cells (Figure 1C, middle panel). H1 cells displayed intractability to being induced and maintained as domed colonies under these conditions (Figure 1C, top panel). Upon removal of ROCKi after 3 passages, H1 cells exhibited extremely poor growth that was even exacerbated under hypoxia (i.e., 3% O_2). For example, H1 cells exhibited a 6.7-fold reduction in colony formation at passage 10 under RSeT and hypoxic growth condition (RH) (Figure 1C, top panel). Interestingly, RSeT hESCs typically propagate well in normoxic conditions compared with the same cells under hypoxia (Figure 1C, RN and RH columns). However, H7 colonies grown in RSeT under normoxia (RN) showed some heterogeneous morphologies (Figure 1C, middle panel). Nonetheless, the data suggest that RSeT hESCs can bypass the hypoxic growth requirement for naïve hPSCs.

RSeT hESCs transcriptomically similar to their primed counterparts

Globally, RSeT hESCs have a transcriptome similar to their primed hESCs counterparts based on microarray profiling of gene expression in this study (Figure 2A, Figure S1). Noticeably, both primed and RSeT transcriptomes comprise the majority (> 65%) of transcripts with low to moderate mRNA expression. Only a small portion of the mRNA transcripts were highly expressed in RSeT cells (Figure 2A). We further utilized principal component analysis (PCA) to define the differences between individual cell lines. PCA transforms a large number of variables into a smaller number of orthogonal variables (e.g., principal components 1, 2, and 3, abbreviated as PC1, PC2, and PC3, respectively). Hence, 96.1% of gene expression variations, in which PC1 and PC2 constitute 88.5% and 7.6% variability respectively, might be used to

depict the differences between primed and RSeT states (Figure 2B). Moreover, H9 cells showed a similar transcriptome to H1 cells under primed growth conditions, but closer to H7 than H1 under RSeT conditions (Figure 2A; 2B, left panel). Interestingly, only a small percentage (i.e., 2.4%) of gene expression variations was responsible for the response to normoxic and hypoxic growth (Figure 2B, right panel), which may be associated with the enhanced growth capacity of RSeT hESCs under normoxia as described above.

RSeT hESCs lack naïve and formative gene markers

To define naïve and formative gene expression, we implemented a supervised analysis of a 112-gene list (Table S1), derived from previous mouse and human pluripotent stem cell reports or reviews (Brons et al., 2007; Chan et al., 2013; Collier et al., 2017; Gafni et al., 2013; Liao et al., 2015; Mallon et al., 2013; Pera and Rossant, 2021; Smith, 2017; Takashima et al., 2014; Tesar et al., 2007; Theunissen et al., 2016; Valamehr et al., 2014; Weinberger et al., 2016), which defines tangible primed and naïve transcriptomes. Briefly, this list includes the ICM and other developmental markers that are not expressed in the implantation epiblast (n = 5), naïve or epiblast-specific genes (n = 60), formative (between pre-implantation and the E6.5 epiblast in mouse) genes (n = 5), primed or metastable lineage specifier genes (n = 37), and other X-linked or imprinted gene expression (n = 5).

Overall, there are approximately 45% of genes within the 112-gene cluster that had altered expression when converted to RSeT cells. Some of the most notable human naïve signature genes (e.g., *DNMT3L*, *DPPA2*, *DPPA5*, *TFCP2L1*, *XIST*, and *ZFP42)* remained unchanged by microarray assays (Figure 2C, left panel; Table S1). Also, there was only 4.3% of genes altered between primed and RSeT states (Table S1), which is consistent with genome-wide microarray analysis showing the limited alterations of gene expression between the two states of hESCs (Figures 2A, 2B).

Notably, only a few naïve pluripotency genes (e.g., *KLF4*) were upregulated (Figure 2C, Tables S1). Interestingly, *NANOG* was downregulated in RSeT cells in both RSeT normoxia (RN) and RSeT hypoxia (RH) conditions (Figure 2C, left panel), which was confirmed by quantitative real-time PCR (qPCR) analysis using two different *NANOG* probes (Figure 2D, n = 6 cell lines, $P < 0.05$). *POU5F1*, another important pluripotency marker gene, was also significantly suppressed in RSeT hESCs as determined by qPCR (Figure 2D, n = 6, $P < 0.05$). Moreover, the expression of *GDF3* and *NODAL,* which were highly expressed in the naïve epiblast of human embryos (Blakeley et al., 2015; Stirparo et al., 2018), was also significantly reduced in RSeT hESCs (Figure 2D, n = 6, $P < 0.05$). These results are clearly opposite to that which was observed in reported naïve hPSCs.

Moreover, several formative pluripotency associated genes (e.g., *DNMT3A*, *DNMT3B, OTX2, SALL2, and SOX3*) and imprinted gene clusters (e.g., *DLK1*, *MEG3*, and *MEST*) had no significant alterations (Figure 2C, middle and right panels; Table S1). Other down-regulated primed genes include *CER1, EGR1, GSC, LEFTY1, NODAL, and ZIC2* (Figure 2C, Table S1). Some of which (e.g., *CER1*, *LEFTY1*, and *NODAL*) were also confirmed by our qPCR analysis (Figure 2D). Taken together, RSeT hESCs do not vary significantly in terms of their global gene signatures when compared with their primed parental cells. Specifically, RSeT hESCs lack human naïve and formative transcriptomic hallmarks as described in other human naive hPSC models, which makes it hard to ascertain the developmental stage of RSeT hESCs.

Meta-analysis revealed a unique transcriptome and plausible developmental stage of RSeT hESCs.

We utilized a comparative meta-analysis based on both PCA and *t*-distributed stochastic neighbor embedding (*t*-SNE) to resolve the above issue (Figure 3; Figures S2 and S3). This analysis revealed that RSeT hESCs had a minimal transcriptomic difference compared with their primed counterparts (Figures 3A-C), which were in contrast to some established naïve hPSCs, particularly to those derived by t2iLGo and 5iLAF protocols (Takashima et al., 2014; Theunissen et al., 2016) (Figures 3A-C).

To directly relate cultured hPSCs to a developmental stage represents a significant challenge because of the presence of substantial batch effects between cultured hESCs and embryo datasets. However, we were able to maximally reduce the inter-laboratory and cell-culture batch effects (Johnson et al., 2021). This was further confirmed by including more primate embryo datasets, where all the primate embryo data are clustered closer together (Figure 3B, S3). Notably, the primed hESCs maintained in our laboratory are significantly different from their primed counterparts from several different laboratories (Figure 3B). Thus, we can associate our primed and RSeT hESCs with the early stage of post-implantation embryos, similar to the previously reported primary hESCs (at passage 0) and early hESC cultures at passage 10 (Figures 3B, S3) (Yan et al., 2013). These primary hESCs were derived from the ICM of blastocysts at day 6 (E6) with the least cell-culture-related changes (Yan et al., 2013).

Consequently, the findings of this analysis suggest that both primed and RSeT hESCs cultivated in our laboratory exhibit a similar transcriptomic state, and further indicate that these cells are not significantly affected by in vitro cell culture artifacts. Nonetheless, this observation emphasizes the presence of heterogeneous primed states under various cell culture conditions. It is worth noting that the primed hESCs grown in our laboratory may not be completely primed

due to their resemblance to primary hESCs (Yan et al., 2013) (Figure 3B, S3), potentially harboring some unprimed naïve hESCs.

Feeder-supported hESCs are not fully primed and heterogeneously express naïve-specific surface markers

We initially showed that RSeT hESCs highly expressed several primed pluripotency surface markers such as SSEA-4, Tra-1-60, and Tra-1-81 (Figure 4A). To verify the primed state of hESCs, we also examined the primed hESCs cells with several well-characterized primed-specific markers (e.g., CD24, CD57, and CD90) and naïve-specific markers (e.g., CD75, SUSD2, and CD130) (Bredenkamp et al., 2019; Collier et al., 2017) by flow cytometric analysis of Hoechst 33258-gated single cells. Surprisingly, the low oxygen level led to the partial expression of CD57 (39% in H1 and H7, and 45% in H9 cells), CD24 (58% in H1, 29% in H7, and 21% in H9 cells), and CD90 (41% in H1, 35% in H7, and 51% in H9 cells) in primed hESCs (Figures 4B-D, 5A1).

Moreover, we found a very low percentage (0.5 to 3.1%) of primed H1 and H7 hESCs expressing naïve-specific markers CD75, SUSD2, and CD130 (Figures 4B-D). Unexpectedly, 14% of primed H9 cells express the naïve-specific marker SUSD2 (Figure 5A1). These data suggest that hESCs maintained on mouse embryonic fibroblasts (MEFs) in hESC medium are likely in a partially primed pluripotent (PPP) state with a heterogeneous expression of naïve-specific markers.

RSeT hESCs are deficient in the expression of naïve-specific surface markers

To further examine the naïve state in RSeT hESCs, we employed the same set of primed- and naïve-specific markers to assess the heterogeneity of RSeT cells (Figures 4B-D, 5B, 5C). When primed hESCs were switched to RSeT medium, we observed a 15% increase in CD75, 17% in SUSD2, and 0.1% in CD130 expression in RSeT H1 cells (Figures 4B-D, second row). Evidently, only a 4% increase was found in CD75, with no increase in both SUSD1 and CD130 expression in RSeT H7 cells (Figure 4B-D, the fourth row). In the case of H9 cells, there is a 9.5% increase in CD75 in RSeT H9 cells, concomitantly with decreased SUSD2 (Figure 5A, middle panel). Moreover, the expression of CD75 tended to decrease under the normoxic conditions (Figure 5A, top panel).

To gain a deeper understanding of the transitions from primed to naïve states, we examined the changes in three pairs of double positive subpopulations, $CD57^{+}CD75^{+}$, $CD24^{+}SUSD2^{+}$, and $CD90^{+}CD130^{+}$. These three pairs were elevated by 7.6%, 4.1%, and 1.1%, respectively, in RSeT hESCs when compared with primed controls (n = 3 cell lines) (Figure 5D). Taken

together, our data suggest that only a subset of H1 cells responded to naïve conversion with RSeT medium. However, the CD75$^+$ and/or SUSD2$^+$ subsets in RSeT H1 cells failed to express the LIF coreceptor CD130/IL6ST (Figure 4D). Hence, the RSeT medium has a limited ability to reprogram primed hESCs into a complete naïve state (Figures 5B-E). This might be due to the absence of some core signaling transduction pathways (e.g., the LIF coreceptor pathway) required to sustain the naïve state.

Primed hESCs display differential sensitivity to the inhibition of core signal transduction pathways

To determine the response of both primed and RSeT hESCs to core signaling factors, we initially evaluated primed hPSC growth patterns in three hESC lines (H1, H7, and H9) under both normoxic and hypoxic conditions. We found that hypoxia did not significantly increase hESC colony numbers (data not shown), suggesting that hypoxia alone is not a major factor that affects the growth patterns of primed hESCs. It is well known that primed hPSCs require FGF2/TGFβ/Nodal/Activin A for cell proliferation and self-renewal (James et al., 2005; Thomson et al., 1998; Vallier et al., 2005; Xu et al., 2005). It is expected that some core signaling pathway-related growth factors might regulate hPSC growth potential.

