Esrrb is a cell cycle dependent priming factor balancing between pluripotency and differentiation

Sapir Herchcovici Levi [1,3], Lee Arnon [1,3], Sharon Feldman [1], Adi Alajem [1], Danny Bavli [1], Xue Sun [1], Yosef Buganim [2] and Oren Ram [1,4]

1. Department of Biological Chemistry, Alexander Silberman Institute of Life Sciences, The Hebrew University, Jerusalem, 91904, Israel.
2. Department of Developmental Biology and Cancer Research, Institute for Medical Research Israel-Canada, The Hebrew University, Hadassah Medical School, Jerusalem 91120, Israel.
3. These authors contributed equally to this work.
4. Correspondence should be addressed to O.R. (oren.ram@mail.huji.ac.il)

Key words: Exit from Pluripotency, Cell Cycle, Embryonic Stem Cells, Cellular differentiation and Lineage specification, Extraembryonic Endoderm Stem Cells, Epiblast Stem Cells, ChIP-seq and Single-cell RNA-seq.

Abstract

Cell cycle and differentiation decisions are tightly linked; however, the underlying principles that drive these decisions are not fully understood. Here, we combined cell-cycle reporter system and single-cell RNA-seq profiling to study the transcriptomes of mouse embryonic stem cells (ESCs) in the context of cell cycle states and differentiation. By applying a retinoic acid-based protocol as a differentiation assay representing exit from pluripotency, we show that Esrrb, a key pluripotent factor is upregulated during G2/M phase and promptly downregulated during exit from pluripotency. Enhancer chromatin states and Esrrb ChIP-seq map expose Esrrb's association with differentiation genes, were in the context of retinoic acid, enhancers associated with differentiation program driving extraembryonic endoderm cells (XENs). We show that only G2/M ESCs were capable of differentiating into XENs, whereas cells in the G1 phase predominantly produce epiblast stem cells (EpiSCs). Furthermore, Cells engineered to overexpress Esrrb allowed G1 ESCs to produce XENs, while Esrrb knockout ESCs completely lost their potential to differentiate into XEN. Interestingly, this phenomenon is unique to the pluripotency state as sorting cells after initiation of differentiation resulted in a cell cycle-independent differentiation decisions. Cells in both G1 and G2/M phases contributed equally to EpiSC and XEN cellular lineages. Taken together, this study reveals an important functional link between Esrrb and cell-cycle states during exit from pluripotency. Our novel approach of

applying single cell RNAseq in a cell cycle dependent manner can be further expand into other differentiation assays and expand our understanding of early differentiation events.

Introduction

Embryonic stem cells (ESCs) are pluripotent cells derived from the inner cell mass of the preimplantation blastocyst [1]. These cells possess unique properties of self-renewal and the ability to give rise to all definitive structures of the fetus. The first cellular decision distinguishes between the epiblast, which produces the embryo body, and the hypoblast, which contributes to the extraembryonic endoderm cells (XEN) [2 3].

The transition from an uncommitted to a differentiated state requires rapid and global execution of specific gene programs including the gradual silencing of pluripotency genes and the activation of lineage-specific genes. External signals sensed by each cell drive its fate decisions. ESCs must coordinately alter their transcriptomes, chromatin architectures, and epigenetic landscapes throughout the differentiation process [4–6].

The cell cycle is a critical processes in the development of an organism, and it is closely linked to cell-fate decisions [7,8]. Cell-cycle consists of four distinct phases dedicated to the replication and transmission of genetic material to daughter cells; in S-phase and M-phase cells, chromosome replication and chromosome transmission occur, respectively. These key events are separated by gap phases, G1 and G2, that serve as regulatory windows to ensure that cell-cycle events occur at the correct time and in the right order [9,10]. The ESC cell-cycle structure is characterized by a short G1 phase and a high proportion of cells in S phase [11]. This is associated with pluripotency factors that influence activities of cyclin-dependent protein kinases (CDKs) [12]. Studies have shown that the characteristic cell-cycle of ESCs is also affected by the culture conditions. ESCs cultured in serum lack G1 control and rapidly progress through the cell cycle, whereas ESCs grown in the presence of two small molecule inhibitors, MEK inhibitor (PD0325091) and Gsk3β inhibitor (CHIR99021), have a longer G1 phase [13].

Previous studies have shown that the cell-cycle stage is a major determinant of cell-fate decisions [8,11,14–18]. Evidence suggests that G1/S is the cellular stage at which differentiation decisions are made, whereas the G2 phase is mostly dedicated to mitosis control [19]. The intersection between cell-cycle regulation and cell-fate decision mechanisms involves developmental signals and CDK activities, which mediate cell-cycle dependent changes in the epigenetic landscape and chromosome architecture of developmental genes [20]. The CDKs are also responsible for recruitment of transcription factors to cell-fate related genes [20]. Activation of defined target genes determines cell-fate decisions and subsequent lineage commitment, as shown in studies of zebrafish and other model organisms [21]. Transcription

factor activities can counter CDK activities and drive the cells to exit the pluripotent state. Thus, the balance between CDK activity and transcription factor activity determines cell fate.

Retinoic acid (RA), the active metabolite of vitamin A, is crucial in early embryonic development and in maintenance of many organ systems in the adult organism [22]. It has been shown that RA represses pluripotency-associated genes and activates linage-specific markers in ESCs [23]. RA promotes a variety of lineage outcomes such as ectodermal [24–26], endodermal and extraembryonic endoderm (XEN) cells [27–30]. The ability of RA to promote various differentiation phenotypes implies that RA is involved in the switch between proliferation and differentiation [26].

Pluripotent cell identity is sustained by the activity of a highly interconnected network of transcription factors such as Oct4, Nanog, and Sox2 [31] and a large group of ancillary factors such as Estrogen related receptor beta (Esrrb) [32]. Esrrb is an orphan nuclear receptor that is required for self-renewal and pluripotency of ESCs [33]. In ESCs, Esrrb function is controlled by extrinsic cues mediated by kinases such as GSK3i [34] and intrinsic regulators such as Nanog [35]. This confers flexibility to the pluripotency network, as changes in the activity of these factors modulate the balance between maintenance and loss of pluripotency [36]. In the early post-implantation mouse embryo, Esrrb is specifically expressed in the extraembryonic ectoderm and plays a crucial role in trophoblast development [32].

Here, by combining a cell-cycle reporter system with single-cell RNA sequencing, we directly studied the cross-talk between cell-cycle state and cell-fate decisions during differentiation. Surprisingly, we found that the cell cycle has a deterministic effect on the differentiation capacity of cells. ESCs that are in G1 of the cell cycle, when exposed to RA, exclusively differentiate to EpiSCs and the more differentiated mesodermal state. Strikingly, if cells are exposed to RA while in the G2/M phase, they have the unique capacity to differentiate into XENs. We identified Esrrb as a central inducer of the XEN lineage and demonstrated Esrrb binding to poised XEN enhancers in ESCs enriched for G2/M state. Furthermore, by overexpressing Esrrb, we could switch on the XEN differentiation program in cells in G1 phase. These results demonstrate a yet unappreciated role for the cell cycle state in regulation of cellular differentiation and cell fate.

Results

ESCs differentiation is coupled with a decrease in proliferation and changes in cell-cycle state composition

In order to visualize and effectively distinguish between cell cycle states without complications of cell-to-cell variation, we produced a clonal ESC cell line expressing the FUCCI cell-cycle reporter system [37]. The FUCCI system is based on cycled translation and

ubiquitination-based degradation of Cdt1 (conjugated to mKO2, red) and Geminin (conjugated to mAG1, green). Immediately after mitosis, in early G1, cells have no color. Nuclei of late-stage G1 cells appear orange, and there is a gradual increase in the intensity of green as cell-cycle progresses, with a peak at G2/M (Fig. 1A). The insertion of the lentiviral-based FUCCI probes into ESCs did not result in detectable differences in pluripotency or in the ability of cells to differentiate. To induce differentiation, we grew ESCs in RA-containing media for 4 days. We chose the RA differentiation protocol for its broad affects on differentiation outcomes [26]. Upon RA treatment, cells underwent morphology changes typical of differentiation 38 (Fig. 1B). To ensure accurate separation between S and G2/M cellular states, we set the FACS sorting parameters using integration of both FUCCI and Hoechst staining of DNA. Confocal imaging and FACS analysis of the FUCCI reporters revealed that prior to addition of RA, approximately 43% of the cells were in the G1 phase; 33% in early G1 and only 10% in late G1 (Fig. 1C-1E). 11 Similar proportions were also measured in cells after 2 days of RA treatment. After 4 days of differentiation, however, the percentages were considerably different, with the majority of the cells in G1 states: 63% in early G1 and 29% in late G1. Less than 10% of the cells were in S or G2/M states. Differentiation is coupled with reduced replication rate: this cellular behavior aligns with changes in cell-cycle state (Fig. 1F and Fig. S1). As ESCs differentiate, the cells halt in G1 state, thus fewer replicating cells are captured 39. Proliferation

Figure 1

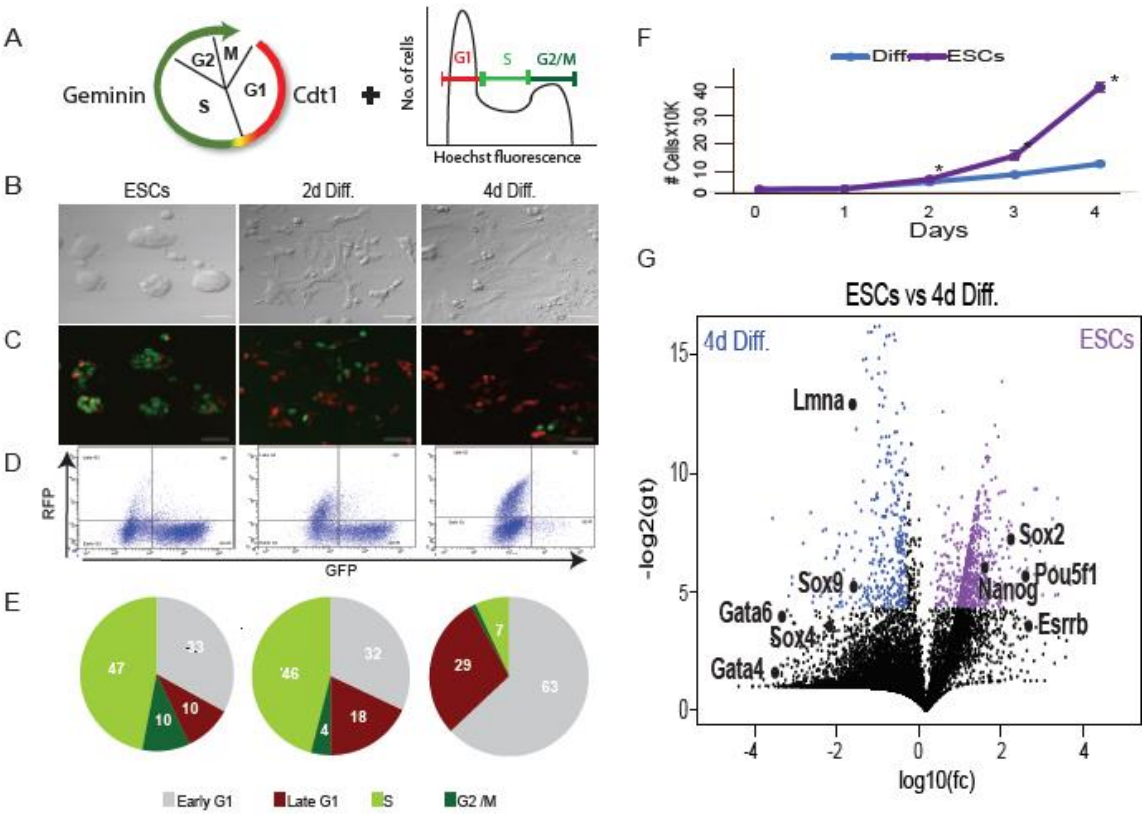

Figure 1: Clonal FUCCI system in differentiated ESCs. (**A**) Colors of FUCCI-expressing ESCs during the cell cycle. (**B-C**) ESCs were maintained in normal ESC medium and then switched to differentiation medium (1 µM RA without LIF and 2i) and analyzed after 2 and 4 days (2d Diff. and 4d Diff., respectively). **B**) Representative X20 bright-field images, scale bar 50 µm. **C**) Representative X20 confocal images, scale bar 50 µm. (**D**) Fluorescence-activated cell sorting of cells in the late G1 (Q1), and G2 (Q4) phases of the cell-cycle in the pluripotent state and after 2 and 4 days of RA treatment. (**E**) Pie charts of percentages of cells in indicated cell-cycle phases in pluripotent state and during differentiation calculated by a combination of FUCCI and Hoechst DNA staining. (**F**) Growth rates of ESCs and during treatment with RA (Diff.). Shown are mean values of two independent experiments ±SEM (* $p < 0.05$, Wilcoxon test). (**G**) RNA-seq expression profiles in ESCs and cells treated with RA for 4 days. The x axis represents expression fold changes in logarithmic scales and the y axis represents p values based on three replicates. Genes upregulated in ESCs are indicated in purple and those upregulated in differentiated cells are indicated in blue.

rate analysis by CellTrace verified that the rate of cell divisions was significantly slower in cells grown in differentiation conditions than in pluripotent cells from day 2 through day 4 (Fig. S1).

