Article
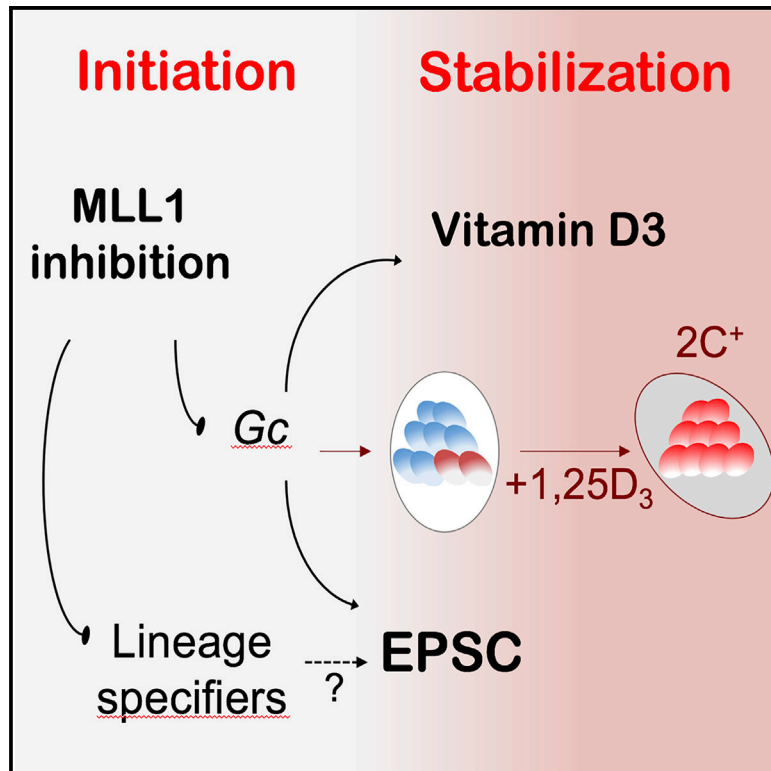
MLL1 Inhibition and Vitamin D Signaling Cooperate to Facilitate the Expanded Pluripotency State

Graphical Abstract

Authors

Hui Zhang, Le Tran Phuc Khoa, Fengbiao Mao, ..., Li Wang, Thomas L. Saunders, Yali Dou

Correspondence

yalid@med.umich.edu

In Brief

Zhang et al. discover that inhibiting histone methyltransferase MLL1 is able to expand the pluripotency potential of naive embryonic stem cells. These findings highlight the function of developmentally regulated histone H3K4 methylation in restricting cell fate plasticity and have direct implication in regenerate medicine.

Highlights

- MLL1 inhibition reprograms naive stem cells to expanded pluripotent stem cells

- MLL1 inhibition or deletion upregulates early blastomere gene signatures

- Gc depletion partially mimics MLL1 inhibition in promoting EPSCs

- MLL1i and vitamin D treatment cooperatively enhance functionality of EPSCs

CellPress

MLL1 Inhibition and Vitamin D Signaling Cooperate to Facilitate the Expanded Pluripotency State

Hui Zhang,[1,2] Le Tran Phuc Khoa,[2] Fengbiao Mao,[2] Hanshi Xu,[1] Bo Zhou,[2] Yu Han,[1] Monique O'Leary,[2] Asma Nusrat,[2] Li Wang,[1] Thomas L. Saunders,[3] and Yali Dou[2,4,*]

[1]State Key Laboratory of Cardiovascular Disease, Fuwai Hospital, National Center for Cardiovascular Diseases, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100037, China
[2]Department of Pathology, University of Michigan, Ann Arbor, MI 48109, USA
[3]Department of Internal Medicine, University of Michigan, Ann Arbor, MI 48109, USA
[4]Lead Contact
*Correspondence: yalid@med.umich.edu
https://doi.org/10.1016/j.celrep.2019.10.074