We next examined the requirements of some key signaling pathways (e.g., FGF2 and TGFβ) in primed hESCs under both normoxic and hypoxic conditions (Figures 6A-C). Inhibition of FGF2 by its specific inhibitor, PD173074, indeed resulted in differential colony growth in the three hESC lines, in which H1 cells were sensitive to FGF2 inhibition under both normoxic and hypoxic conditions (Figure 6A, lanes 1-4, 13-16). H7 cells showed a moderate FGF2 inhibition under the above two conditions. However, H9 cells were susceptible to FGF2 inhibition only under Primed and Normoxic growth conditions (PN) (Figure 6A, lanes 9-12). This inhibitory pattern in H9 cells was partially restored by the FGF2 inhibitor under the Primed and Hypoxic growth condition (PH) (Figure 6A, lanes 21-24).

With respect to the effect of TGFβ inhibition, the three hESC lines exhibited similar growth patterns under PN conditions, and relatively insensitive to TGFβ inhibition mediated by the TGFβ inhibitor SB431542, in the range of 0.5-2 μM range (Figure 6B, lanes 1-12). Interestingly, H1 cells appeared not to be affected by TGF-β inhibition in both PN and PH conditions, an unexpected result that has not been previously reported (Figure 6B, lanes 1-4, 13-16). However, both H7 and H9 cells showed a dose-dependent TGFβ inhibition only under PH (Figure 6B, lanes 17-24).

To examine the intrinsic naïve growth capacity in primed hESCs, we performed a Janus kinase I (JAK) inhibitory experiment under the assumption that naïve ES colonies (e.g., mESCs)

would be greatly reduced or eliminated due to their hypersensitivity to JAK inhibition (Tesar et al., 2007). As shown in Figure 6C, the three hESC lines showed a similar JAK inhibition under 0.5 μM JAK inhibitor I (JAKi) treatment (Figure 6C, lanes 1-12). However, H1 exhibited much lower sensitivity to JAKi under PH, which is in contrast with both H7 and H9 cells, which manifested an enhanced response to JAKi (Figure 6C, lanes 13-24).

Collectively, we identified several different intrinsic properties peculiar to these three primed hESC lines. The differential requirement of FGF2 in primed hESCs is both cell-line- and oxygen-tension-specific. Compared with both H7 and H9, H1 cells are relatively sensitive to FGF2 inhibition and resistant to JAKi, two of the hallmarks of primed pluripotency. These data suggest that the intrinsic naïve-compatible activity is higher in both H7 and H9 than H1 cells. This differential naïve signaling compatibility might enable H7 and H9 cells to readily adapt themselves in some naïve-like growth media such as RSeT medium (Figure 1C, middle and bottom panels). Thus, the intractability of primed H1 cells to be induced and maintained as domed colonies might also be consistent with the lower intrinsic JAK activity, which is associated with the LIF-CD130/IL6ST-JAK pathway (Collier et al., 2017) in these H1 cells (Figure1C, top panel; Figure 6C, lanes 13-15).

Heterogeneous growth of RSeT hESCs depicted by their differential sensitivity to the inhibition of core signal transduction pathways

In parallel, we also determined the response of RSeT H1 cells to core signaling factors in SCPE assays. Consistently, H1 SCPE under RN conditions was undetectable (Figure 6D, lanes 1-3). It was also difficult to obtain H1 cell colonies under RH conditions for a comparative analysis in this assay due to their extremely slow growth rate (Figure 6D, lanes 13-16). Noticeably, under normoxia, both H7 and H9 RSeT cells were completely resistant to JAK inhibition, highly sensitive to FGF2 inhibition, and moderately sensitive to TGFβ inhibition (Figure 6D, lanes 5-12). Similar signaling response patterns were also observed in these RSeT cells under hypoxia (i.e., RH conditions), except that the partial sensitivity to TGFβ inhibition in these cells was abrogated, comparable with the levels of untreated controls (Figure 6D, lanes 17-24). Of note, compared with hypoxia, a significantly higher number of colonies formed in RSeT H7 cells under normoxia (Figure 6D: lanes 5-8, 17-20), further demonstrating the hypoxia-independent cell growth feature in RSeT hESCs.

DISCUSSION

Increasing evidence suggests the existence of a broad spectrum of pluripotency between naïve and primed pluripotency among different mammalian species (Rossant and Tam, 2017; Wu et

al., 2017). Various hPSC models derived *in vitro* cell culture systems may echo these distinct pluripotent states (Figures 1A, 3, and 7). It is possible to achieve a naïve pluripotent state in hPSCs, as indicated in both t2iLGo and 5iLA models (Figures 2C, 7A). The t2iLGo protocol utilizes titrated 2iL in the presence of the PKC inhibitor Go6983 and transgene expression of *KLF2* or *NANOG* (Takashima et al., 2014) or OSKM (i.e., the combination of OCT4, SOX2, KLF4, and c-MYC) (Liu et al., 2017) (Figure 3C). In a revised version of t2iLGo without transgene expression, known as PXGL (Bredenkamp et al., 2019; Guo et al., 2017), GSK3 inhibition is replaced by the tankyrase inhibitor, XAV939. This modified version is now being widely used to promote the establishment of the human naïve pluripotent state (Kagawa et al., 2022; Khan et al., 2021; Zhou et al., 2023). Moreover, the 5iLA protocol evolved from the 2iL protocol into a five-inhibitor-based protocol, which impedes MEK, GSK3, ROCK, BRAF, and SRC pathways (in the presence of LIF and activin), to orchestrate profound genomic and epigenomic changes in primed hPSCs (Collier et al., 2017; Theunissen et al., 2016).

However, several naïve growth methods only enable the emergence of some LIF-dependent signatures similar to those of native preimplantation epiblasts (Chan et al., 2013; Cornacchia et al., 2019; Gafni et al., 2013; Valamehr et al., 2014; Zimmerlin et al., 2016). Of note, without the expression of powerful transgenes, the naive conversion with t2iLGo would be compromised, possibly ending up with a primed-to-naïve intermediate state (Figure 3C) (Guo et al., 2016). The pluripotent states in these hPSC models are fundamentally similar to the executive phase of pluripotency, coined as the formative state (Smith, 2017).

In this study, we have redefined and characterized RSeT-based hESCs, which exhibit several new features that have not been previously characterized. These features include hypoxia-independent cell growth, a unique pluripotent state with heterogeneous expression of primed- and naïve-specific markers, differential dependency of FGF2, JAK, and TGFβ signaling in a cell-line-specific manner. These prominent characteristics distinguish these cells from those of primed and naïve hPSCs, which warrants further discussion as follows.

The converted hESCs appear to acquire a specific state downstream of the formative state with differential growth potentials in individual RSeT hESC lines (Figures 3C, 7A). Interestingly, either RSeT medium-converted (cRSeT) cells or RSeT plus transgene (OKSM) reprogrammed (rRSeT) cells show global similarity to primed hESCs (Figure 3C) (Johnson et al., 2021; Liu et al., 2017). Also interestingly, rRSeT cells are resistant to transgenic reprogramming effects, suggesting that the restricted reprogramming capacity in RSeT hPSCs is difficult to circumvent even in the presence of powerful transgene expression. These data further support the unique pluripotent state in RSeT cells, which possess limited pluripotent plasticity and lower naive reprogramming efficiency.

Thus, the unique pluripotent state in RSeT hESCs might help to address several fundamental developmental questions on the divisibility of the pluripotent window and the state plasticity of hPSCs. Indeed, RSeT hESCs represent a novel state characterized by FGF2-dependency and JAK- and TGFβ co-independency (Figure 6), between the formative and primed states, as evidenced by our detailed characterization and genome-wide profiling (Figures 1-6). It was also shown that RSeT cells did not undergo X chromosome reactivation, which was unveiled by RNA FISH, a more conclusive assay (Vallot et al., 2017). Consistently, we reveal that RSeT hESCs do not acquire many naïve pluripotency hallmarks such as global hypomethylation and transcriptomic signatures as documented in multiple naive protocols.

Additionally, substantial differences between laboratories in the way they handle primed hPSC cultures can significantly affect the outcomes of converting them to naïve states in vitro and make data interpretation confusing (Johnson et al., 2019). The H9 cell line is an example of this (Johnson et al., 2021). H9 cells have been widely used for hPSC genetic engineering and lineage differentiation (www.wicell.org) (Johnson et al., 2021; Mallon et al., 2013; Takashima et al., 2014; Wang et al., 2011). Primed H9 cells might have acquired substantial genetic and epigenetic changes to adapt themselves to various pluripotent states or differentiation conditions. Hence. it is conceivable that H9 cells exhibit much higher growth advantage and adaptability under various RSeT conditions. A high-resolution genome sequencing may provide a definite answer in the future. The primed hESC lines used in this study have a PPP state and closely resemble primary hESCs. This might reflect a more physiological state of hESCs in vitro. Therefore, hESCs with the PPP state are especially suitable for hPSC research.

In conclusion, we have redefined RSeT as FGF2-dependent and JAK- and TGFβ co-independent cellular models, with a unique pluripotent state between formative and primed pluripotency. RSeT cells are relatively insensitive to hypoxic tension. Primed hPSCs may have a restricted plasticity to be converted to the formative and naïve states under certain hPSC growth conditions (such as RSeT). It is imperative to have a thorough understanding and mastery of physiologically relevant growth conditions in order to maintain desired pluripotent states for disease modeling, drug discovery, and future clinical applications.

Materials and Methods

Cell lines

Three NIH registered human embryonic stem cell (hESC) lines, H1 (WA01, registration number 0043), H7 (WA07, registration number 0061), and H9 (WA09, registration number 0062), were used in this study. All three hESC lines were obtained from WiCell Research Institute (Madison, WI) and previously extensively characterized in our NIH Stem Cell Unit (Chen et al., 2014; Chen

et al., 2012; Mallon et al., 2013). More detailed information regarding these hESC lines is available (stemcells.nih.gov/research/).

MEF isolation and cell culture

MEFs were isolated from the mouse CF1 strain and grown on 0.1% gelatin-coated culture plates in DMEM/F12 medium supplemented with 10% fetal bovine serum (Gemini Bio-Products, West Sacramento, CA), 2 mM L-glutamine, and 0.1 mM non-essential amino acids.

hESC culture on MEF feeder

All hESC lines were initially cultured according to the supplier's protocols and adapted to a simple protocol as outlined below. The MULTIWELL™-6 WELL polystyrene plates were coated with 0.1% gelatin for 1 hour at 37°C. MEFs, at passage number 5 and 6 (designated as p5 and p6, respectively), were irradiated at a dose of 8,110 rads with an X-ray machine (Faxitron Bioptics, LLC., Tucson, AZ). The irradiated cells were plated at a density of 2 x 10^4 cells/cm^2 and incubated at 37°C for 24 hours. HESCs were plated on MEF feeder layer as small clumps (~ 50 to 100 µm in diameter) in the hESC medium containing 80% DMEM/F12 medium, 20% knock-out serum Replacement (KSR), 2 mM L-glutamine, and 0.1 mM non-essential amino acids, 0.1 mM β-mercaptoethanol, and 4 ng/mL of FGF2. Detailed descriptions about hESC passage, expansion, and cryopreservation are available in our previous reports (Chen et al., 2014; Chen et al., 2012; Mallon et al., 2013).

RSeT conversion, passages, and maintenance

RSeT medium has been advertised as a "defined" medium (StemCell Technologies Inc. Vancouver, catalog number 05970), without the inclusion of both FGF2 and TGFβ. RSeT converts and maintains primed hPSCs in a "naïve-like" state (www.stemcell.com). This medium was based on NHSM (Gafni et al., 2013), but made by a different formulation under a license from the Weizmann Institute of Science (Israel). The primed H1, H7, and H9 hESCs were converted to a new pluripotent state on MEFs in RSeT medium as recommended by the vendor's protocol as described below.