Over the 4-day period, we observed transcriptional activation of genes that regulate EpiSC- and XEN- like states in the embryo, such as *Sox4*, *Sox9*, *Lmna*, *Gata4*, and *Gata6*. Concomitantly, there was downregulation of pluripotency factors such as *Pou5f1*, *Nanog*, and *Esrrb* (Fig. 1G). Additional genes associated with differentiation or pluripotent state with altered expression over the time course are listed in Supplementary Table 1.

Extraembryonic endoderm differentiation is dependent on cell-cycle state of ESCs

To map the potential links between cell-cycle states of ESCs and differentiation trajectories, ESCs were sorted into early G1, late G1, and G2/M states. Immediately after sorting, we initiated differentiation by replacing the standard medium with medium containing RA (Fig. 2A). To analyze the heterogeneous population of cells resulting from RA treatment, we used the InDrop scRNA-seq system [40] to sequence the transcriptomes of cells from each condition. Based on two biological replicates and analysis of transcriptomes of about 2500 cells, we detected three main subpopulations (Fig. 2B, 2E and 2F): (1) EpiSCs (cluster 0), (2) extraembryonic endoderm lineage (clusters 1 and 2), and (3) mesodermal lineage (cluster 3). These results suggest that the differentiation capacity of ESCs is strongly influenced by the state of the cell-cycle at the point of induction. While early and late G1 cells show relatively similar differentiation capacities, G2/M cells show different differentiation trends (Fig 2C and 2D). G1 cells differentiated predominantly into cells with expression profiles characteristic of EpiSCs, with high expression of markers such as *Sox4*, *Vimentin*, *Ecm1*, *Smad6*, and *Gata2* (Fig. 2E and 2F). *Gata2* marks a specific mesodermal subpopulation[41]. On the contrary, only small proportion of G2/M ESCs were able to differentiate into EpiSCs (Fig. 2C and 2D). Most of the G2/M cells differentiated into cells that expressed markers of early endoderm and, intriguingly, some had characteristics of the subpopulation that initiates the XEN lineage including upregulation of *Gata4*, *Foxq1*, *Foxa2*, and *Dab2* (Fig. 2E and 2F).

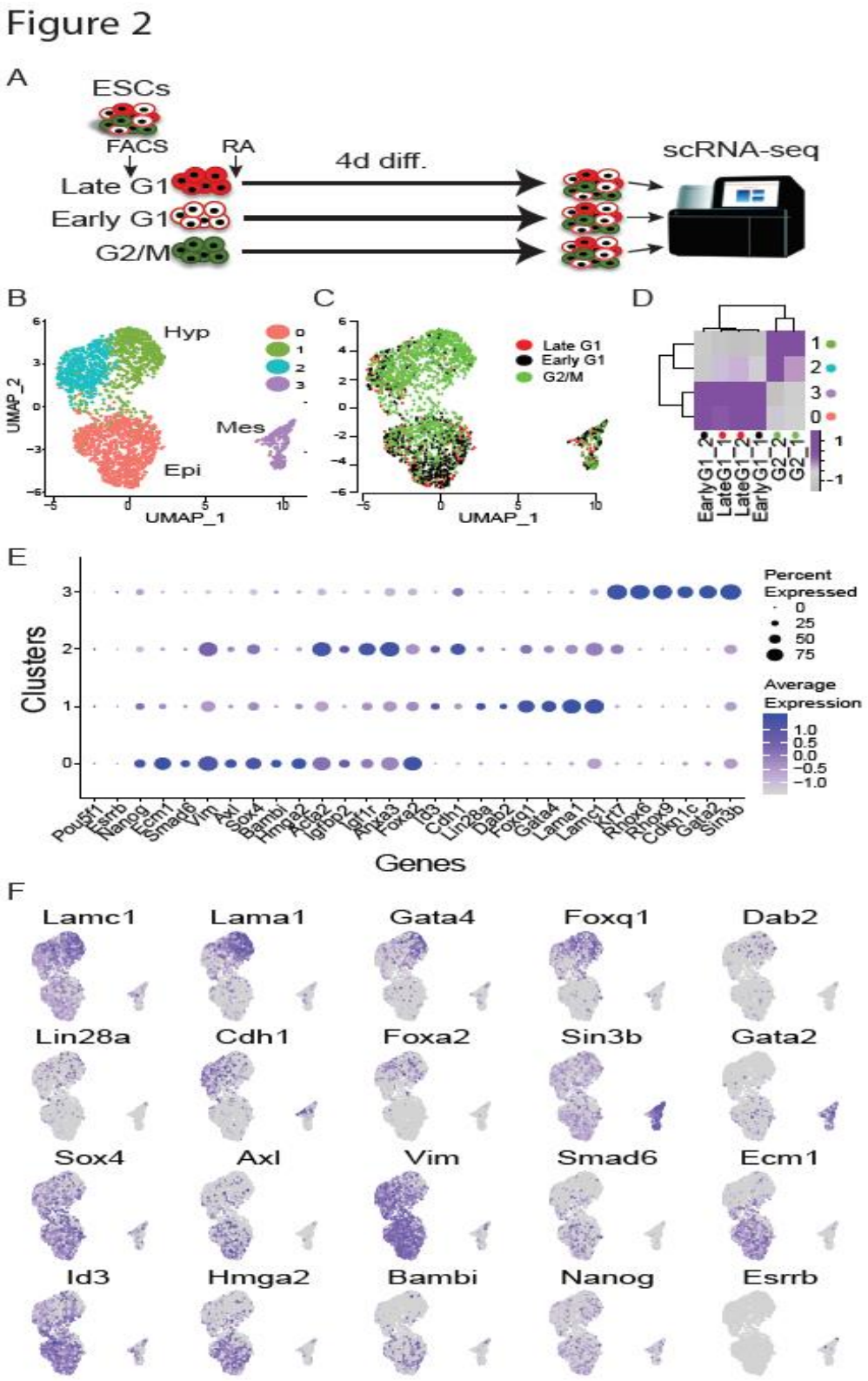

Figure 2: scRNA-seq analysis of sorted early G1, late G1, and G2/M after 4 days of RA treatment. (**A**) Schematic illustration of the sorting and scRNA-seq experiments. (**B**) UMAP visualization of ~2000 single cells clustered into four groups using Seurat pipeline [42]; groups correspond to EpiSC, two XEN states, and mesodermal like cluster. (**C**) UMAP visualization of single-cell data colored based on the

initial sorting of ESCs; early G1 are in black, late G1 are in red, and G2/M are in green. (**D**) Heatmap and hierarchical clustering of scRNA-seq from each of two replicates from early G1, late G1, and G2/M sorted populations treated for 4 days of differentiation in columns and the scRNA-seq based clusters labeled as in panel B in the rows. (**E**) Dot plot of differentially expressed genes that best explain the four scRNA-seq clusters. Dot size indicates percent of cells that express the gene and gray to purple scale indicates average expression within the cluster. (**F**) UMAP plots of expression of indicated genes in EpiSC, XEN and mesodermal like clusters as shown in panels B and C. Gray to purple scale (as in panel E) indicates average gene expression in each single cell.

Esrrb is a potent inducer of the XEN lineage

To identify genes expressed by pluripotent cells that promote XEN differentiation in the context of cell cycle, we first profiled bulk mRNA of cells in the G1 and G2/M states during pluripotency. Based on three replicates, we detected 2100 genes present at higher levels and 404 genes present at lower levels in G2/M compared to late G1 cells (Fig. 3A, Supplementary Table 1). The majority of these genes were cell-cycle regulators, and, as a result, only cell-cycle GO terms were significantly enriched. Some of the identified genes are involved in regulation of pluripotency and early differentiation. For example, genes expressed at higher levels in ESCs in the G1 state compared to G2/M included *Nanog*, *Otx2*, *Leafty1*, and *Pax3*, which are central regulators of EpiSCs initiation [43]. In the group of genes that show higher expression in G2/M cells we did not find any known direct drivers of the XEN lineage. However, we identified two potential candidates: *Sall4* and *Esrrb* (Fig. 3A). *Sall4* is expressed in both ESCs and XEN cells and encodes a transcription factor that regulates expression of key XEN lineage-associated genes such as *Gata4*, *Gata6*, *Sox7*, and *Sox17* [44]. *Esrrb* encodes a pluripotent transcription factor that is involved in regulation of the trophectodermal lineage [45] and, together with Oct4, Sox2, and Klf4, in maintaining ESCs by mitotic bookmarking of pluripotent genes [46,47]. In addition, Esrrb is associated with Tfe3, a non-cycled factor that regulates the hypoblast/PrE circuitry[48], and together with Gata3, Eomes, Tfat2c and Myc induces a XEN-like state during reprogramming [49].

Esrrb expression is quickly downregulated during exit from pluripotency. Therefore, to be involved in XEN regulation of differentiation, we hypothesized that Esrrb should specifically mark poised enhancers of XEN genes. To test this hypothesis, we mapped H3K4me1 and H3K27ac enhancers of ESCs and cells after 4 days of differentiation. We also obtained and re-analyzed Esrrb ChIPseq maps of interphase ESCs [50]. Both histone marks and Esrrb obtained from ESCs grown on LIF-serum medium condition without 2i. As a result, ESC G1 state is significantly reduced (<10%) while G2/M state enriched [13,39]. Indeed, when candidate enhancers of key XEN markers were examined (e.g. Gata6, Gata4, Foxa2, Dab2 and

Foxq1), Esrrb was found to be associated with many of them, which were expectedly poised during pluripotency (H3K4me1 positive and H3K27ac negative) but were activated (H3K27ac positive) upon differentiation (Fig. 3B).

Next, to further determine and validate Esrrb involvement in G2/M-specific XEN induction, we produced FUCCI ESCs that constitutively overexpress Esrrb fused to YFP. We used FACS to select late G1 cells that express moderate levels of Esrrb, as indicated by YFP signal. We also imaged these selected cells by confocal microscopy to verify nuclear expression of Esrrb-YFP (Fig. 3C) and validated Esrrb over expression using western blot (Fig. S4). After selection, we differentiated the cells for 4 days by treatment with RA and then analyzed 3024 cells from three biological replicates with scRNA-seq. Clustering and UMAP visualization of the data indicate that the subpopulations observed after differentiation of the late G1 cells that express Esrrb-YFP include the same subpopulations observed for wild-type G2/M cells with EpiSCs, *Gata2*-expressing mesodermal cells, and a smaller population of XEN cells (Fig. 3D). This suggests that Esrrb is involved in inducing the XEN state in a cell cycle dependent manner.