After 3-day induction or selection of primed hESCs, the cell colonies (in each well) were dissociated as single-cell suspension by TrypLE™ Express (Thermo Fisher Scientific, REF number 12604-013) and passaged on MEF feeder in 6-well plates in a 1:6 split ratio. The cells were grown in a humidified incubator with 5% CO_2, 3% O_2 (hypoxia), or 20% O_2 (normoxia). The ROCK inhibitor Y-27632 (10 µM) was added to enhance SCPE for the initial 3 passages. The RSeT cell lines were fully characterized at passage 10 in this study.

Primed and RSeT cell growth, proliferation, and pluripotency assays were examined in downstream experiments (or analyses) such as alkaline phosphatase, immunofluorescence assays, microarray, real-time PCR, and cell growth assays. RSeT cell lines were routinely analyzed with their primed counterparts. We refer to samples from different cell lines (e.g., H1, H7, and H9) with different passages (batches) as biological replicates and those triplicate or quadruplicate determinants in one independent experiment as technical replicates in this study.

Clonogenicity of primed and RSeT hESCs under the inhibition of JAK, TGFβ, and FGFR pathways

RSeT lines (at passage number 15) were dissociated in TrypLE solution, counted, and seeded at a density of 3 x 10^4 cells per well (in 12-well plate) on MEF feeder in 0.5 mL of RSeT complete medium. Drugs and growth factor-containing medium were added next day, incubated for 96 hours, and then cells in 4% formaldehyde. The cell colonies were stained with anti-NANOG antibody and 5 regions were randomly counted at in three wells (technical replicates) under an immunofluorescence microscope. Data are represented as the mean ± SD (standard deviation).

Alkaline phosphatase assay

Alkaline phosphatase staining (Alkaline Phosphatase Staining Kit II, catalog number. 00-0055, Stemgent, CA) was used to assess the number of pluripotent stem cells and their differentiated or partially differentiated counterparts. Cells were fixed in 4% formaldehyde for 20 minutes and permeabilized in 0.05% Tween-20 in D-PBS for 5 minutes at room temperature. Cells were then incubated with the AP Staining Solution A, B, and mix at room temperature for 20 minutes followed by rinsing cells twice in D-PBS. All samples were analyzed in the presence of D-PBS. Red, purple, and dark black colored colonies, which were NANOG and OCT4 positive, were counted as pluripotent colonies. Data are represented as the mean ± SD (standard deviation).

RNA preparations

RNA was extracted by using the basic Trizol protocol (Invitrogen Inc.) with slight modifications. Briefly, hESCs per sample were dissolved in 1 mL of the Trizol solution and stored in -80°C freezer prior to use. HESC lysates were further subjected to chloroform extraction and isopropanol precipitations. A DNA-Free kit (Applied Biosystems/Ambion, Austin, TX) was used to remove genomic DNA contamination. The quality of RNAs was agarose gel-verified and the concentration determined by using the ND-1000 UV-Vis Spectrophotometer (NanoDrop Technologies Wilmington, DE).

Complementary DNA (cDNA) microarray

Genome-wide gene expression analysis in primed and RSeT hESCs (at passage 10) was performed and analyzed with the Agilent chip (Agilent Technologies Inc., Santa Clara, CA). Global gene expression microarray was performed and analyzed using Agilent software, reagents according to the manufacturer's instructions (Mallon et al., 2013). Sample amplification and labeling were carried out using the Low Input Quick Amp labeling Kit (Catalog number 5190-2305, Agilent Technologies, Santa Clara, CA). Briefly, total RNAs (200 ng) were reverse transcribed into cDNAs using T7 primer mix in 5X First Strand Buffer, 0.1 M DTT, 10 mM dNTP Mix, and Affinity Script RNase Block Mix. The cDNAs were used to synthesize Cy3-labeled cRNAs in a reaction containing nuclease-free water, 5X Transcription Buffer, 0.1 M DTT, NTP Mix, T7 RNA polymerase blend, and Cyanine 3-CTP. The RNeasy Mini Kit (QIAGEN) was used to purify the linearly amplified cRNA samples. Cy3-labeled cRNAs were measured using the NanoDrop Spectrophotometer.

Cy3-labeled cRNAs were further hybridized onto Agilent Whole Human Genome kit 4 x 44K slides (Catalog number G4112F, Agilent Technologies) containing 44,397 oligonucleotide probes in 2X Hi-RPM Hybridization buffer (Catalog number 5188-5242, Agilent Technologies) at 65°C for 17 hours at 10 r.p.m. as recommended. Slides were scanned using an Agilent DNA microarray compatible scanner, with a one-color scan setting for 4 x 44k array slides. Quantification files were processed using Agilent Feature Extraction Software. We also used quantile normalization and background correction for further data analysis.

Quantitative real-time PCR (qPCR)

For qRT-PCR, genomic DNA-free RNAs (2 µg) were reversely transcribed (in a 20-µl volume) into cDNAs using the SuperScript™ VILO™ cDNA Synthesis Kit (Thermo Fisher catalog number 11754) and diluted in 380 µl nuclease-free H2O. All TaqMan™ MGB probes were purchased from Thermo Fisher Scientific Inc. (see Key Resource Table). Both β-actin (*ATCB*) and glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) assays were used as internal controls for normalization. Approximately, 100 ng of cDNAs were amplified in 1X TaqMan™ Fast Advanced Master Mix and 1X TaqMan™ Assay in a 20-µL reaction using the QuantStudio 6 Flex real-time PCR Platform (Thermo Fisher Scientific Inc.) according to the manufacturer's instructions. Thermal cycling was done with 2 minutes at 50°C, 20 seconds at 95°C followed by 40 cycles of PCR amplification (i.e., 1 second denature at 95°C and 20 seconds at 62°C). Equal amplification efficiencies in TaqMan assays were confirmed by serial cDNA dilutions. Data were analyzed by a three-step comparative cycle threshold (Ct) method, which includes (1)

normalization to endogenous control ($Ct_{Target\ gene} - Ct_{Endogenous\ control} = \Delta Ct$), (2) normalization to calibrator sample ($Ct_{Sample} - Ct_{calibrator} = \Delta\Delta Ct$), and (3) the final use of formula ($2^{-\Delta\Delta Ct}$) to calculate the relative fold change of gene expression.

Flow cytometry

Flow cytometric analysis of conventional pluripotent stem cell surface markers for primed and RSeT hESCs was initially performed as described previously (Mallon et al., 2013). For primed- or naïve-specific cell surface marker expression, we modified the above method as follows. Briefly, hESC colonies were grown on MEF feeder in hESC and RSeT media in a 6-well plate. For hESCs on feeders, the cells were rinsed once in 1X D-PBS and dissociated by Collagenase type IV for 40 min. Only floating colonies were collected, sedimented once in D-PBS, further digested in 1 mL of TrypLE for 10 min into single-cell suspension followed by single cell confirmation under the microscope. The enzymatic activity was inactivated by adding FACS buffer (DMEM/F12 medium containing 10% FBS). Then, 2 mL of cells (~1 x 10^6 cells) were transferred into a Falcon 5 mL-polypropylene round-bottom tube (REF 352063, Corning) and centrifuged in a swing bucket at 300 *g* for 5 minutes. The above wash step with FACS buffer was repeated once. The cell pellets were resuspended in 100 µL of FACS buffer and incubated with Alexa Fluor® 647 (AF647)- or FITC-conjugated primary antibodies or isotype controls (4 µg/mL, Key Resource Table) on ice for 1 hour with gentle flickering the tubes every 20 minutes. The incubation was terminated by adding 4 mL of FACS buffer followed by centrifugation at 300 *g*. The cell pellets were then resuspended in 0.5 mL of D-PBS and stained with 2.5 µL of bis benzamide (Hoechst 33258) solution (0.2 mg/mL, Sigma) at room temperature for 20 minutes. The cells were further fixed (in 4% of formaldehyde solution) for 20 minutes at room temperature. The fixative was removed by rinsing the cells twice in D-PBS followed by filtering the cell solution into a Falcon 5-mL polypropylene round-bottom tube with a cell-strainer cap (REF 352253, Corning). Finally, the cells were analyzed by a Beckman Coulter machine at the NINDS Core Facility (National Institutes of Health, Bethesda, MD) or by a Guava easyCyte™ device (Luminex Corporation, Austin TX). We used DAPI 405-448/59-area (at X-axis) versus Y-axis DAPI-405-448/59-width (at Y-axis) to gate nucleated single cells for data analysis by FlowJo v10.9.0 (www.flowjo.com). All positive staining cells are related to their isotype (negative) controls, which are included in every independent experiment. These controls are partially shown in Figure S4.

Immunofluorescence

Primed and RSeT hESCs were grown either in 6- or 24-well plates and fixed in 4% formaldehyde at room temperature for 20 minutes. The cells were also grown on ibidi chamber slides (ibidi USA, Inc., Fitchburg, WI) for confocal microscopic analysis. Cell samples were blocked with 10% normal goat serum and 0.1% bovine serum albumin (BSA) in the presence or absence of 0.1% Triton X-100 (diluted in D-PBS) at room temperature for 1 hour, then reacted with primary antibodies at 4°C overnight or at room temperature for 2 hours, and subsequently incubated with Alexa Fluro® conjugated secondary antibodies (Thermo Fisher Scientific Inc., Waltham, MA) in 5% blocking solutions for 1 hour at room temperature. Finally, the cells were stained with Hoechst 33258 solution or DAPI. All images used for quantitative analysis were obtained under unsaturated exposure conditions. The detailed procedures are also available in our previous reports (Chen et al., 2014; Chen et al., 2012; Mallon et al., 2013).

Microscopy

Cell samples were examined under a series of Zeiss microscopes, including Axiovert 25 light microscope, Axiovert 200 fluorescence, Axio Observer, and LSM 810 confocal microscopes (Zeiss, Jena, Germany). These microscopes were equipped with corresponding software such as Adobe Photoshop® (Adobe Inc., Mountain View, CA), AxioVision Rel 4.6 (Zeiss), and ZenBlue (Zeiss) for image acquisitions and analysis.

Microarray data analysis

The microarray data analysis was performed essentially as described previously (Chen et al., 2012; Mallon et al., 2013). ANOVA (analysis of variance) was used to assess the number of genes (probes) (n = 43,376) enumerated in the datasets, in which 32,563 probes had the expression values greater than background, 26,317 probes had the values greater than the noise cut-off value 6.5 in at least one sample, and 1,958 probes showed type III ANOVA corrected *P*-values less than 0.05. Gene probes selected for use represent those without noise-, processing-, and gender-biased ones. These probes passed multiple tests, including ANOVA under multiple comparison correction condition ($P < 0.05$) using class as the factor, *post-hoc* tests for at least one pair-wise class comparison (Tukey HSD $P < 0.05$), and a mean/difference criterion for the same pair-wise class comparison having a Tukey HSD $P < 0.05$. Pearson correlation heatmap depicts 12 cell lines using log (base = 2) transformed and quantile normalized mRNA expression. PCA was performed as described previously (Mallon et al., 2013).

Meta-analysis of human pluripotent stem cell (hPSC) and primate early embryo datasets

Comparative meta-analysis of hPSC datasets generated from various protocols was performed by integrating three interrelated analytic tools including PCA, *t*-SNE, and SC3 consensus clustering methods, which were described previously (Johnson et al., 2021). This multivariate meta-analysis platform consists of a series of independent modules that constitute a gene expression matrix, which comprises 16 datasets, including the microarray dataset D21 in this study.

All naïve protocols were based on the modifications of the 2iL protocol [i.e., MEK and GSK3 inhibitors in the presence of leukemia inhibitory factor (LIF)]. The related datasets (with GSE or EMBL-EBI accession numbers in parentheses) include: D3 (E-MTAB-2031) from 3iL: MEK, GSK3, and BMP pathway inhibitors in the presence of LIF in TeSR1 medium (Chan et al., 2013); D5 (GSE46872) from NHSM (Gafni et al., 2013); D6 (GSE50868) fate maintenance medium (FMM) (Valamehr et al., 2014); D7 (GSE59435) from MEK, GSK3, ROCK, BRAF, and SRC pathway inhibitors in the presence of LIF and Activin (5iLA) (Theunissen et al., 2016; Theunissen et al., 2014); D22B (E-MTAB-2857) from titrated MEK and GSK3 inhibitors with LIF plus the PKC inhibitor Go6983 (t2iLGo) (Takashima et al., 2014); D23 (E-MTAB-2856) from t2iLGo (Takashima et al., 2014); D24 (E-MTAB-4461) also from t2iLGo (Guo et al., 2016); D25 (GSE36552) containing human embryo datasets (Yan et al., 2013); D26 (GSE29397) representing the second human embryo dataset (Vassena et al., 2011); D27 (SRP115256) encompassing a panel of datasets from multiple naïve protocols at the same laboratory (Liu et al., 2017); D28 (GSE44430) from 2iL in the presence of the tankyrase inhibitor XAV939 (XAV939/2iL) (Zimmerlin et al., 2016); D29 (GSE141639) from XAV939/2iL (Park et al., 2020); D30 (GSE66507) containing human blastocysts at embryonic day 6 to 7 (E6-E7) (Blakeley et al., 2015); D31 (E-MTAB-3929) having human embryos at E3-E7 (Petropoulos et al., 2016); and D32 (GSE74767) encompassing Cynomolgus monkey embryos at E7-E17 (Nakamura et al., 2016).

To integrate the *Cynomolgus* monkey dataset D32, Fastq files for 34 *Macaca fascicularis* samples were downloaded from NCBI SRA (PRJNA301445) and used as input into the nf-core RNA-Seq pipeline (https://nf-co.re/rnaseq/3.10.1, --clip_r1 5 --clip_r2 5) to generate salmon-enumerated gene expression in transcript per million (tpm). The reference genome and annotations used were those for Macaca_fascicularis_6.0 (GCA_011100615.1). Expression was then imported into R (https://cran.r-project.org/), pedestalled by 2, and log2 transformed. To map homologous gene symbols from *Macaca fascicularis* to *Homo sapiens*, biomaRt and SignacX functions were used as described in the "Crabeating_vignette" (https://cran.r-project.org/web/packages/SignacX/vignettes/). All datasets were quantile-normalized, limited to unique genes, percentile-scaled, and batch-corrected (Johnson et al., 2021).

For PCA in meta-analysis, normalized datasets were used to construct a Pearson correlation-based matrix and employed to calculate the principal components using the R programming software. For visualization, PC values were used to map the data points in PCA scatter plots, in which each data point contains the genome-wide gene expression profile of one sample (or cell line). The *t*-SNE analysis was also implemented by the R program using the "Rtsne" function under default parameters (dims = 2, perplexity = 30, theta = 0.5, pca = T, momentum = 0.5, final momentum = 0.8, eta = 200) (Johnson et al., 2021).

Data and code availability

Data reported in this study will be available in supplemental figures and tables. The microarray data were deposited in the Gene Expression Omnibus (GEO) database at the National Center for Biotechnology Information (NCBI accession: GSE217275). We provide the link below (with a security token: klclsoqmdtuhxuf), www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE217275, to facilitate the access to microarray data. Any additional information required to reanalyze the data reported in this study would be available upon request.

ACKNOWLEDGMENTS

This work was supported by the Intramural Research Program of the National Institutes of Health (NIH) at the National Institute of Neurological Disorders and Stroke (K.G.C., K.R.J., K.P., D.M., Y.C.F., B.S.M.) and the National Institute of Dental and Craniofacial Research (P.G.R).

AUTHOR CONTRIBUTIONS

K.G.C., B.S.M., and P.G.R.: Conceptualization; K.G.C. and B.S.M.: methodology; K.G.C., B.S.M. P.G.R.: validation and formal analysis; K.G.C., K.R.J., K.P., D.M., F.Y., W.F.L., Y.C.F., B.S.M.: experimentation and investigation; K.G.C.: writing – original draft; K.G.C., B.S.M., P.G.R.; writing – review and editing; K.G.C., K.R.J.: visualization. All authors approved the final manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

FIGURE LEGENDS

Figure 1. **RSeT conversion of human embryonic stem cells (hESCs) under normoxia and hypoxia.** (**A**) Linear presentation of diverse hPSC pluripotent states determined by meta-

analysis (Johnson et al., 2021), with current naïve conversion or reprogramming protocols (labeled in italicized characters with first authors' names), t2iLGo, t2iLGoY, 5iLA, 5iLAF, XAV939/2iL, FMM, TeSR1/3iL, NHSM, and RSeT, from the laboratories of AS (Austin Smith), EZ (Elias Zambidis), HN (Huck-Hui Ng), JH (Jacob Hanna), JP (Jose Polo), PF (Peter Flynn), and RJ (Rudolf Jaenisch). The hPSCs (in red-font characters) are related to RSeT protocols (SCT RSeT and NHSM). (**B**) Schema of hESC lines (H1, H7, and H9) treated under different primed and RSeT conditions as indicated. (**C**) Phase images of primed and RSeT hESCs grown on MEFs: PN, primed growth under normoxia; PH, primed growth under hypoxia; RN, RSeT growth under normoxia; and RH, RSeT growth under hypoxia. Additional abbreviations: 2c, 4c, and 8c: 2-cell, 4-cell, and 8 cell embryos, respectively; B, blastocyst; I, intermediate transitions between the naive and primed states; M, morulae; N, cells with the naïve pluripotent state; P, primed pluripotent states; and Zy, zygote.

Figure 2. **Genome-wide transcriptomic analysis of RSeT hESCs**. **(A)** cDNA microarray heatmap of mRNA expression in RSeT hESCs. (**B**) Covariance-based principal component analysis (PCA) of the variability of gene expression as shown in Figure 2A. (**C**) Representative naïve, formative, primed, and imprinted gene markers from microarray analysis. (**D**) Quantitative real-time PCR (qPCR) confirmation of gene expression patterns based on microarray.

Figure 3. **Genome-wide transcriptomic analysis of RSeT hESCs and meta-analysis of human naïve pluripotency**. (**A**) Comparative meta-analysis of hPSC datasets generated from various protocols by Pearson correlation-based PCA and *t*-distributed stochastic neighbor embedding (*t*-SNE), which depict cellular and pluripotent states (color-coded dots) with dataset identification numbers as labeled (D3 to D29). The datasets used for comparison in this study were designated based on the name of first authors of published reports as detailed previously (Johnson et al., 2021). (**B**) PCA and *t*-SNE of detailed pluripotent states and developmental stages with the addition of multiple early primate embryo datasets (D30 to D32). (**C**) Mapping cellular and pluripotent states based on the Euclidean distance of gene-wide transcriptomic datasets. Abbreviations for cellular states and human naïve protocols (including the expression of transgenes) are indicated, which are also detailed in Materials and Methods and in the legend to Figure 1A. Additional abbreviations: Dif, lineage differentiation; I, intermediate transitions between the naïve and primed states; N, cells with naive pluripotent states; OKSM, transgene expression of the Yamanaka factors OCT4, KLF4, SOX2, and c-MYC; P, primed pluripotent states; R, hPSCs with a distinct pluripotent state in RSeT medium.

Figure 4. Flow cytometric analysis of human pluripotent stem cell surface marker expression. (A) Representative analysis of the expression of the human pluripotent stem cell surface markers SSEA-4, Tra-1-60, and Tra-1-81 in RSeT H1 and H9 cells. (**B to D**) Expression of the primed-specific markers (CD57, CD24, and CD90) as well as the naïve-specific surface markers (CD75, SUSD2, and CD130) in Hoechst 33258-gated H1 and H7 single cells. PH, primed hESCs grown under hypoxia; RH, RSeT-converted hESC grown under hypoxia.

Figure 5. Flow cytometric analysis of primed- and naïve-specific surface expression. (A) Expression of the primed-specific markers (CD57, CD24, and CD90) as well as the naïve-specific surface markers (CD75, SUSD2, and CD130) in Hoechst 33258-gated primed and RSeT H9 cells under hypoxia (H9PH and H9RH respectively), H9RH cells grown under normoxia for 6 days (i.e., H9RN, p0, d6) and 1 passage (H9RN, p1). (**B to D**) Statistical analysis of the primed-specific markers (CD57, CD24, and CD90), naïve-specific surface markers (CD75, SUSD2, and CD130), double positive and double negative subsets in Hoechst 33258-gated H1, H7, and H9 single cells. Data are represented as the mean ± SD (standard deviation) of fluorescence intensity in H1, H7, and H9 lines (n = 3). PH, primed hESCs grown under hypoxia; RH, RSeT-converted hESC grown under hypoxia.

Figure 6. **Single-cell colony growth assays of FGF2- TGFβ-, and JAK-dependent hESC colonies**. (**A**-**C**) Primed H1, H7, and H9 cells, grown on MEFs under normoxic and hypoxic conditions, treated with the FGF2 inhibitor PD173074, TGFβ1 inhibitor SB431542, and JAK inhibitor I (JAKi) as indicated. The hESC colonies were fixed at day 3 in 4% formaldehyde, stained for alkaline phosphatase positive (undifferentiated) colonies, and counted. The asterisk signs indicate the major changes observed between primed and/or RSeT cell lines. (**D**) JAK, TGFβ, and FGFR pathway inhibition assays in RSeT hESC lines, with JAK inhibitor I (0.5 μM), the FGF2 inhibitor PD173074 (0.1 μM), and the TGFβ inhibitor SB431542 (0.5 μM). The cell colonies were stained with the anti-NANOG antibody. Data are represented as the mean ± SD (standard deviation) of NANOG positive colonies counted randomly at 5 regions (n = 5) under an immunofluorescence microscope. inh, inhibitor or inhibition; NA, cells not available at the time of the assay due to the significantly slow growth of cells.

Figure 7. Regulation of pluripotent states by integrated signal inputs from growth media. (**A**) Reprogramming the primed pluripotent stem cells to naïve, formative, and RSeT-based states. Differential growth signaling molecules in media may limit the reprogramming efficiency and trajectories observed in different naïve conversion protocols such as t2iLGo plus

reprogramming transgenes (t2iLGo + Rp), 5iLA, t2iLGo, 3iL, NHSM, and RSeT, resulting in diverse pluripotent states upstream or downstream of the formative state. (**B**) Color key, symbols, and abbreviations for cell types, cellular or pluripotent states, and integrated reprogramming trajectories of naïve (ITN) or RSeT (ITR) states.