Figure 3

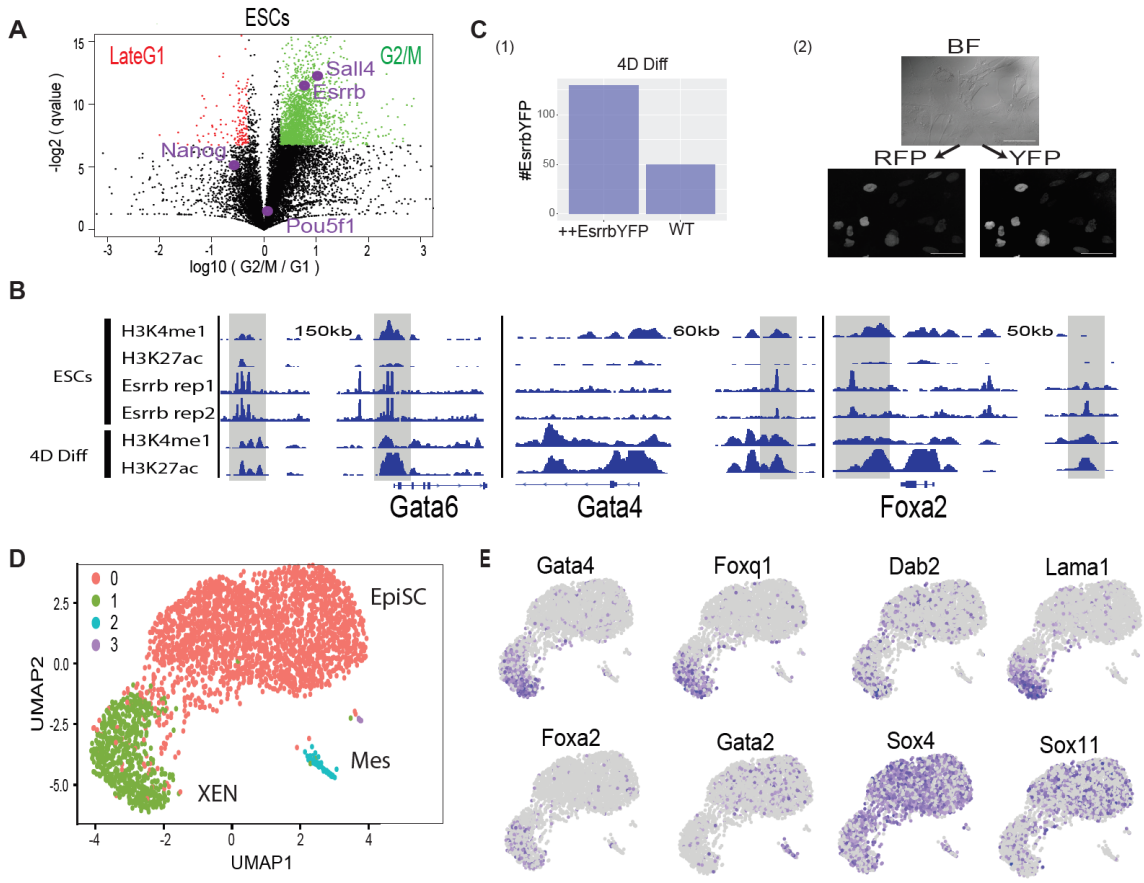

Figure 3: Esrrb is an inducer of XEN like cells. (**A**) Bulk RNA-seq data of ESCs in the late G1 state (left) vs. ESCs in the G2/M state (right). The fold change in G2/M vs. G1 ESCs is plotted vs. t-test p values (n=3). *Esrrb*, *Sall4*, *Nanog*, and *Pou5f1* genes are marked showing G2/M upregulation of Esrrb

and Sall4 expression levels. (**B**) ChIP-seq IGV tracks of H3K4me1, H3K27ac and G2/M enriched ESCs (grown without 2i) and differentiated cells. Esrrb (Festuccia N. *et al.* [46]) are shown only for ESCs. Enhancer and promoter regions of XEN markers Gata6, Gata4 and Foxa2 bound by Esrrb marked in grey boxes. Candidate distal enhancers are indicated by Nkb genomic fragments removal. (**C**) (1) FUCCI cells non infected vs infected with *Esrrb-YFP* lentiviral construct. Y axis represent expression levels in ESCs based on RNA-seq read counts of lentiviral transcripts. (2) Confocal imaging of cells treated with RA showing EsrrbYFP and RFP G1 in the same cells. **Upper-** Representative X20 bright-field image, scale bar 50 µm. Lower- Representative X20 confocal images, scale bar 50 µm. (**D**) UMAP visualization of 3025 single cells clustered into three groups using Seurat pipeline [42] corresponding to EpiSC in red, XENs in green, and mesoderm like in blue. (**D**) UMAP plots highlighting expression levels of indicated genes. Gray (low) to purple (high) scale indicates average expression signal.

Finally, to determine the function of ERRβ (*Esrrb*) during the exit from pluripotency, we differentiated ESCs lacking ERRβ for 4 days with RA. The KO was done using CRISPR/CAS9 targeting ERRβ[51]. We validated ERRβ KO using western blot shown in figure 4A and S4. Excitingly, regardless to ESC cell cycle states, ESCs KO contributed only to the formation of EpiSC (Fig. 4B-D). This highly supports our overall observation that EpiSC and XEN differentiation potential are dictated by Esrrb which in WT ESCs is being upregulated during G2/M state. In addition, pluripotent genes such as *Nanog* and *Pou5f1* (*Oct4*) shown higher expression levels in the EpiSC state, suggesting that EpiSC is a less differentiated state comparing to XEN. In addition, in line with our previous single cell experiments of differentiated ESCs, a small population of Gata2 cells (mesodermal like) could also be detected even though unbiased clustering did not highlighted it as a standalone cluster.

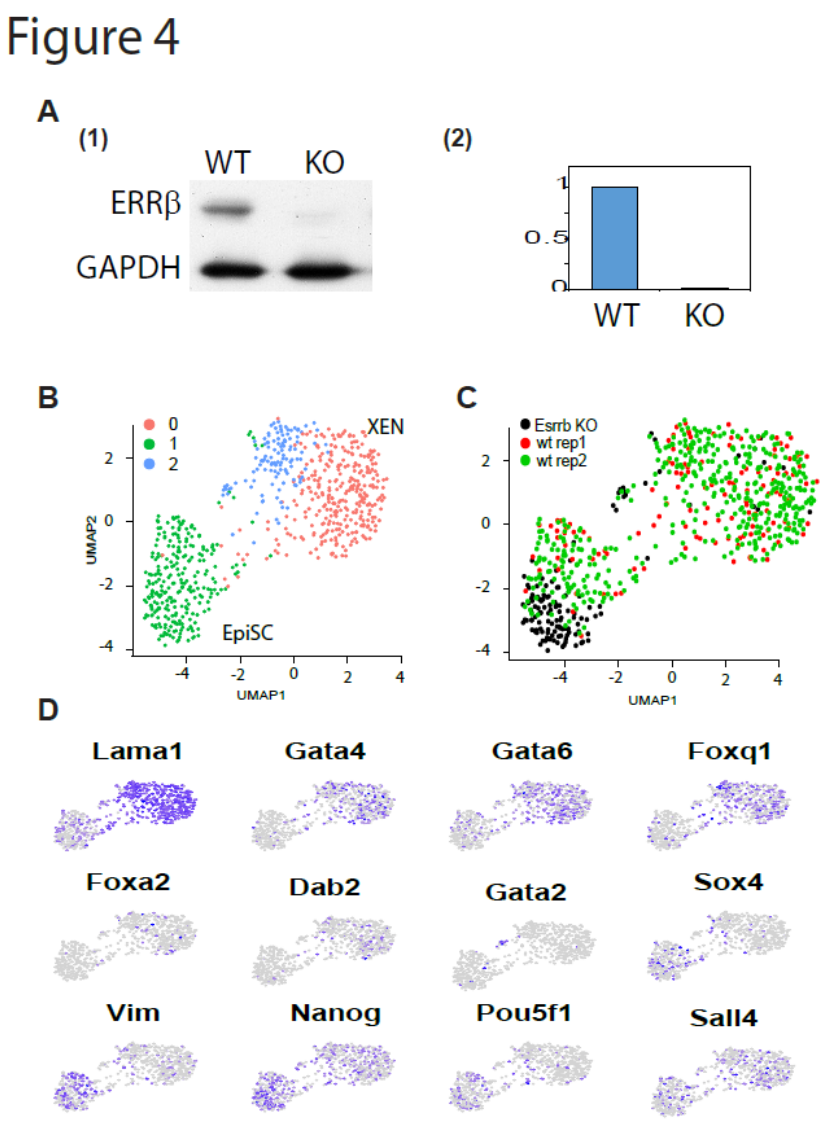

Figure 4: ERRβ knockout ESCs cannot differentiate to XEN like cells. (**A**) Western blot of WT and ERRβ (*Esrrb*) KO ESCs in (1) and fold change of ERRβ translation levels based on imageJ in (2) (**B**) UMAP visualization of 700 single ESCs cells after 4 days of differentiation with RA clustered into three groups using Seurat pipeline [42] corresponding to EpiSC in green and XEN states in red and blue. Blue cluster 2 also includes Gata2 positive mesodermal cells. (**C**) UMAP visualization of single-cell data as shown in 'B', colored in red and green for two replicates of WT ESCs and black for ERRβ KO ESCs. (**D**) UMAP plots highlighting expression levels of selected genes. Gray (low) to purple (high) scale indicates average expression signal.

During cellular differentiation, linage outcomes are not affected by cell-cycle states

During early differentiation, when progenitor cells emerge, the cells replicate constantly. Therefore, we aimed to explore whether cell-cycle states also influence fate

decisions in cells that already initiated the differentiation process. To that end, we treated ESCs with RA for 2 days, and then sorted the cells based on cell-cycle state as described above. Sorted cells were subjected to single-cell RNA sequencing directly, or allowed to differentiate for additional 2 days followed by scRNA-seq (Fig. 5A). To validate the accuracy of our cell-cycle separation using FACS, we first focused the analysis on previously identified cell-cycle genes [52]. We calculated the principal component matrix, performed clustering [53], and visualized the results using tSNE. The cells clustered based on our FUCCI/Hoechst sorting gates and GO annotations show a very clear separation based on G1 and G2/M states (Fig. 5B).