REFERENCES

Blakeley P, Fogarty NM, Del Valle I, Wamaitha SE, Hu TX, Elder K, Snell P, Christie L, Robson P and Niakan KK (2015) Defining the three cell lineages of the human blastocyst by single-cell RNA-seq. *Development* **142**:3613.

Bredenkamp N, Stirparo GG, Nichols J, Smith A and Guo G (2019) The Cell-Surface Marker Sushi Containing Domain 2 Facilitates Establishment of Human Naïve Pluripotent Stem Cells. *Stem Cell Reports* **12**:1212-1222.

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 and Vallier L (2007) Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* **448**:191-195.

Chan YS, Goke J, Ng JH, Lu X, Gonzales KA, Tan CP, Tng WQ, Hong ZZ, Lim YS and Ng HH (2013) Induction of a human pluripotent state with distinct regulatory circuitry that resembles preimplantation epiblast. *Cell Stem Cell* **13**:663-675.

Chen KG, Hamilton RS, Robey PG and Mallon BS (2014) Alternative cultures for human pluripotent stem cell production, maintenance, and genetic analysis. *J Vis Exp*. 2014 Jul 24;(89):51519. doi: 10.3791/51519.

Chen KG, Mallon BS, Hamilton RS, Kozhich OA, Park K, Hoeppner DJ, Robey PG and McKay RD (2012) Non-colony type monolayer culture of human embryonic stem cells. *Stem Cell Res* **9**:237-248.

Collier AJ, Panula SP, Schell JP, Chovanec P, Plaza Reyes A, Petropoulos S, Corcoran AE, Walker R, Douagi I, Lanner F and Rugg-Gunn PJ (2017) Comprehensive Cell Surface Protein Profiling Identifies Specific Markers of Human Naïve and Primed Pluripotent States. *Cell Stem Cell* **20**:874-890 e877.

Cornacchia D, Zhang C, Zimmer B, Chung SY, Fan Y, Soliman MA, Tchieu J, Chambers SM, Shah H, Paull D, Konrad C, Vincendeau M, Noggle SA, Manfredi G, Finley LWS, Cross JR, Betel D and Studer L (2019) Lipid Deprivation Induces a Stable, Naïve-to-Primed Intermediate State of Pluripotency in Human PSCs. *Cell Stem Cell* **25**:120-136 e110.

Dodsworth BT, Flynn R and Cowley SA (2015) The Current State of Naïve Human Pluripotency. *Stem Cells* **33**:3181-3186.

Dong C, Fischer LA and Theunissen TW (2019) Recent insights into the naïve state of human pluripotency and its applications. *Exp Cell Res* **385**:111645.

Gafni O, Weinberger L, Mansour AA, Manor YS, Chomsky E, Ben-Yosef D, Kalma Y, Viukov S, Maza I, Zviran A, Rais Y, Shipony Z, Mukamel Z, Krupalnik V, Zerbib M, Geula S, Caspi I, Schneir D, Shwartz T, Gilad S, Amann-Zalcenstein D, Benjamin S, Amit I, Tanay A, Massarwa

R, Novershtern N and Hanna JH (2013) Derivation of novel human ground state naïve pluripotent stem cells. *Nature* **504**:282-286.

Guo G, von Meyenn F, Rostovskaya M, Clarke J, Dietmann S, Baker D, Sahakyan A, Myers S, Bertone P, Reik W, Plath K and Smith A (2017) Epigenetic resetting of human pluripotency. *Development* **144**:2748-2763.

Guo G, von Meyenn F, Santos F, Chen Y, Reik W, Bertone P, Smith A and Nichols J (2016) Naïve Pluripotent Stem Cells Derived Directly from Isolated Cells of the Human Inner Cell Mass. *Stem Cell Reports* **6**:437-446.

Hanna JH, Saha K and Jaenisch R (2010) Pluripotency and cellular reprogramming: facts, hypotheses, unresolved issues. *Cell* **143**:508-525.

James D, Levine AJ, Besser D and Hemmati-Brivanlou A (2005) TGFbeta/activin/nodal signaling is necessary for the maintenance of pluripotency in human embryonic stem cells. *Development* **132**:1273-1282.

Johnson KR, Mallon BS, Fann YC and Chen KG (2021) Multivariate meta-analysis reveals global transcriptomic signatures underlying distinct human naïve-like pluripotent states. *PLoS One* **16**:e0251461.

Johnson KR, Mallon, B.S., Fann, Y.C., Chen, K.G. (2019) Interlaboratory Data Variability Contributes to the Differential Principal Components of Human Primed and Naïve-like Pluripotent States in Multivariate Meta-Analysis, bioRxiv. doi: https://doi.org/10.1101/822163

Kagawa H, Javali A, Khoei HH, Sommer TM, Sestini G, Novatchkova M, Scholte Op Reimer Y, Castel G, Bruneau A, Maenhoudt N, Lammers J, Loubersac S, Freour T, Vankelecom H, David L and Rivron N (2022) Human blastoids model blastocyst development and implantation. *Nature* **601**:600-605.

Khan SA, Park KM, Fischer LA, Dong C, Lungjangwa T, Jimenez M, Casalena D, Chew B, Dietmann S, Auld DS, Jaenisch R and Theunissen TW (2021) Probing the signaling requirements for naïve human pluripotency by high-throughput chemical screening. *Cell Rep* **35**:109233.

Kilens S, Meistermann D, Moreno D, Chariau C, Gaignerie A, Reignier A, Lelievre Y, Casanova M, Vallot C, Nedellec S, Flippe L, Firmin J, Song J, Charpentier E, Lammers J, Donnart A, Marec N, Deb W, Bihouee A, Le Caignec C, Pecqueur C, Redon R, Barriere P, Bourdon J, Pasque V, Soumillon M, Mikkelsen TS, Rougeulle C, Freour T, David L and Milieu Interieur C (2018) Parallel derivation of isogenic human primed and naïve induced pluripotent stem cells. *Nat Commun* **9**:360.

Liao J, Karnik R, Gu H, Ziller MJ, Clement K, Tsankov AM, Akopian V, Gifford CA, Donaghey J, Galonska C, Pop R, Reyon D, Tsai SQ, Mallard W, Joung JK, Rinn JL, Gnirke A and Meissner A (2015) Targeted disruption of DNMT1, DNMT3A and DNMT3B in human embryonic stem cells. *Nat Genet* **47**:469-478.

Liu X, Nefzger CM, Rossello FJ, Chen J, Knaupp AS, Firas J, Ford E, Pflueger J, Paynter JM, Chy HS, O'Brien CM, Huang C, Mishra K, Hodgson-Garms M, Jansz N, Williams SM, Blewitt ME, Nilsson SK, Schittenhelm RB, Laslett AL, Lister R and Polo JM (2017) Comprehensive characterization of distinct states of human naïve pluripotency generated by reprogramming. *Nat Methods* **14**:1055-1062.

Mallon BS, Chenoweth JG, Johnson KR, Hamilton RS, Tesar PJ, Yavatkar AS, Tyson LJ, Park K, Chen KG, Fann YC and McKay RD (2013) StemCellDB: the human pluripotent stem cell database at the National Institutes of Health. *Stem Cell Res* **10**:57-66.

Nakamura T, Okamoto I, Sasaki K, Yabuta Y, Iwatani C, Tsuchiya H, Seita Y, Nakamura S, Yamamoto T and Saitou M (2016) A developmental coordinate of pluripotency among mice, monkeys and humans. *Nature* **537**:57-62.

Nichols J and Smith A (2009) Naïve and primed pluripotent states. *Cell Stem Cell* **4**:487-492.

Park TS, Zimmerlin L, Evans-Moses R, Thomas J, Huo JS, Kanherkar R, He A, Ruzgar N, Grebe R, Bhutto I, Barbato M, Koldobskiy MA, Lutty G and Zambidis ET (2020) Vascular progenitors generated from tankyrase inhibitor-regulated naïve diabetic human iPSC potentiate efficient revascularization of ischemic retina. *Nat Commun* **11**:1195.

Pera MF and Rossant J (2021) The exploration of pluripotency space: Charting cell state transitions in peri-implantation development. *Cell Stem Cell* **28**:1896-1906.

Petropoulos S, Edsgard D, Reinius B, Deng Q, Panula SP, Codeluppi S, Plaza Reyes A, Linnarsson S, Sandberg R and Lanner F (2016) Single-Cell RNA-Seq Reveals Lineage and X Chromosome Dynamics in Human Preimplantation Embryos. *Cell* **165**:1012-1026.

Rossant J and Tam PPL (2017) New Insights into Early Human Development: Lessons for Stem Cell Derivation and Differentiation. *Cell Stem Cell* **20**:18-28.

Smith A (2017) Formative pluripotency: the executive phase in a developmental continuum. *Development* **144**:365-373.

Stirparo GG, Boroviak T, Guo G, Nichols J, Smith A and Bertone P (2018) Integrated analysis of single-cell embryo data yields a unified transcriptome signature for the human pre-implantation epiblast. *Development* **145**.

Szczerbinska I, Gonzales KAU, Cukuroglu E, Ramli MNB, Lee BPG, Tan CP, Wong CK, Rancati GI, Liang H, Goke J, Ng HH and Chan YS (2019) A Chemically Defined Feeder-free System for the Establishment and Maintenance of the Human Naïve Pluripotent State. *Stem Cell Reports* **13**:612-626.

Takashima Y, Guo G, Loos R, Nichols J, Ficz G, Krueger F, Oxley D, Santos F, Clarke J, Mansfield W, Reik W, Bertone P and Smith A (2014) Resetting transcription factor control circuitry toward ground-state pluripotency in human. *Cell* **158**:1254-1269.

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.

Theunissen TW, Friedli M, He Y, Planet E, O'Neil RC, Markoulaki S, Pontis J, Wang H, Iouranova A, Imbeault M, Duc J, Cohen MA, Wert KJ, Castanon R, Zhang Z, Huang Y, Nery JR, Drotar J, Lungjangwa T, Trono D, Ecker JR and Jaenisch R (2016) Molecular Criteria for Defining the Naïve Human Pluripotent State. *Cell Stem Cell* **19**:502-515.

Theunissen TW, Powell BE, Wang H, Mitalipova M, Faddah DA, Reddy J, Fan ZP, Maetzel D, Ganz K, Shi L, Lungjangwa T, Imsoonthornruksa S, Stelzer Y, Rangarajan S, D'Alessio A, Zhang J, Gao Q, Dawlaty MM, Young RA, Gray NS and Jaenisch R (2014) Systematic

identification of culture conditions for induction and maintenance of naïve human pluripotency. *Cell Stem Cell* **15**:471-487.

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* **282**:1145-1147.

Valamehr B, Robinson M, Abujarour R, Rezner B, Vranceanu F, Le T, Medcalf A, Lee TT, Fitch M, Robbins D and Flynn P (2014) Platform for induction and maintenance of transgene-free hiPSCs resembling ground state pluripotent stem cells. *Stem Cell Reports* **2**:366-381.

Vallier L, Alexander M and Pedersen RA (2005) Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. *J Cell Sci* **118**:4495-4509.

Vallot C, Patrat C, Collier AJ, Huret C, Casanova M, Liyakat Ali TM, Tosolini M, Frydman N, Heard E, Rugg-Gunn PJ and Rougeulle C (2017) XACT Noncoding RNA Competes with XIST in the Control of X Chromosome Activity during Human Early Development. *Cell Stem Cell* **20**:102-111.

Vassena R, Boue S, Gonzalez-Roca E, Aran B, Auer H, Veiga A and Izpisua Belmonte JC (2011) Waves of early transcriptional activation and pluripotency program initiation during human preimplantation development. *Development* **138**:3699-3709.

Wang P, Rodriguez RT, Wang J, Ghodasara A and Kim SK (2011) Targeting SOX17 in human embryonic stem cells creates unique strategies for isolating and analyzing developing endoderm. *Cell Stem Cell* **8**:335-346.

Ware CB, Nelson AM, Mecham B, Hesson J, Zhou W, Jonlin EC, Jimenez-Caliani AJ, Deng X, Cavanaugh C, Cook S, Tesar PJ, Okada J, Margaretha L, Sperber H, Choi M, Blau CA, Treuting PM, Hawkins RD, Cirulli V and Ruohola-Baker H (2014) Derivation of naive human embryonic stem cells. *Proc Natl Acad Sci U S A* **111**:4484-4489.

Weinberger L, Ayyash M, Novershtern N and Hanna JH (2016) Dynamic stem cell states: naïve to primed pluripotency in rodents and humans. *Nat Rev Mol Cell Biol* **17**:155-169.

Wray J, Kalkan T and Smith AG (2010) The ground state of pluripotency. *Biochem Soc Trans* **38**:1027-1032.

Wu J, Platero-Luengo A, Sakurai M, Sugawara A, Gil MA, Yamauchi T, Suzuki K, Bogliotti YS, Cuello C, Morales Valencia M, Okumura D, Luo J, Vilarino M, Parrilla I, Soto DA, Martinez CA, Hishida T, Sanchez-Bautista S, Martinez-Martinez ML, Wang H, Nohalez A, Aizawa E, Martinez-Redondo P, Ocampo A, Reddy P, Roca J, Maga EA, Esteban CR, Berggren WT, Nunez Delicado E, Lajara J, Guillen I, Guillen P, Campistol JM, Martinez EA, Ross PJ and Izpisua Belmonte JC (2017) Interspecies Chimerism with Mammalian Pluripotent Stem Cells. *Cell* **168**:473-486 e415.

Xu RH, Peck RM, Li DS, Feng X, Ludwig T and Thomson JA (2005) Basic FGF and suppression of BMP signaling sustain undifferentiated proliferation of human ES cells. *Nat Methods* **2**:185-190.

Yan L, Yang M, Guo H, Yang L, Wu J, Li R, Liu P, Lian Y, Zheng X, Yan J, Huang J, Li M, Wu X, Wen L, Lao K, Li R, Qiao J and Tang F (2013) Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nat Struct Mol Biol* **20**:1131-1139.

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.

Zhou J, Hu J, Wang Y and Gao S (2023) Induction and application of human naïve pluripotency. *Cell Rep* **42**:112379.

Zimmerlin L, Park TS, Huo JS, Verma K, Pather SR, Talbot CC, Jr., Agarwal J, Steppan D, Zhang YW, Considine M, Guo H, Zhong X, Gutierrez C, Cope L, Canto-Soler MV, Friedman AD, Baylin SB and Zambidis ET (2016) Tankyrase inhibition promotes a stable human naïve pluripotent state with improved functionality. *Development* **143**:4368-4380.

Zimmerlin L, Park TS and Zambidis ET (2017) Capturing Human Naïve Pluripotency in the Embryo and in the Dish. *Stem Cells Dev* **26**:1141-1161.

Figure 1

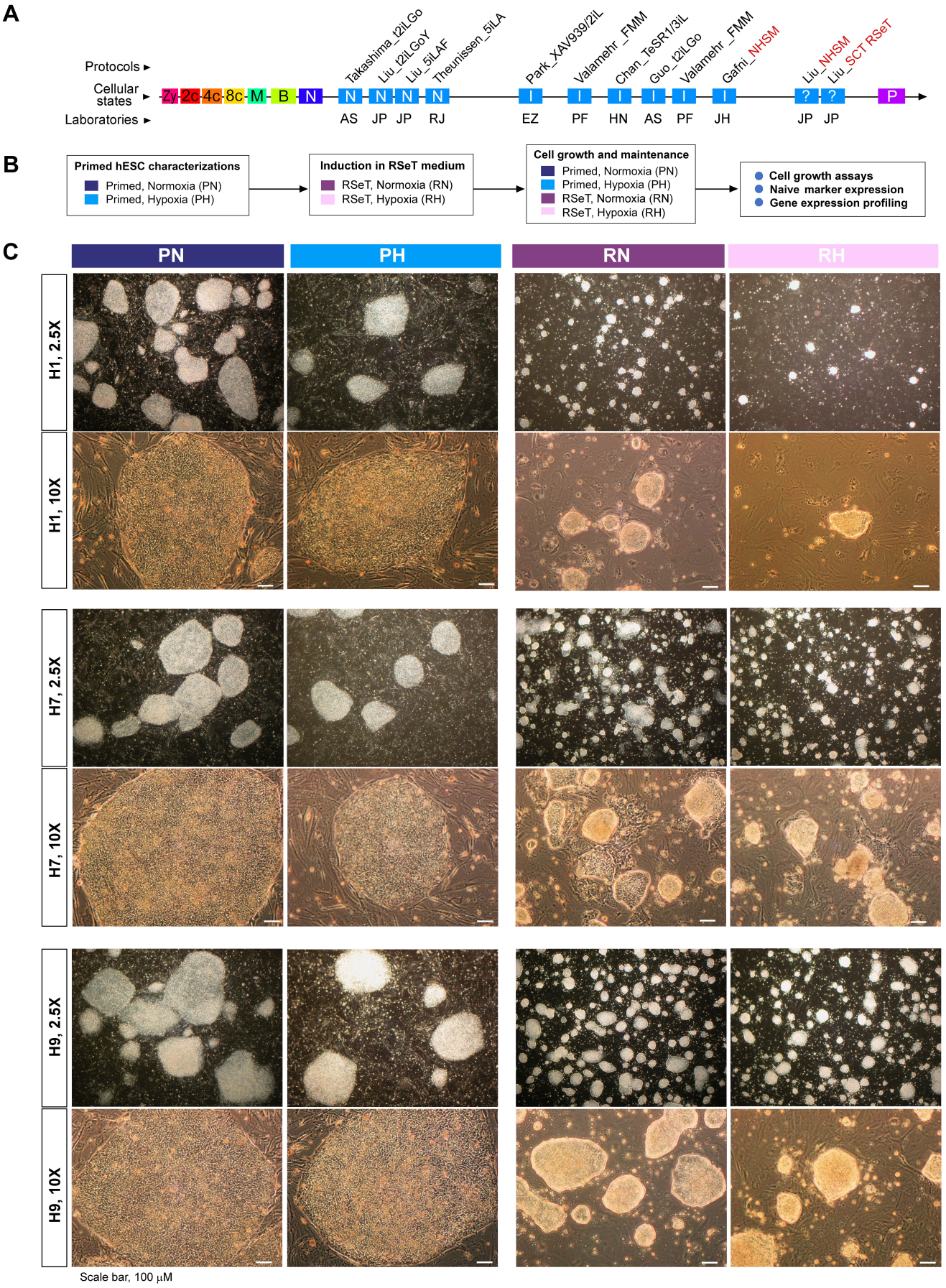

A
Protocols
Cellular states
Laboratories
Zy 2c 4c 8c M B N
Takashima_t2iLGo
Liu_t2iLGoY
Liu_5iLAF
Theunissen_5iLA
Park_XAV939/2iL
Valamehr _FMM
Chan_TeSR1/3iL
Guo_t2iLGo
Valamehr _FMM
Gafni_NHSM
Liu_NHSM
Liu_SCT RSeT
N N N N
I I I I I I
? ?
P
AS JP JP RJ
EZ PF HN AS PF JH
JP JP
B
Primed hESC characterizations
Primed, Normoxia (PN)
Primed, Hypoxia (PH)
Induction in RSeT medium
RSeT, Normoxia (RN)
RSeT, Hypoxia (RH)
Cell growth and maintenance
Primed, Normoxia (PN)
Primed, Hypoxia (PH)
RSeT, Normoxia (RN)
RSeT, Hypoxia (RH)
Cell growth assays
Naive marker expression
Gene expression profiling
C
PN
PH
RN
RH
H1, 2.5X
H1, 10X
H7, 2.5X
H7, 10X
H9, 2.5X
H9, 10X
Scale bar, 100 μM