Unbiased single-cell analysis of cells sorted after 2 days of differentiation revealed three major clusters (Fig. 5C). Interestingly, G1 and G2/M cells were present at similar

Figure 5

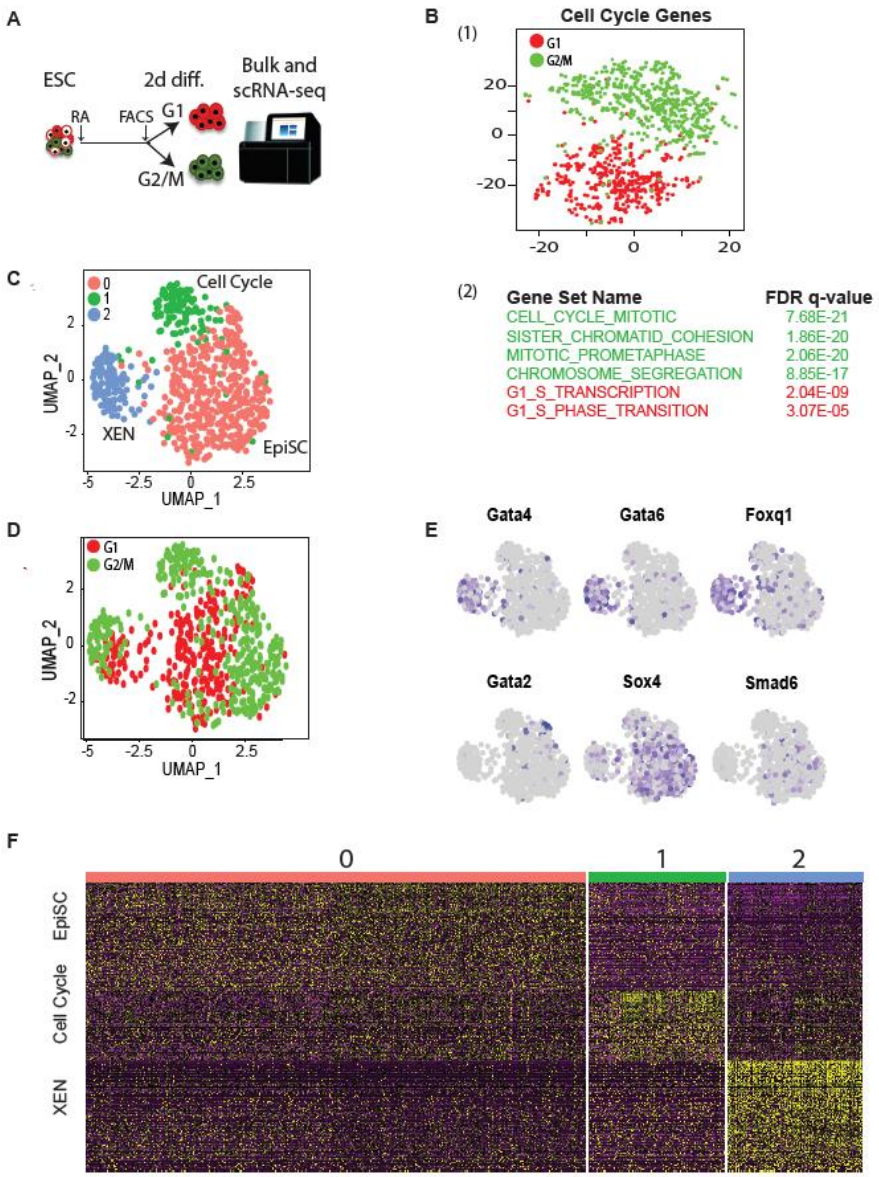

Figure 5: **Bulk and scRNA-seq analysis of sorted late G1 and G2/M cells after 2 days with RA.** (**A**) Schematic illustration of the sorting and scRNA-seq experiments. (**B**) Cells were treated with RA for 2 days and sorted into G2/M and late G1 subpopulations prior to bulk RNA-seq. The x axis shows fold change of expression levels between G2/M cells and G1 cells, and the y axis shows t-test p values (n=2). (**C**) (1) tSNE visualization and MsigDB annotations of single-cell data analyzed using only known cell-cycle genes. The two clusters marked in red and green are based on the MsigDB annotations for G1 vs. G2/M cell-cycle states. (2) Gene set names of G1 and G2/M clusters with their FDR-q value (**D**) UMAP visualization of sequencing data from 940 single cells clustered using the Seurat pipeline [42] corresponding to EpiSC in red, XEN cells in blue and cell cycle related subpopulation in green. (**E**) UMAP visualization of single-cell data colored by late G1 cells in red and G2/M cells in green. (**F**) UMAP plots highlighting expression levels of indicated genes. Gray (low) to purple (high) scale indicates average expression signal. (**G**) Heatmap of scaled expression based on Seurat normalization and scaling pipeline[42] of the marker genes for each of the clusters as defined in fig.4C.

frequencies in two of the clusters (Fig. 5D). Cells in cluster 0 were characterized by expression of genes associated with earlier stages of EpiSC differentiation, and cluster 1 cells showed a XEN signature, which includes upregulation of both *Gata4* and *Gata6* (Fig. 5E and 5F). Cluster 2, predominantly composed from G2/M cells, showed reduced differentiation state compared to cells of cluster 0, as suggested by expression levels of *Gata2*, *Sox4* and *Smad6* which are associated with EpiSC state (Fig. 5E and 5F). This is in agreement with previous reports that indicated that the G2/M state is more associated with pluripotency than the G1 state[54].

Interestingly, single cell profiling of cells at day 4 of RA treatment, sorted on the basis of cell-cycle state at day 2 (Supplementary Fig.2A), show two major clusters of EpiSC and XEN like cells (Supplementary Fig. 2B) with similar cell cycle derived proportion of EpiSC and XEN differentiation (Supplementary Fig. 2C and 2D). These results suggest that at 2 days following the differentiation signal, the differentiation potential is no longer influenced by the cell cycle state. However, in EpiSC cluster 0 (Supplementary Fig. 2B), a thin separation between G1 and G2/M cells could be detected. G1 derived cells expressed higher levels of *Vimentin*, *Ecm1*, and *Acta2*,[55,56] which are associated with later stages of differentiation, whereas G2/M expressed higher levels of *Nanog*, which is associated with premature differentiation status[57] (Supplementary Fig. 2D). This result aligned with bulk expression of G1 vs G2/M cells sorted after 4 days of differentiation, showing strong signature of cell cycle genes in both subpopulation of cells and upregulation of differentiation signature associated only with G2/M cells, suggesting that G2/M cells include cells less matured in their differentiation state (Supplementary Fig. S3). Taken together, our study suggests that the differentiation fate of ESCs is strongly influenced by the cell-cycle state at the moment of exposure to the

differentiation signal. However, once cells start to differentiate and cellular commitments are made, the cell-cycle becomes insignificant to fate decisions.

Discussion

The molecular mechanisms that specify embryo and extraembryonic germ layer identities are only partially understood. Different transcription factors act to regulate extensive networks in ESCs [38,57,58]. Gata, Fox, and Sox protein families, among others, are associated with early differentiated states [59]. Factors that trigger exit from pluripotency remain largely unknown and this process is thought to be regulated by "stochastic" events. To identify whether the cell cycle state plays a role in this process, we used the FUCCI system [7] that allowed us to monitor cell-cycle progression in living cells. Using this approach, we followed ESCs along the path of early differentiation. The pluripotent state was characterized by a high percentage of S-phase cells and rapid proliferation. During the 4-day treatment with RA, there was a gradual increase in the number of G1 cells and proliferation rate slowed down. The difference in proliferation rate between treated and untreated cells was significant after 2 days, and the rate was dramatically slower under differentiation conditions after 4 days. In addition, by integrating the FUCCI system with DNA staining, we were able to narrow the error range that typically characterizes cell-cycle separation methods, thus allowing more accurate and cleaner sorting outcomes than obtained with either system alone.

Our bulk transcriptome analysis showed that G2/M cells express higher levels of cell-cycle genes and lower levels of differentiation-related genes than do cells in the G1 state. This suggests that the cells isolated on the basis of cell-cycle markers are not homogeneous and may include cells in different differentiation states. In support of this, after 2 days of differentiation a small subpopulation of G2/M cells were captured that express genes associated with a less differentiated state. Our analysis of gene expression in these populations suggests that the G2/M population contains a higher proportion of slow responders to the differentiation signals than the G1 population.

Our results demonstrate that cell-cycle states of ESCs at the moment they are exposed to differentiation signals dictate the decision to differentiate into XENs or EpiSCs. We further demonstrated that Esrrb is a key factor that regulates the XEN differentiation pathway. In agreement with this data, a previous study showed that Esrrb, in conjunctions with Gata3, Eomes, Tfap2c and Myc, can induce pluripotency by the activation of a unique XEN-like state [60]. In addition, Esrrb regulates expression of many transcription factors that are critical for maintaining pluripotency and self-renewal [60]. Betschinger et al. showed that there is significant

overlap between chromatin binding of Esrrb and Tfe3, which is a key regulator of the hypoblast/PrE circuitry [61]. Our analysis revealed that *Esrrb* is specifically upregulated in pluripotent ESCs in the G2/M state. We functionally validated that Esrrb is an inducer of the XEN lineage by overexpressing Esrrb: In G1 ESCs that overexpress Esrrb, differentiation into XEN cells was observed after treatment with RA; in G1 cells that do not overexpress Esrrb, no XEN-like cells were detected. Thus, the integration of Esrrb into the core transcriptional network of G1 cells stimulated XEN initiation and overcame cell-cycle dependency. In addition, ERRβ KO ESCs lack the potential to form XEN like cells at all, emphasizing that Esrrb is a key regulator in XEN differentiation and cell cycle dependent expression of Esrrb allow only G2/M cells to become XEN. Other known members of the XEN differentiation circuitry are Tfap2c, Sox17, Eomes, and Cdx2 [62]. However, we did not detect differential expression of mRNAs encoding these proteins in our scRNA-seq experiments. Moreover, these genes are expressed at very low levels in ESCs, indicated that they are not likely to be involve in the exit from pluripotency step.

Sall4 is critical for XEN differentiation through regulation of expression of *Gata4*, *Gata6*, *Sox7*, and *Sox17*, and due to interconnections in the pluripotent regulatory circuitry with Oct4, Sox2, and Nanog [44]. We observed that *Sall4* expression has a prominent cell-cycle dependency, and its expression was upregulated during G2/M phase in ESCs. The Sall4 chromatin binding profile is correlated with that of Esrrb, suggesting that Sall4 and Esrrb may be co-regulated during XEN differentiation. Transcription factor recruitment can be modulated by epigenetic modifications to the chromatin [63,64], and cell-cycle-specific Esrrb enhancer occupancy could be regulated by methylation of DNA and covalent modification of histone proteins. Further experiments are needed to identify additional epigenetic players that may be involved in the exit from pluripotency specifically to the XEN state.

Interestingly, down regulation of *Esrrb* at the mRNA level was specific to late G1 cells and not to early G1. Comparison of *Esrrb* expression in early and late-stage G1 cells showed that 156 genes were upregulated in early G1 (p value $2x10^{-40}$) and that these genes were enriched for those involved in mitotic cell-cycle regulation. This suggests that, in terms of transcriptional state, early G1 is highly similar to G2/M. High levels of *Esrrb* and *Sall4* were detected in early G1 as well, suggesting a gradual decrease in Sall4 and Esrrb transcripts. Surprisingly, like late G1 cells, early G1 cells only contributed to EpiSC state, so the positive association between *Esrrb* and XEN is not correlated directly to mRNA levels. We therefore hypothesize that the observed *Esrrb* upregulation during G2/M and downregulation during late G1 involves other regulatory proteins in this pathway.

Overall, the presented results support the tight association between cell-cycle stage and cell-fate determination. We demonstrated that the cell-cycle state affect linage specification only at the exit from pluripotency and although the complete cellular signaling is yet to be comprehensively revealed, Esrrb plays a key role in this regulatory pathway. We therefore suggest a model for which Esrrb accumulation during interphase, from G1 towards G2, expands its enhancer binding capacity towards XEN poised enhancers and potentially involves co-binding of Sall4. Thus, allows the preparation of G2/M specific ESCs towards XEN differentiation (Fig. 6).

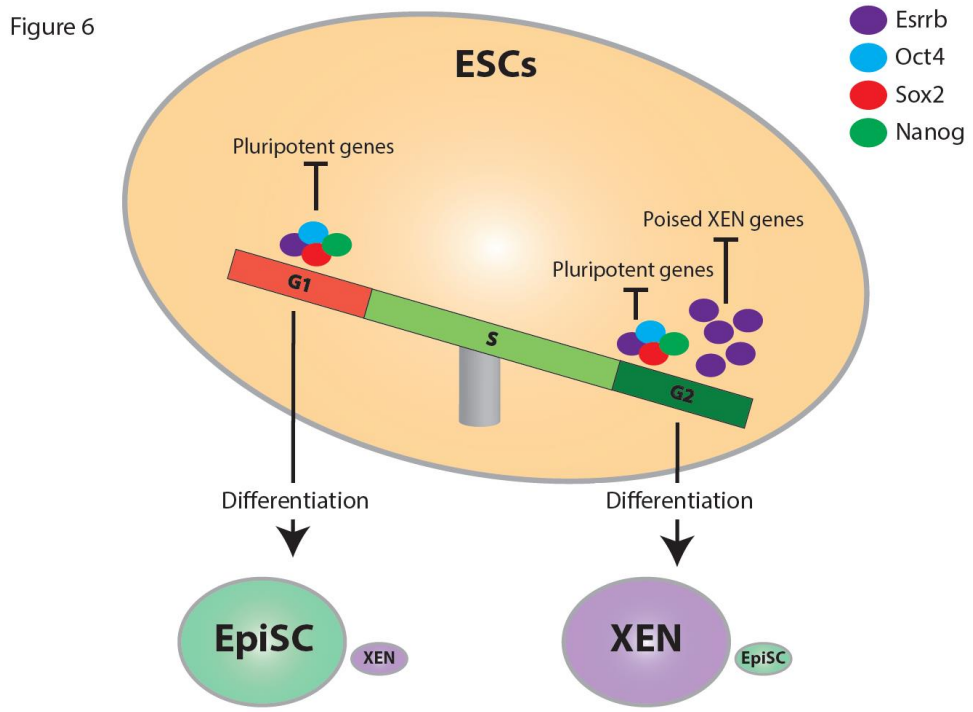

Figure 6: A model for Esrrb dependent XEN induction during ESCs G2/M cell cycle state. Esrrb accumulates in ESCs during interphase with low expression level in G1 and high expression during G2. When cells are in G1, Esrrb limited abundance is dedicated for pluripotency maintenance through co-binding with pluripotent factors Oct4, Sox2 and Nanog. Excess accumulation of Esrrb during G2/M, allows Esrrb to bind XEN specific enhancers potentially with Sall4.