Figure 2

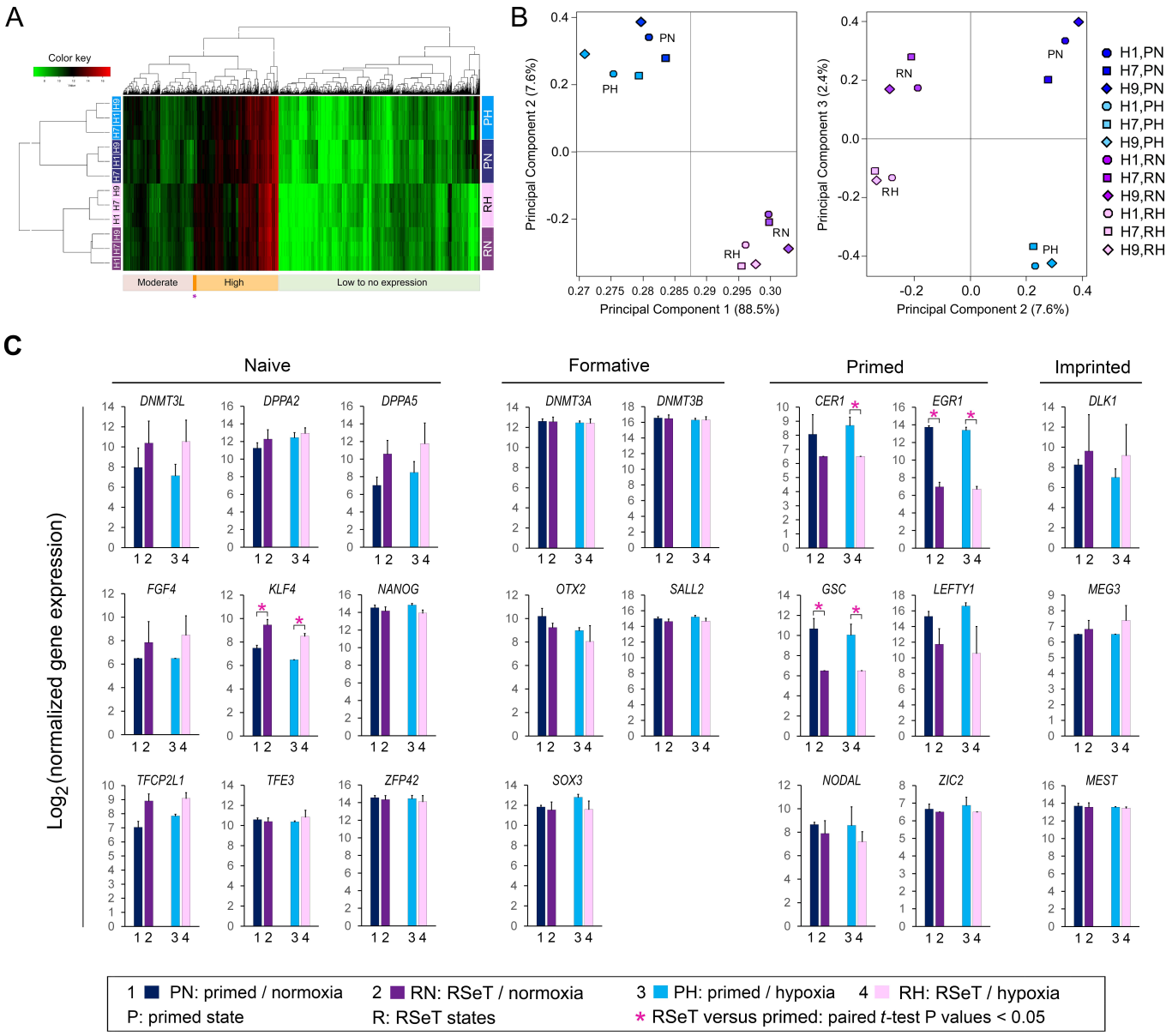
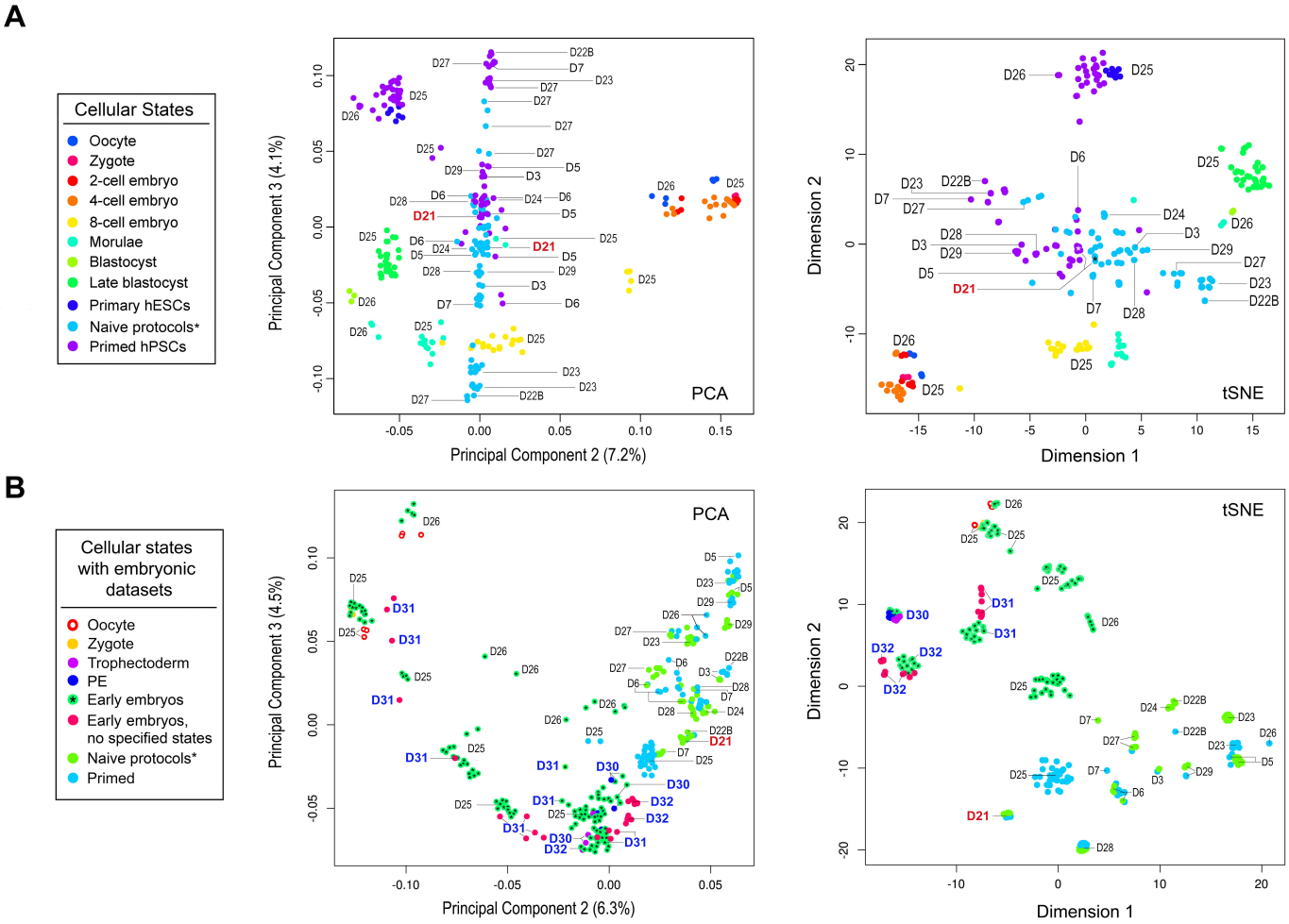
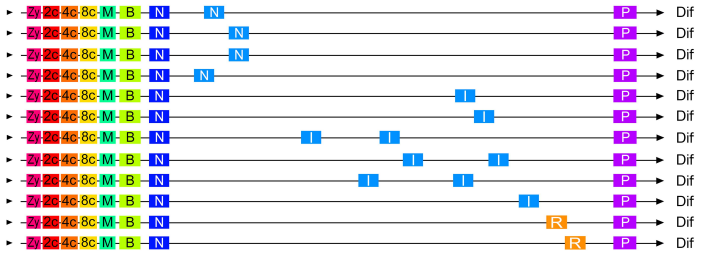

A
Color key
Moderate
High
Low to no expression
PH
PN
RH
RN
B
Principal Component 2 (7.6%)
Principal Component 1 (88.5%)
Principal Component 3 (2.4%)
Principal Component 2 (7.6%)
H1,PN
H7,PN
H9,PN
H1,PH
H7,PH
H9,PH
H1,RN
H7,RN
H9,RN
H1,RH
H7,RH
H9,RH
C
Naive
Formative
Primed
Imprinted
DNMT3L
DPPA2
DPPA5
DNMT3A
DNMT3B
CER1
EGR1
DLK1
FGF4
KLF4
NANOG
OTX2
SALL2
GSC
LEFTY1
MEG3
TFCP2L1
TFE3
ZFP42
SOX3
NODAL
ZIC2
MEST
Log2(normalized gene expression)
1 PN: primed / normoxia
2 RN: RSeT / normoxia
3 PH: primed / hypoxia
4 RH: RSeT / hypoxia
P: primed state
R: RSeT states
* RSeT versus primed: paired t-test P values < 0.05
D
NANOG (A1)
NANOG (A2)
POU5F1
GDF3
NODAL
LEFTY1
LEFTY2
CER1
Normalized gene expression (relative fold changes)
All H9
Primed
RSeT

Figure 3



C

Transcriptomic differences between naive and primed states



Datasets	Author_protocol	Transgene expression
D22B, D23	Takashima_t2iLGo	Yes, *KLF2* or *NANOG*
D27	Liu_t2iLGoY	Yes, OSKM in ct2iLGoY
D27	Liu_5iLAF	No, in c5iLAF
D7	Theunissen_5iLA	No
D28, D29	Zimmerlin_XAV939/2iL	No
D3	Chan_TeSR1/3iL	No
D24	Guo_t2iLGoY	No
D5	Gafni_NHSM	No
D6	Valamehr_FMM	No
D27	Liu_NHSM	No, in cNHSM
D27	Liu_RSeT	No, in cRSeT
D21	Chen_RSeT (this study)	No

Figure 4

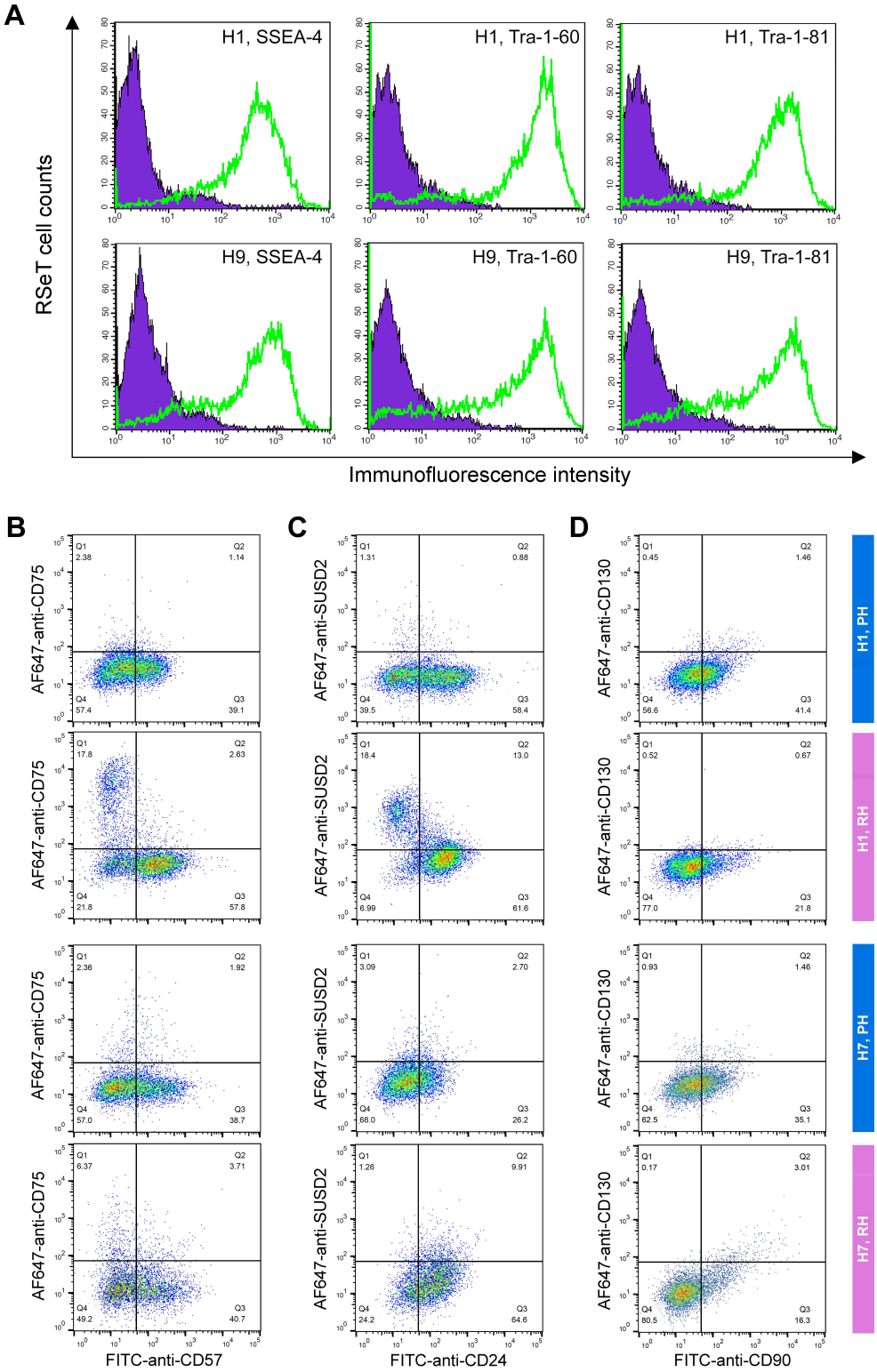

A
H1, SSEA-4
H1, Tra-1-60
H1, Tra-1-81
H9, SSEA-4
H9, Tra-1-60
H9, Tra-1-81
RSeT cell counts
Immunofluorescence intensity
B
C
D
AF647-anti-CD75
AF647-anti-SUSD2
AF647-anti-CD130
FITC-anti-CD57
FITC-anti-CD24
FITC-anti-CD90
H1, PH
H1, RH
H7, PH
H7, RH