Further validation experiments and expanding our cell cycle based single cell RNA-seq protocol on other differentiation pathways such as trophoblast, neuronal and other linage specific differentiations will allow to further increase our understanding of the link between cell cycle and differentiation outcomes exposing other cell cycle or cycling genes that are central for early ESC differentiation decisions.

Acknowledgments

O.R. is supported by research grants from the European Research Council (ERC, # 715260 SC-EpiCode), the Israeli Center of Research Excellence (I-CORE) program, the Israel Science Foundation (ISF, #1618/16), and Azriely Foundation Scholar Program for Distinguished Junior Faculty. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 765966 - EpiSyStem. We offer special thanks to Prof. Eran Mesorer, Dr. Yonatan Tzur, Prof. Itamar Simon from The Hebrew University for helpful discussions and critical reading of the manuscript.

Author Contributions

S.H.L., L.A., S.F., and O.R. conceived the study and prepared the figures. S.H.L., L.A., S.F., Y.B., and O.R. designed the experiments. S.H.L., L.A., and S.F. performed tissue culture, FACS sorting, and library preparation for bulk and single-cell RNA-seq. D.B. and X.S. prepared the microfluidics system and performed scRNA-seq experiments. O.R. preformed the computational analysis for bulk and single-cell RNA-seq data. S.H.L., L.A., S.F., and O.R. wrote the manuscript.

Author Information

All bulk and single-cell RNA-seq data will be deposited in the Gene Expression Omnibus database (GEO). Reviewers, please use this link: in process. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to O.R. (oren.ram@mail.huji.ac.il).

Materials and Methods

Cell culture

Mouse R1 ESCs were a gift from A. Nagy (Lunenfeld-Tanenbaum Research Institute at Mount Sinai Hospital, Toronto, Ontario, Canada). ESCs were seeded on 0.1% gelatin-coated plates (Sigma-Aldrich, G1393) and grown in ES medium (Dulbecco's modified Eagle's medium (Sigma-Aldrich, D5671), 15% fetal bovine serum (FBS, Biological Industries, 04-007-1A), 1 mM sodium pyruvate (Biological Industries, 03-042-1B), 0.1 mM nonessential amino acids (Biological Industries, 01-340-1B), 0.1 mM β-mercaptoethanol (Sigma-Aldrich, M3148), 1000 U/ml LIF (Mercury, ESG1107)). The cells were maintained at an incubator at 37 °C with 5% CO_2 humidified air. For pluripotent conditions, 3 μM CHIR99021 (PeproTech, SM-2520691-B) and 0.2 μM PD0325901 (PeproTech, SM-3911091-B) were added to the ES medium. Ectodermal differentiation was induced using RA (Sigma-Aldrich, R2625) dissolved in dimethyl sulfoxide (ENCO, 0219605580) at 1 μM in LIF-free media with 10% FBS, on 0.1%

gelatin-coated plates. Cells were passaged every 2 to 3 days using trypsin EDTA (Biological Industries, 03-050-1A).

Establishment of cell line stably transfected with FUCCI vector plasmids

For cell cycle-phase visualization, the FUCCI (for fluorescent ubiquitination-based cell-cycle indicator) expression system was used [37]. Plasmids expressing mKO2-hCdt1 (orange-red fluorescent protein) or mAG-hGem (green fluorescent protein) were a kind gift from Prof. Itamar Simon (The Hebrew University, Jerusalem, Israel). 293T cells were kindly gifted from Prof. Meshorer (The Hebrew University, Jerusalem, Israel). Plasmids were transfected into the 293T cells using TransIT-LT1 Transfection Reagent (Mirusbio, MC-MIR-2300) at a ratio of 1:7:3.5 of PMD2.G to psPAX.2 to FUCCI plasmids. The medium was changed after 4 hours to ES medium. After 24 hours, the samples were filtered through a 0.45-mm filter, and cells were resuspended in polybrene-supplemented medium (8 µg/ml PB (Sigma-Aldrich, 107689) in ES medium). After 24 hours, the medium was again replaced with PB-supplemented medium. The cells were then subjected to a two-step FACS sorting. In the first step S, G2, and M cells were sorted by green fluorescence (488 filter) using a FACSAria II cell sorter (Becton Dickinson). The sorted cells were then reseeded and after a week were sorted for red fluorescence (461 filter). For clonal selection, colonies were trypsinzed and seeded in 96-well plates at a density of 1 cell per well.

Establishment of cell line stably transfected with Esrrb expression vector

Plasmid for expression of Esrrb fused to YFP was a kind gift from Prof. Yosef Buganim (The Hebrew University, Jerusalem, Israel). Plasmids were transfected into 293T p.3 cells using TransIT-LT1 Transfection Reagent (Mirusbio, MC-MIR-2300) at a ratio of 1:7:3.5 of PMD2.G to psPAX.2 to Esrrb-expression plasmid. Medium was changed after 4 hours to ES medium. After 24 hours, samples were filtered through a 0.45-mm filter, and cells were resuspended in PB-supplemented medium. This step was repeated after 24 and 72 hours.

Fluorescence-activated cell sorting (FACS)

For cell-cycle analysis, cells were trypsinzed and washed twice in PBS. Next, cells were stained with solution containing Hoechst dye (Sigma-Aldrich, B2261). A solution of 9 µl of 25 mg/ml stock of dye in 1 ml of PBS with 2% FBS was used for $5x10^6$ to $6x10^6$ cells. The G1 and S/G2 population were sorted using FACSAria II cell sorter (Becton Dickinson) in PBS supplemented with RNase inhibitor (NEB, M0314L). Sorted samples were processed for scRNA-seq or cDNA-seq.

Quantification of cellular proliferation rates

To assess cellular proliferation, ESCs cultured with RA were stained with CellTrace Violet Cell Proliferation Kit (Rhenium, C34571) following the manufacturer's instructions for labeling of adherent cells. Stained cells were analyzed using FACS and by confocal imaging.

mRNA extraction

mRNA extraction was performed using Invitrogen Dynabeads mRNA DIRECT Purification Kit (61011) according to the protocol included with the kit.

Library preparation

The library for differential gene expression analysis was prepared using 100-1000 ng of mRNA. First, RNA was fragmented using an RNA fragmentation kit (Ambion, AM8740) following the manufacturer's protocols. Fragmented RNA was purified with 1.4x reaction volume of AMPure XP beads (Beckman Coulter, A63881) and eluted in 10 µl TE buffer (10 mM Tris–HCl, pH 8.0, 0.1 mM EDTA). The resulting fragmented RNA was reverse transcribed using an oligo dT primer: First, 10 µl of RNA were mixed with 2 µl of 50 µM oligo dT primer (Bio-Rad, 1725038) and 1 µl of 25 mM dNTP mix (NEB, N0447S), and samples were incubated for 3 minutes at 65 °C and transferred to ice. The following components were added to the reaction, for a total volume of 20 µl: 4 µl of 5x First-Strand SuperScript III buffer (Invitrogen, 18080044), 2 µl of 0.1 M DTT (Invitrogen, 18080044), 1 µl murine RNase inhibitor (NEB, M0314L), and 1 µl of 200 U/µL SuperScript III RT enzyme (Invitrogen, 18080044). The reaction product (in the form of a cDNA:RNA hybrid) was purified with 1.5x reaction volume of AMPure XP beads (Beckman Coulter, A63881) and eluted in 16 µl TE buffer as described above. For second-strand synthesis, 16 µl of the digestion reaction product were combined with 2 µl second-strand synthesis (SSS) buffer and 1 µl of SSS enzyme mix of mRNA Second Strand Synthesis Module kit (NEBNext, E6111S) and incubated at 16 °C for 2.5 hours, followed by 20 minutes at 65 °C. The resulting product was purified with 1.5x reaction volume of AMPure XP beads (Beckman Coulter, A63881) and eluted with 20.4 µl TE buffer. cDNA edges were repaired using End-It DNA End-Repair Kit (Danyel Biotech, ER81050) by combining the 20.4 µl sample with 3 µl 10X End-repair Buffer, 3 µl 2.5 mM dNTPs, 3 µl 10 mM ATP, and 0.6 µl END-IT enzyme mix at room temperature for 45 minutes, followed by purification with 1.5x reaction volume of AMPure XP beads (Beckman Coulter, A63881) and elution with 32 µl TE buffer. Klenow-mediated addition of an adenine to the 3' end of the DNA fragments was performed using the Klenow fragment (3'→5' exo-) kit (NEB, M0212L) by combining the 32 µl sample with 5 µl 10X Klenow Buffer NEB 2, 10 µl 1 mM dATP, and 0.6 µl 5 U/µl Klenow (3'-5' exo-) at 37 °C for 30 minutes, followed by purification

with 1.5x reaction volume of AMPure XP beads (Beckman Coulter, A63881) and elution with 8 µl TE buffer. Illumina library adaptors were added to the resulting cDNA fragment using DNA ligase (NEB, M2200S) by incubation with 5 µl 2x Ligase Buffer, 0.5 µl adapter Oligo mix (1:10 in H_2O), and 0.5 µl 1 U/ml DNA Ligase at room temperature for 15 minutes. The ligated product was purified with 1.5x reaction volume of AMPure XP beads (Beckman Coulter, A63881) and eluted in 11.5 µl TE buffer. The resulting libraries were PCR amplified using standard PE1/PE2 full-length primer mix containing Illumina library indices for multiplexing (sequences are given in Supplementary Table X). Each PCR reaction contained 11.5 µl post-RT cDNA library, 12.5 µl 2x KAPA HiFi HotStart PCR mix (Zotal, KK-KK2601), and 1 µl of 25 µM 2p Fixed (+barcode) and 2p Fixed primer mix (Supplementary Table 3). Amplified libraries were purified using 0.7x reaction volume of AMPure XP beads (Beckman Coulter, A63881) and eluted in 32 µl TE buffer. A 15-µl aliquot of each resulting library was run in a 2% agarose gel, and size selection for the desired 200-800 bp DNA library fragments was performed using PureLink DNA gel extraction kit (Invitrogen, K210012). Library quality was confirmed by Agilent 2200 TapeStation Nucleic Acids System.

The in vitro transcription sequencing (IVT) library was prepared using 75-100 ng of mRNA, extracted as described above. First-strand synthesis was performed using SuperScript III Reverse Transcriptase (Invitrogen, 18080044) and the T7RTPolyT primer (Supplementary Table X) at 50 °C for 2 hours followed by 15 minutes at 70 °C and 1 minute at 4 °C. To remove unused primers and primer dimers, the reaction product was combined with 20 µl digestion mix containing 5 µl ExoI (NEB, M0293S), 5 µl HinFI (NEB, R0155S), 5 µl ExoI buffer (NEB, B0293S), 5 µl CutSmart buffer (NEB, B7204S), and 30 µl nuclease-free water and incubated for 1 hour at 37 °C and 10 minutes at 80 °C. The reaction product (in the form of a cDNA:RNA hybrid) was purified with 1.5x reaction volume of AMPure XP beads (Beckman Coulter, A63881) and eluted in 13.5 µl TE buffer. For second-strand synthesis, the 13.5 µl digestion reaction product was combined with 1.5 µl SSS buffer and 1 µl of SSS enzyme mix from the mRNA Second Strand Synthesis Module kit (NEBNext, E6111S) and incubated at 16 °C for 2.5 hours, followed by 20 minutes at 65 °C. For linear amplification by in vitro transcription, 16 µl of SSS reaction products were combined with 24 µl HiScribe T7 High Yield RNA Synthesis Kit (NEB, E2040S) reagent mix containing 4 µl T7 Buffer, 4 µl ATP, 4 µl CTP, 4 µl GTP, 4 µl UTP, and 4 µl T7 enzyme mix. The reaction was incubated at 37 °C for 13 hours, and the resulting RNA was purified with 1.3x reaction volume of AMPure XP beads (Beckman Coulter, A63881) and eluted with 20 µl TE buffer. Aliquots of 9 µl of each sample were frozen for backup at -80 °C, 2 µl of each sample was directly analyzed, and the remaining 9 µl were used in subsequent library preparation steps. RNA was fragmented using the RNA fragmentation kit (Ambion, AM8740) by combining 9 µl sample with 1 µl of RNA

fragmentation reagent and incubating at 70 °C for 2 minutes. After transfer to ice, 40 µl fragmentation stop mix containing 5 µl fragmentation stop solution and 35 µl TE buffer was added. Fragmented RNA was purified with 1.4x reaction volume of AMPure XP beads (Beckman Coulter, A63881) and eluted in 10 µl TE buffer. The resulting amplified and fragmented RNA was reverse transcribed using a random hexamer primer (IDT) as follows: First, 10 µl of RNA were mixed with 2 µl of 100 µM PvG748-SBS12-RT random hexamer primer (Supplementary Table X) and 1 µl of 10 mM dNTP mix (NEB, N0447S), incubated for 3 minutes at 65 °C and transferred to ice. The following components were then added to the reaction for a total volume of 20 µl: 4 µl of 5x First-Strand SuperScript III buffer (Invitrogen, 18080044), 1 µl 0.1 M DTT (Invitrogen, 18080044), 1 µl murine RNase inhibitor (NEB, M0314L), and 1 µl of 200 U/µL SuperScript III RT enzyme (Invitrogen, 18080044). Following reverse transcription, the reaction volume was raised to 50 µl by adding 30 µl nuclease-free water, and the resulting cDNA was purified with 1.2x reaction volume of AMPure XP beads (Beckman Coulter, A63881) and eluted in 11.5 µl TE buffer. Library amplification was performed using 12.5 µl 2x KAPA HiFi HotStart ReadyMix PCR kit (Zotal, KK-KK2601) and 1 µl of 25 µM 2p fixed primers (Supplementary Table X). Amplified libraries were purified using 0.7x reaction volume of AMPure XP beads (Beckman Coulter, A63881) and eluted in 30 µl nuclease-free water. Aliquots of 15 µl of each library were run in 2% agarose gels, and size selection for the desired 200-600 bp DNA library fragments was performed using PureLink DNA gel extraction kit (Invitrogen, K210012). Library quality was confirmed using the Agilent 2200 TapeStation Nucleic Acids System (Agilent).

For scRNA-seq, encapsulation of cells with Reverse Transcription mix (IGEPAL CA-630- Sigma-Aldrich I8896-50ML , SuperScript III Reverse Transcriptase Invitrogen 18080044, Deoxynucleotide (dNTP) Solution Set Ornat N0446S, D,L-DITHIOTHREITOL Bio-Lab 000448235200, TRIS-HCL 1 M STOCK SOLUTIONS PH 7.5 B 1L Sigma-Aldrich T2319-1L, RNase Inhibitor, Murine NEB M0314L) and gels with unique molecular identifiers was performed using the FLUIGENT Smart Microfluidics Pump System with the following flow rate parameters: Cells 100 µL/hour, RT mix 100 µL/hour, BHMs [barcoding hydrogels] 10-20 µL/hour, Oil flow 80 µL/hour. The 4-nL drops were released at a frequency of 15 droplets per second. To release photocleavable barcoding primers from the barcoding beads, the collection tubes were exposed to 6.5 J/cm^2 of 365-nm light for 10 minutes. Next, to the collection tubes containing the UV-exposed emulsion was added SuperScript III Reverse Transcriptase (Invitrogen, 18080044), and samples were incubated at 50 °C for 2 hours followed by 15 minutes at 70 °C and 1 minute at 4 °C. Each sample was then demulsified by addition of 50 µl perfluoro-1-octanol (Sigma-Aldrich 370533-25G) to release the barcoded cDNA from the

droplets. The aqueous phase containing the barcoded cDNA (~50 µl) was processed for IVT library preparation as described above.

ChIP

ChIP Antibodies (H3K4me1 abcam ab8895, H3K27Ac abcam ab4729, H3K36me abcam ab9050, 2 µg antibody per 2x107 cells) were incubated with 25 µL protein A Dynabeads (Thermo Scientific) previously washed twice with blocking buffer (0.5% BSA, 0.5% Tween-20 in PBS) for 2 h at 4 °C in blocking buffer. The conjugated beads were washed twice with blocking buffer before adding the lysate. Lysates were prepared as follows: Cells were trypsinized, washed with PBS, and resuspended with 500 µL PBS per 2x107 cells. The same volume of lysis buffer with 100 U/mL MNase was added, and the cells were incubated for 10 min on ice and then for 15 min at 37 °C. Reactions were stopped by adding 20 mM EGTA followed by centrifugation (20,000 g, 2 min). Supernatants were added to the conjugated beads, and samples were incubated overnight with rotation. Supernatants were removed, and the beads were washed twice with RIPA buffer (10 mM Tris-HCl [pH 8], 140 mM NaCl, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS, 1 mM EDTA), twice with RIPA buffer high salt (10 mM Tris-HCl [pH 8], 360 mM NaCl, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS, 1 mM EDTA), twice with LiCl wash buffer (10 mM Tris-HCl [pH 8], 250 mM LiCl, 0.5% sodium deoxycholate, 1 mM EDTA, 0.5% IGEPAL CA-630 [Sigma]), and twice with 10 mM Tris-HCl [pH 8]. The beads were resuspended in 45 µL 10 mM Tris-HCl [pH 8] and 20 µg RNase A. The beads were incubated for 30 min at 37 °C, and 60 µg proteinase K was added. The beads were further incubated for 2 h at 37 °C. Proteinase K was inactivated by heating to 65 °C for 15 min, and 50 µL elution buffer (10 mM Tris-HCl [pH 8], 300 mM NaCl, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS, 1 mM EDTA) was added, and the samples were incubated in 65 °C for 15 min. DNA was purified from supernatants using Ampure XP beads (Beckman Coulter). DNA was eluted with 22 µL 10 mM Tris-HCl [pH 8].

ChIP library preparation

DNA samples were converted to blunt-ended, phosphorylated DNA using the End-It DNA End-Repair Kit (Lucigen). DNA was purified using ratio of 1:1.8 sample to AMPure XP beads. Adenosine was added to the 3' end of the DNA fragments using Klenow (3'-5' exo-) (New England Biolabs). DNA was purified using ratio of 1:1.8 sample to AMPure XP beads. Ilumina adapters were added by ligation using DNA ligase (New England Biolabs). DNA was purified using ratio of 1:1.2 sample to AMPure XP beads. DNA samples were amplified using KAPA HiFi HotStart (Roche) with 25 µM PE1 and PE2 full-length primers containing Illumina library indices for multiplexing (Table 1). Amplified libraries were purified using 0.7x reaction volume

of AMPure XP beads and eluted in 22 µL 10 mM Tris-HCl [pH 8]. Aliquots of 15 µL of resulting libraries were run in a 2% agarose gel, and the desired 200-600 bp DNA library fragments were selected and isolated using the PureLink DNA gel extraction kit (Invitrogen). Library quality was confirmed using the Agilent 2200 TapeStation nucleic acids system and the Agilent High Sensitivity D1000 DS DNA kit. The resulting libraries had an average size of 350-550 bp. Size-selected libraries were diluted to 4 nM concentration and combined for paired-end, single index sequencing on the Illumina NextSeq 550 instrument using an illumina 550 High Output v2 (75 cycles) kit. Cycle distribution was 45 cycles for Read 1, 35 cycles for Read 2, and 8 cycles for the library index read.

ChIP-seq data analysis

Sequencing data were aligned using Bismark and Bowtie (https://www.bioinformatics.babraham.ac.uk/projects/bismark/Bismark_User_Guide.pdf) using paired-ended approach. TDF genomic browser files were produced using IGV count. We applied HOMER to find peaks using ChIP-seq criteria and used BEDTools to intersect bins with genomic intervals such as promoters, genes, and predicted enhancers.

Western Blot and Antibodies

Proteins were separated by SDS-PAGE on 4%–20% polyacrylamide gradient gels and transferred to 0.45-µm nitrocellulose membranes (iBlot2, PVDF, mini Transfer Stacks, Thermo Scientific; IB24002v). The membranes were incubated with the appropriate primary and secondary antibodies and washed with PBS-Tween 20. Horseradish-peroxidase-conjugated secondary antibodies were detected by SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific; PI-34080). Antibodies used were anti-human ERRβ/NR3B2 (R&D systems; PP-H6705-00), anti-GAPDH (Abcam; AB-ab8245) Goat Anti-Mouse (Jackson ImmunoResearch; 115-035-062).

Supplementary Table 3. Primer sequences

Primer name	Sequence (5' to 3')
2p Fixed (+barcode)	CAAGCAGAAGACGGCATACGAGATNNNNNNNNGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
2p Fixed	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT
T7RTPolyT	CGATGACGTAATACGACTCACTATAGGGATACCACCATGGCTCTTTCCCTACACGACGCTCTTCCGATCT NNNNNNNNNNTTTTTTTTTTTTTTTTTTTVN
PvG748-SBS12-RT	AGACGTGTGCTCTTCCGATCTNNNNNN

Deep sequencing

Deep sequencing was carried out on an Illumina NextSeq using commercially available kits from Illumina (Danyel Biotech FC-404- 2005) following the manufacturer's protocols.

Data analysis

The Illumina output was analyzed using an in-house Perl script that produced a reads matrix that was aligned using RSEM [65] with Bowtie [66]. The resulting matrix was analyzed in R. For bulk data analysis the transcript per million (TPM) values were used to compare between libraries. Differential gene expression was visualized using volcano plots. Statistical analysis was performed for the two replicates using a two-sided t-test, and p values of <0.05 were deemed significant.

scRNA-seq data was analyzed using the Seurat v2.4 pipeline [67]. First, cells with more than 10,000 unique molecular identifiers were retained for further analysis. A global-scaling normalization was performed on the filtered dataset using "LogNormalize" with a scale factor of 10,000. Identification of highly variable genes was performed with the following parameters: x.low.cutoff = 0.2, x.high.cutoff = 5, y.cutoff = 0.5, and y.high.cutoff = 10. Cell-to-cell variation in gene expression driven by batch, cell alignment rate, and the number of detected molecules were regressed out and a linear transformation was applied. A principal component analysis was performed on the scaled data with 12 principal components. Clustering was done with resolution of 0.6, and tSNE or UMAP was using for visualization.

Gene Set Enrichment Analysis

Gene Set Enrichment Analysis was done using GSEA software [68,69] with an false negative discovery q value <0.01.

Statistical analyses

Chi-square goodness-of-fit test was performed using "chisq.test" function in R, with the following parameters:

	G1-phase	**G2-phase**
Specific cluster: X	No. of G1 cells in X	No. of G2 cells in X
All other clusters	Total no. of G1 cells - no. of G1 cells in X	Total no. of G2 cells - no. of G2 cells in X

For small numbers, Fisher's exact test was performed instead.