Figure 5

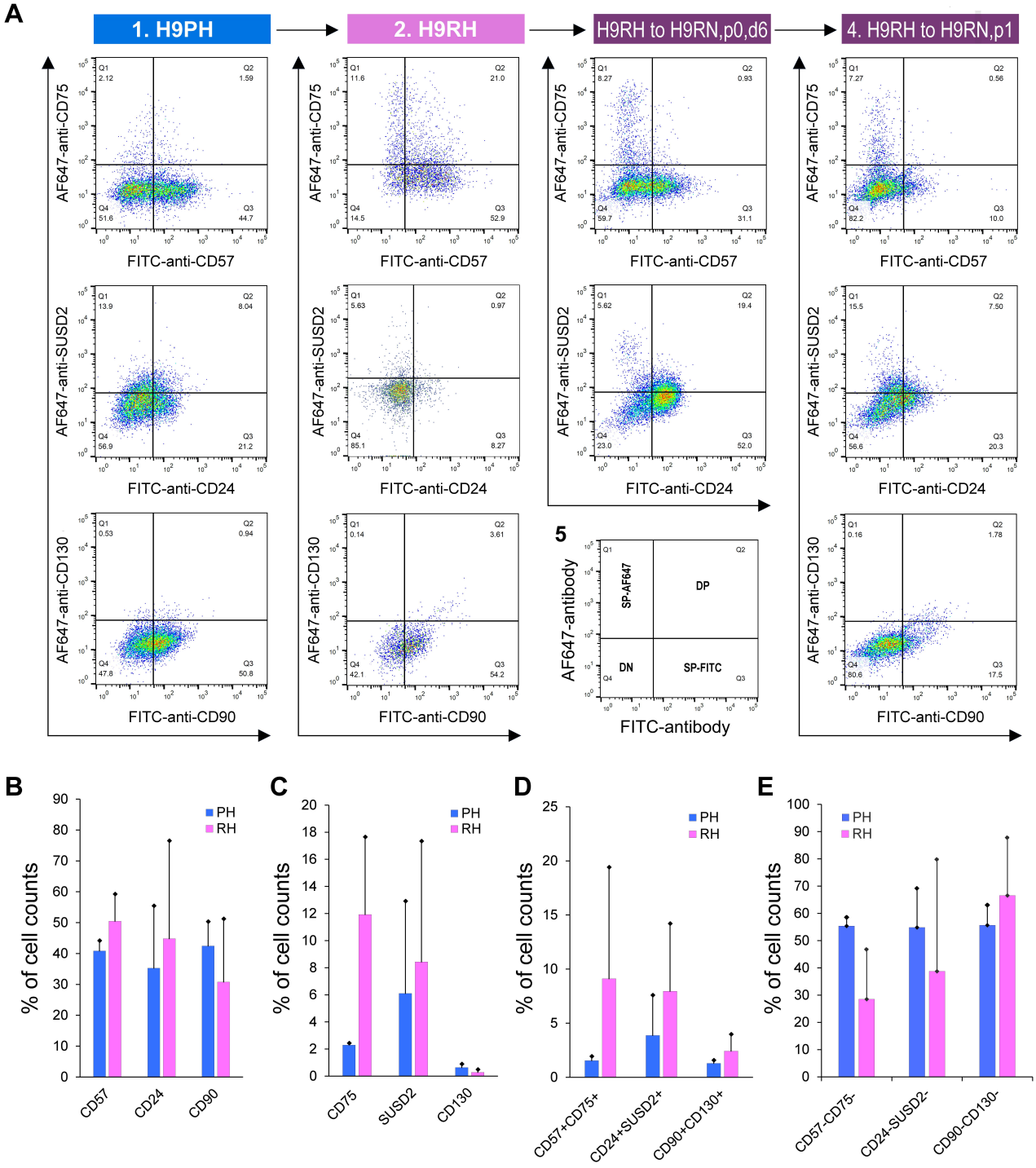

A
1. H9PH
2. H9RH
H9RH to H9RN,p0,d6
4. H9RH to H9RN,p1
AF647-anti-CD75
FITC-anti-CD57
AF647-anti-SUSD2
FITC-anti-CD24
AF647-anti-CD130
FITC-anti-CD90
5
AF647-antibody
FITC-antibody
SP-AF647
DP
DN
SP-FITC
B
C
D
E
% of cell counts
PH
RH
CD57
CD24
CD90
CD75
SUSD2
CD130
CD57+CD75+
CD24+SUSD2+
CD90+CD130+
CD57-CD75-
CD24-SUSD2-
CD90-CD130-

Figure 6

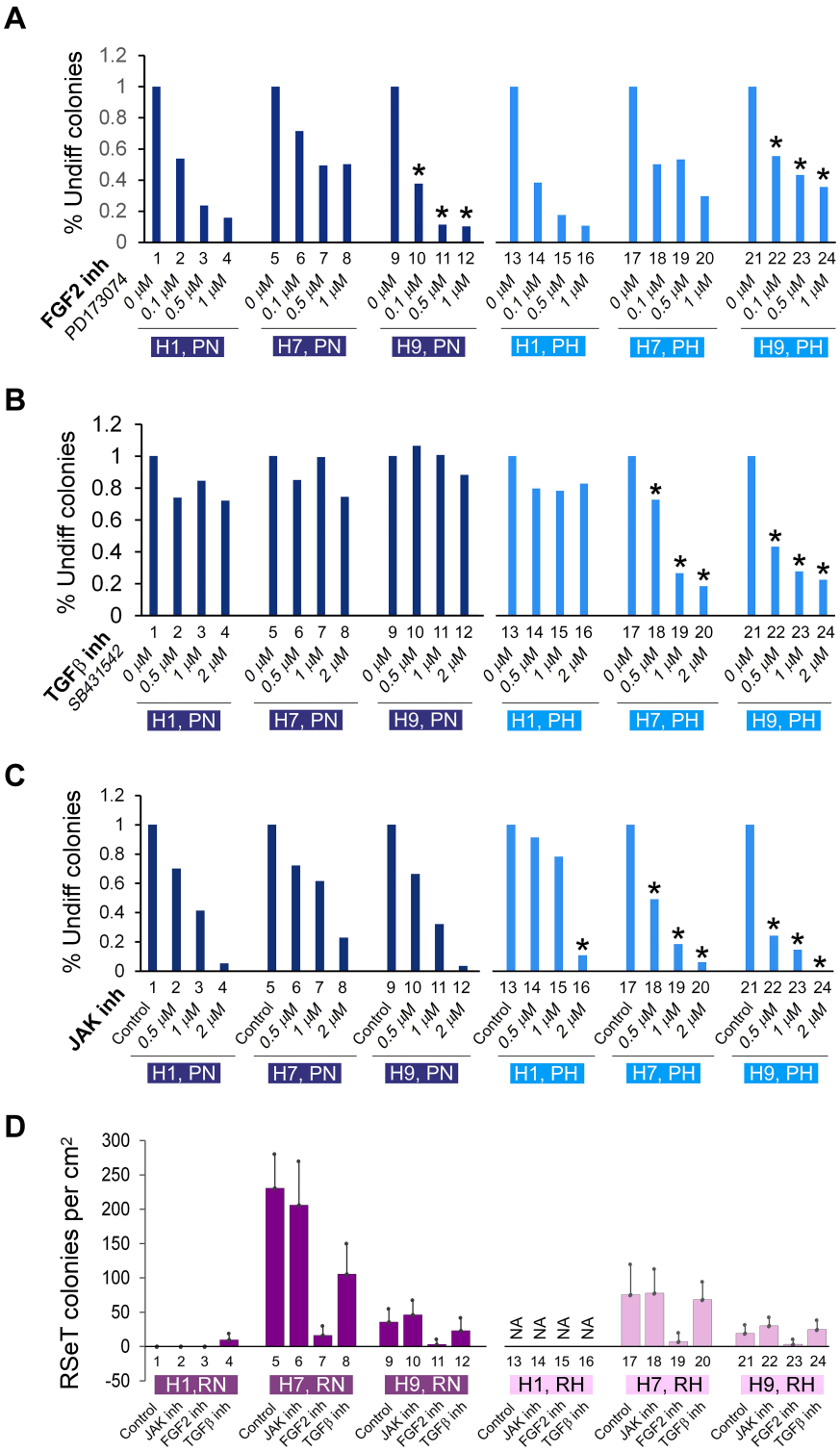
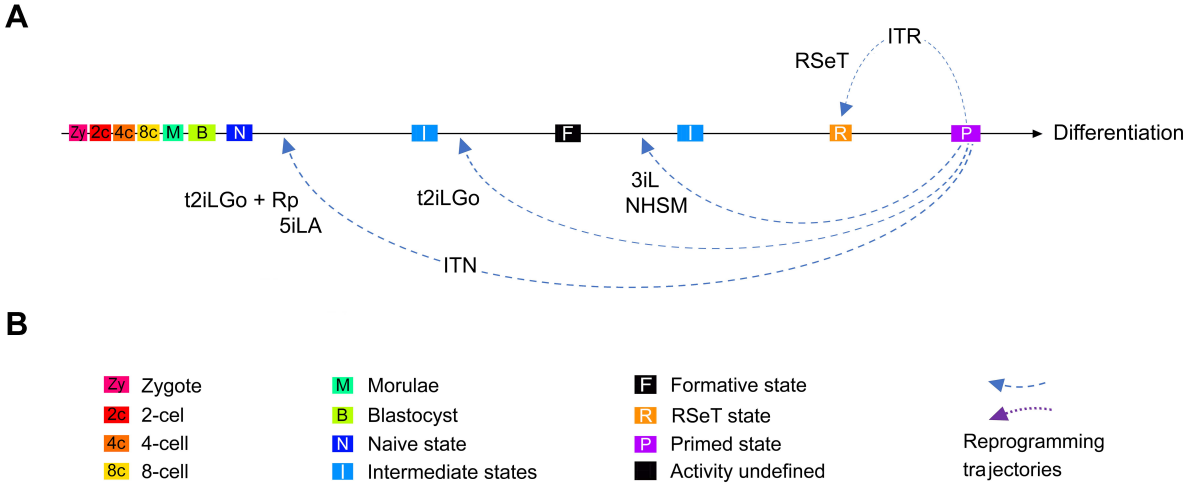

A
% Undiff colonies
FGF2 inh
PD173074
0 μM
0.1 μM
0.5 μM
1 μM
H1, PN
H7, PN
H9, PN
H1, PH
H7, PH
H9, PH
B
TGFβ inh
SB431542
2 μM
C
JAK inh
Control
D
RSeT colonies per cm²
H1,RN
H7, RN
H9, RN
H1, RH
H7, RH
H9, RH
NA
JAK inh
FGF2 inh
TGFβ inh

Figure 7


A
ITR
RSeT
Zy
2c
4c
8c
M
B
N
I
F
I
R
P
Differentiation
t2iLGo + Rp
5iLA
t2iLGo
3iL
NHSM
ITN
B
Zy Zygote
2c 2-cel
4c 4-cell
8c 8-cell
M Morulae
B Blastocyst
N Naive state
I Intermediate states
F Formative state
R RSeT state
P Primed state
Activity undefined
Reprogramming
trajectories