Supplementary

Table 1 – Bulk Gene expression analysis during differentiation. Excel sheet is attached.

Table 2- Single Cell Gene expression analysis during differentiation. Excel sheet is attached.

FigS1

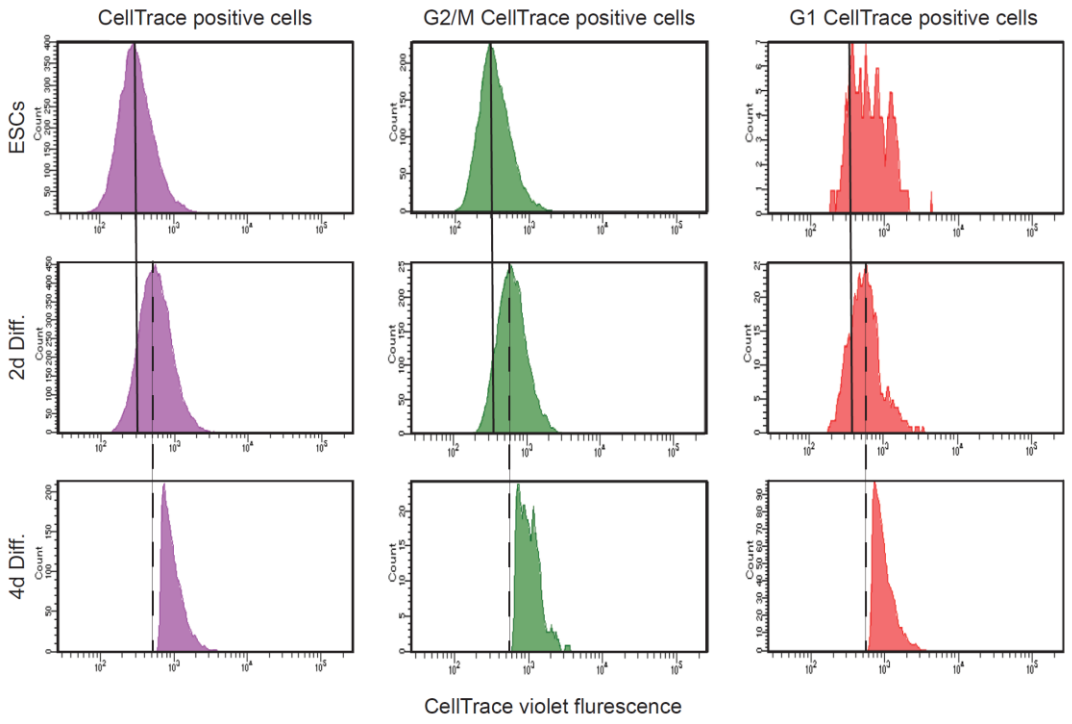

Fig.S1 – CellTrace Violet proliferation flow cytometry analysis of FUCCI expressing cells. ESCs were stained with CellTrace Violet (CTV) dye and then treated with RA for 48h (middle panel) and for 96h (lower panel). The vertical line indicates the mean fluorescence for ESCs. Dashed line indicates mean fluorescence for 2d differentiated cells. Left panel- cell proliferation histogram of all CTV positive cells. Middle panel- CTV positive cell proliferation histogram of G2/M cells. Right panel- CTV positive cell proliferation histogram of G1 cells.

FigS2

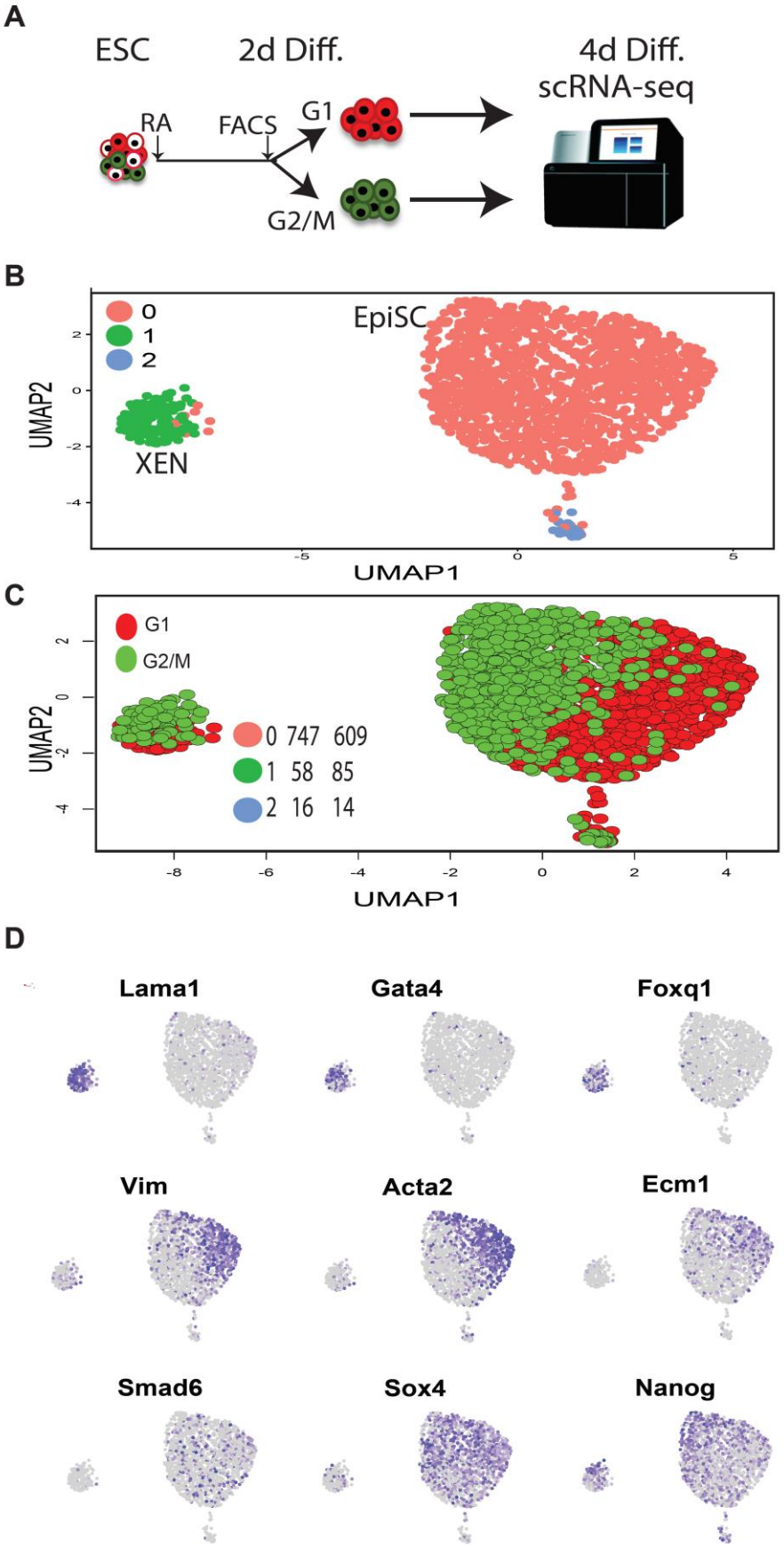
A
ESC
2d Diff.
4d Diff.
scRNA-seq
RA
FACS
G1
G2/M
B
0
1
2
EpiSC
XEN
UMAP2
UMAP1
C
G1
G2/M
0 747 609
1 58 85
2 16 14
UMAP2
UMAP1
D
Lama1
Gata4
Foxq1
Vim
Acta2
Ecm1
Smad6
Sox4
Nanog

Fig.S2- (**A)** Schematic illustration of the sorting and scRNA-seq experiments. **(B)** UMAP visualization of sequencing data from 1110 G1 and 1862 G2/M single cells sorted on day 2 and clustered using the Seurat pipeline [42] on day 4 of the differentiation. Clusters correspond to EpiSC in red and XEN like cells in green. **(C)** UMAP visualization of single-cell data colored by late G1 cells in red and G2/M cells in green. **(D)** UMAP plots highlighting expression levels of indicated genes. Gray (low) to purple (high) scale indicates average expression signal.

FigS3

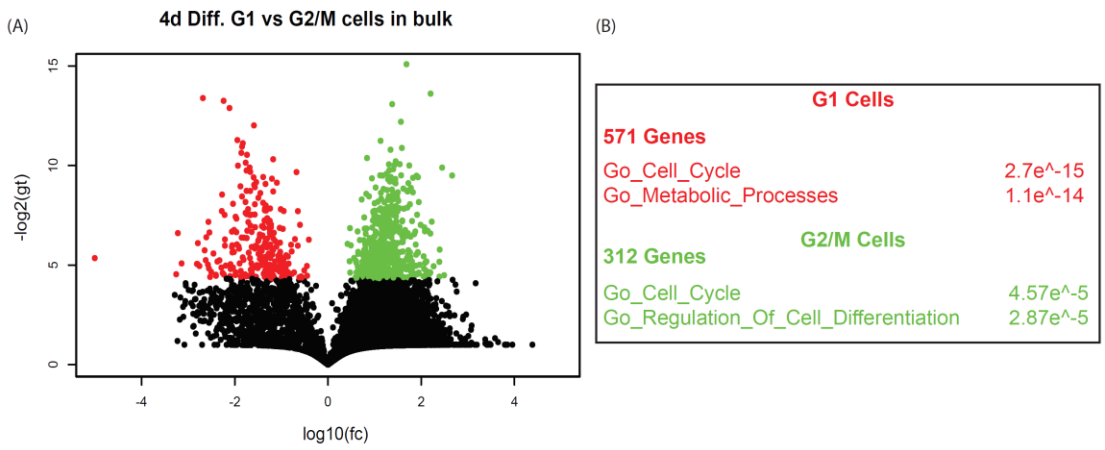

Fig.S3- (A) Volcano plot of bulk RNA-seq data of ESCs in the G1 state vs. ESCs in the G2/M state after four days of differentiation. The fold change in G2/M vs. G1 ESCs is plotted vs. t-test p values (n=3). **(B)** Gene Set Enrichment Analysis of GO terms for significant genes obtained in the volcano plot.

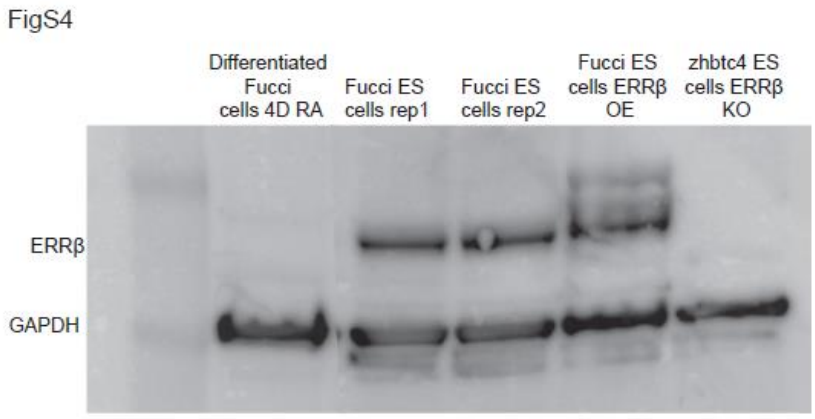

Fig.S4- Western blot of differentiated ESCs after 4 days shows total depletion of ERRβ. Two replicates of WT Fucci ESCs shows basal levels of ERRβ during pluripotency. ERRβ over expression (OE) shows same and heavier isoform of ERRβ. ERRβ Knock out (KO) ESCs shown on Zhbtc4 cell line produced by the Smith lab[51].

References

1. Martin, G. R. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc. Natl. Acad. Sci. U. S. A.* **78**, 7634–7638 (1981).
2. Tam, P. P. . & Behringer, R. R. Mouse gastrulation: the formation of a mammalian body plan. *Mech. Dev.* **68**, 3–25 (1997).
3. O'Shea, K. S. Self-renewal vs. Differentiation of Mouse Embryonic Stem Cells1. *Biol. Reprod.* **71**, 1755–1765 (2004).
4. Graf, T. & Enver, T. Forcing cells to change lineages. *Nature* **462**, 587–594 (2009).
5. Kurimoto, K. *et al.* Quantitative Dynamics of Chromatin Remodeling during Germ Cell Specification from Mouse Embryonic Stem Cells. *Cell Stem Cell* **16**, 517–532 (2015).
6. Dixon, J. R. *et al.* Chromatin architecture reorganization during stem cell differentiation. *Nature* **518**, 331–336 (2015).
7. Pauklin, S. & Vallier, L. XThe cell-cycle state of stem cells determines cell fate propensity. *Cell* **155**, 135 (2013).
8. Lu, Y. C. *et al.* The Molecular Signature of Megakaryocyte-Erythroid Progenitors Reveals a Role for the Cell Cycle in Fate Specification. *Cell Rep.* **25**, 2083-2093.e4 (2018).

9. Nurse, P., Masui, Y. & Hartwell, L. Understanding the cell cycle. *Nature Medicine* **4**, 1103–1106 (1998).
10. Mac Auley, A., Werb, Z. & Mirkes, P. E. Characterization of the unusually rapid cell cycles during rat gastrulation. *Development* **117**, (1993).
11. Salomoni, P. & Calegari, F. Cell cycle control of mammalian neural stem cells: Putting a speed limit on G1. *Trends in Cell Biology* **20**, 233–243 (2010).
12. Stead, E. *et al.* Pluripotent cell division cycles are driven by ectopic Cdk2, cyclin A/E and E2F activities. *Oncogene* **21**, 8320–8333 (2002).
13. ter Huurne, M., Chappell, J., Dalton, S. & Stunnenberg, H. G. Distinct Cell-Cycle Control in Two Different States of Mouse Pluripotency. *Cell Stem Cell* (2017). doi:10.1016/j.stem.2017.09.004
14. Arai, Y. *et al.* Neural stem and progenitor cells shorten S-phase on commitment to neuron production. *Nat. Commun.* **2**, (2011).
15. Boward, B., Wu, T. & Dalton, S. Concise Review: Control of Cell Fate Through Cell Cycle and Pluripotency Networks. *Stem Cells* **34**, 1427–1436 (2016).
16. Soufi, A. & Dalton, S. Cycling through developmental decisions: How cell cycle dynamics control pluripotency, differentiation and reprogramming. *Development (Cambridge)* **143**, 4301–4311 (2016).
17. Chen, X., Hartman, A. & Guo, S. Choosing Cell Fate Through a Dynamic Cell Cycle. *Current Stem Cell Reports* **1**, 129–138 (2015).
18. Dalton, S. XG1 compartmentalization and cell fate coordination. *Cell* **155**, 13 (2013).
19. Hunter, G. L. *et al.* Coordinated control of notch/delta signalling and cell cycle progression drives lateral inhibition-mediated tissue patterning. *Dev.* **143**, 2305–2310 (2016).
20. Singh, A. M. *et al.* Cell-Cycle Control of Bivalent Epigenetic Domains Regulates the Exit from Pluripotency. *Stem Cell Reports* **5**, 323–336 (2015).
21. Bouldin, C. M., Snelson, C. D., Farr, G. H., Kimelman, D. & Kimelman, D. Restricted expression of cdc25a in the tailbud is essential for formation of the zebrafish posterior body. *Genes Dev.* **28**, 384–95 (2014).
22. Niederreither, K. & Dollé, P. Retinoic acid in development: Towards an integrated view. *Nature Reviews Genetics* **9**, 541–553 (2008).
23. Zhang, J. *et al.* Retinoic Acid Induces Embryonic Stem Cell Differentiation by Altering Both Encoding RNA and microRNA Expression. *PLoS One* **10**, e0132566 (2015).
24. Okada, Y., Shimazaki, T., Sobue, G. & Okano, H. Retinoic-acid-concentration-dependent acquisition of neural cell identity during in vitro differentiation of mouse

embryonic stem cells. *Dev. Biol.* **275**, 124–142 (2004).

25. Takahashi, J., Palmer, T. D. & Gage, F. H. Retinoic acid and neurotrophins collaborate to regulate neurogenesis in adult-derived neural stem cell cultures. *J. Neurobiol.* **38**, 65–81 (1999).
26. Janesick, A., Wu, S. C. & Blumberg, B. Retinoic acid signaling and neuronal differentiation. *Cell. Mol. Life Sci.* **72**, 1559–1576 (2015).
27. Borowiak, M. *et al.* Small Molecules Efficiently Direct Endodermal Differentiation of Mouse and Human Embryonic Stem Cells. *Cell Stem Cell* **4**, 348–358 (2009).
28. Simandi, Z., Balint, B. L., Poliska, S., Ruhl, R. & Nagy, L. Activation of retinoic acid receptor signaling coordinates lineage commitment of spontaneously differentiating mouse embryonic stem cells in embryoid bodies. *FEBS Lett.* **584**, 3123–3130 (2010).
29. Cho, L. T. Y. *et al.* Conversion from mouse embryonic to extra-embryonic endoderm stem cells reveals distinct differentiation capacities of pluripotent stem cell states. *Dev.* **139**, 2866–2877 (2012).
30. McDonald, A. C. H., Biechele, S., Rossant, J. & Stanford, W. L. Sox17-mediated XEN cell conversion identifies dynamic networks controlling cell-fate decisions in embryo-derived stem cells. *Cell Rep.* **9**, 780–793 (2014).
31. Young, R. A. Control of Embryonic Stem Cell State. (2011). doi:10.1016/j.cell.2011.01.032
32. Okamura, E. *et al.* Esrrb function is required for proper primordial germ cell development in presomite stage mouse embryos. *Dev. Biol.* **455**, 382–392 (2019).
33. Percharde, M. *et al.* Ncoa3 functions as an essential Esrrb coactivator to sustain embryonic stem cell self-renewal and reprogramming. *Genes Dev.* **26**, 2286–2298 (2012).
34. Adachi, K. & Niwa, H. A liaison between intrinsic and extrinsic regulators of pluripotency. *EMBO J.* **32**, 2531–2532 (2013).
35. Pfeuty, B., Kress, C. & Pain, B. Network Features and Dynamical Landscape of Naive and Primed Pluripotency. *Biophys. J.* **114**, 237–248 (2018).
36. Festuccia, N., Owens, N. & Navarro, P. Esrrb, an estrogen-related receptor involved in early development, pluripotency, and reprogramming. *FEBS Lett.* **592**, 852–877 (2018).
37. Sakaue-Sawano, A. *et al.* Visualizing Spatiotemporal Dynamics of Multicellular Cell-Cycle Progression. *Cell* (2008). doi:10.1016/j.cell.2007.12.033
38. Chen, X. *et al.* Directed neuronal differentiation of mouse embryonic and induced pluripotent stem cells and their gene expression profiles. *Int. J. Mol. Med.* **32**, 25–34 (2013).

39. Coronado, D. *et al.* A short G1 phase is an intrinsic determinant of naïve embryonic stem cell pluripotency. *Stem Cell Res.* **10**, 118–131 (2013).
40. Zhang, X. *et al.* Comparative Analysis of Droplet-Based Ultra-High-Throughput Single-Cell RNA-Seq Systems. *Mol. Cell* **73**, 130-142.e5 (2019).
41. Lugus, J. J. *et al.* GATA2 functions at multiple steps in hemangioblast development and differentiation. *Development* **134**, 393–405 (2007).
42. Gribov, A. *et al.* SEURAT: Visual analytics for the integrated analysis of microarray data. *BMC Med. Genomics* **3**, (2010).
43. Mitsui, K. *et al.* The homeoprotein nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* **113**, 631–642 (2003).
44. Lim, C. Y. *et al.* Sall4 Regulates Distinct Transcription Circuitries in Different Blastocyst-Derived Stem Cell Lineages. *Cell Stem Cell* **3**, 543–554 (2008).
45. Jaber, M., Sebban, S. & Buganim, Y. Acquisition of the pluripotent and trophectoderm states in the embryo and during somatic nuclear reprogramming. *Curr. Opin. Genet. Dev.* **46**, 37–43 (2017).
46. Festuccia, N. *et al.* Mitotic binding of Esrrb marks key regulatory regions of the pluripotency network. *Nat. Cell Biol.* **18**, 1139–1148 (2016).
47. Gao, S. wei & Liu, F. Novel insights into cell cycle regulation of cell fate determination. *J. Zhejiang Univ. Sci. B* **20**, 467–475 (2019).
48. Betschinger, J. *et al.* Exit from pluripotency is gated by intracellular redistribution of the bHLH transcription factor Tfe3. *Cell* **153**, 335–47 (2013).
49. Benchetrit, H. *et al.* Direct Induction of the Three Pre-implantation Blastocyst Cell Types from Fibroblasts. *Cell Stem Cell* **24**, 983-994.e7 (2019).
50. Festuccia, N. *et al.* Mitotic binding of Esrrb marks key regulatory regions of the pluripotency network. *Nat. Cell Biol.* **18**, 1139–1148 (2016).
51. Niwa, H., Miyazaki, J. I. & Smith, A. G. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat. Genet.* **24**, 372–376 (2000).
52. Blasi, T. *et al.* Label-free cell cycle analysis for high-throughput imaging flow cytometry. *Nat. Commun.* **7**, 1–9 (2016).
53. Turkmen, A. S., Yuan, Y. & Billor, N. Evaluation of methods for adjusting population stratification in genome-wide association studies: Standard versus categorical principal component analysis. *Ann. Hum. Genet.* **83**, 454–464 (2019).
54. Waisman, A. *et al.* Cell cycle dynamics of mouse embryonic stem cells in the ground state and during transition to formative pluripotency. *Sci. Rep.* **9**, 8051 (2019).
55. Chung, S. H. *et al.* Differentiation of Retinal Glial Cells From Human Embryonic

Stem Cells by Promoting the Notch Signaling Pathway. *Front. Cell. Neurosci.* **13**, (2019).

56. Zhang, Z. *et al.* Activin a promotes myofibroblast differentiation of endometrial mesenchymal stem cells via STAT3-dependent Smad/CTGF pathway. *Cell Commun. Signal.* **17**, (2019).
57. Kim, J., Chu, J., Shen, X., Wang, J. & Orkin, S. H. An Extended Transcriptional Network for Pluripotency of Embryonic Stem Cells. *Cell* **132**, 1049–1061 (2008).
58. Boyer, L. A. *et al.* Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122**, 947–956 (2005).
59. Gökbuget, D. & Blelloch, R. Epigenetic control of transcriptional regulation in pluripotency and early differentiation. *Dev.* **146**, (2019).
60. Benchetrit, H. *et al.* Direct Induction of the Three Pre-implantation Blastocyst Cell Types from Fibroblasts. *Cell Stem Cell* **24**, 983-994.e7 (2019).
61. Betschinger, J. *et al.* Exit from pluripotency is gated by intracellular redistribution of the bHLH transcription factor Tfe3. *Cell* **153**, 335–347 (2013).
62. Kojima, Y. *et al.* Evolutionarily Distinctive Transcriptional and Signaling Programs Drive Human Germ Cell Lineage Specification from Pluripotent Stem Cells. *Cell Stem Cell* **21**, 517-532.e5 (2017).
63. Guccione, E. *et al.* Myc-binding-site recognition in the human genome is determined by chromatin context. *Nat. Cell Biol.* **8**, 764–770 (2006).
64. Lupien, M. *et al.* FoxA1 Translates Epigenetic Signatures into Enhancer-Driven Lineage-Specific Transcription. *Cell* **132**, 958–970 (2008).
65. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).
66. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, (2009).
67. Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902.e21 (2019).
68. Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15545–15550 (2005).
69. Ray, H. N., Mootha, V. K. & Boxwala, A. A. *Building an Application Framework for Integrative Genomics*.