SUMMARY

Dynamic establishment of histone modifications in early development coincides with programed cell fate restriction and loss of totipotency beyond the early blastocyst stage. Causal function of histone-modifying enzymes in this process remains to be defined. Here we show that inhibiting histone methyltransferase MLL1 reprograms naive embryonic stem cells (ESCs) to expanded pluripotent stem cells (EPSCs), with differentiation potential toward both embryonic and extraembryonic lineages *in vitro* and *in vivo*. MLL1 inhibition or deletion upregulates gene signatures of early blastomere development. The function of MLL1 in restricting induction of EPSCs is mediated partly by *Gc*, which regulates cellular response to vitamin D signaling. Combined treatment of MLL1 inhibitor and 1α,25-dihydroxyvitamin D$_3$ (1,25-(OH)$_2$D$_3$) cooperatively enhanced functionality of EPSCs, triggering an extended 2C-like state *in vitro* and robust totipotent-like property *in vivo*. Our study sheds light on interplay between epigenetics and vitamin D pathway in cell fate determination.

INTRODUCTION

Totipotency is a transient property of cells in the early embryo. Cells in the mouse zygote or 2-cell (2C) stage are totipotent, capable of giving rise to the whole organism (Baker and Pera, 2018; Tarkowski, 1959). This ability is lost when embryonic development goes beyond the early blastocyst stage (Papaioannou et al., 1989; Tarkowski, 1959; Wu et al., 2017). Cells isolated from the inner cell mass (ICM) of the blastocyst are pluripotent, losing the ability to contribute to extraembryonic tissues (e.g., placenta and yolk sac) in later development (Beddington et al., 1989; Evans and Kaufman, 1981). Several studies propose that the totipotent-like state occurs spontaneously in the *in vitro* culture

of pluripotent stem cells (Macfarlan et al., 2012), indicated by elevated expression of 2C-specific MERVL (Evsikov et al., 2004; Macfarlan et al., 2012), as well as other early cleavage gene markers such as *Zscan4* and *Dux* (De Iaco et al., 2017; Evsikov et al., 2004; Falco et al., 2007; Hendrickson et al., 2017). Embryonic stem cells (ESCs) with deletion of miR-34a (Choi et al., 2017) or treated by a chemical cocktail of signaling pathway inhibitors (Yang et al., 2017a, 2017b) are able to acquire expanded pluripotency, contributing to three embryonic germ layers and, with less efficiency, to extraembryonic tissues (Abad et al., 2013; Choi et al., 2017; Macfarlan et al., 2012; Morgani et al., 2013; Yang et al., 2017a, 2017b). Molecular characterization of emergence of totipotent-like or 2C-like cells at single-cell resolution suggests that reprogramming of ESCs to 2C-like cells is a multistep process, which is likely regulated by chromatin modifiers (e.g., CAF1, PRC1, and EP400-TIP60) (Ishiuchi et al., 2015; Rodriguez-Terrones et al., 2018). However, it remains to be tested whether chromatin modification enzymes play a causal role in promoting expanded pluripotent stem cells (EPSCs) and what the underlying mechanism is for cell fate restriction in early development.

Among post-translational histone modifications (PTMs), histone H3 lysine 4 tri-methylation (H3K4me3) marks broad chromatin domains during transcription activation of the zygotic gene program (Dahl et al., 2016; Eckersley-Maslin et al., 2018; Liu et al., 2016; Zhang et al., 2016a). It is increasingly restricted to the gene promoters in early cleavage stage embryos, coinciding with gradual reduction of cell fate plasticity (Dahl et al., 2016; Eckersley-Maslin et al., 2018; Liu et al., 2016; Zhang et al., 2016a). While little is known about dynamics of H3K4 mono-methylation (H3K4me1) and H3K4 di-methylation (H3K4me2), the main histone modifications at distal regulatory enhancers, during early embryogenesis, several studies report dynamic establishment and decommissioning of H3K4me1/2 in cell fate transition, including reprogramming of epiblast stem cells (EpiSCs) and somatic cells (Buecker et al., 2014; Factor et al., 2014; Gifford et al., 2013). Using the MLL1 inhibitor (MM-401) we recently developed (Cao et al., 2014; Karatas et al., 2013), we showed that inhibiting H3K4 methyltransferase MLL1 (also called KMT2A and MLL) makes it possible to reprogram EpiSCs to the naive pluripotent state (Zhang et al.,
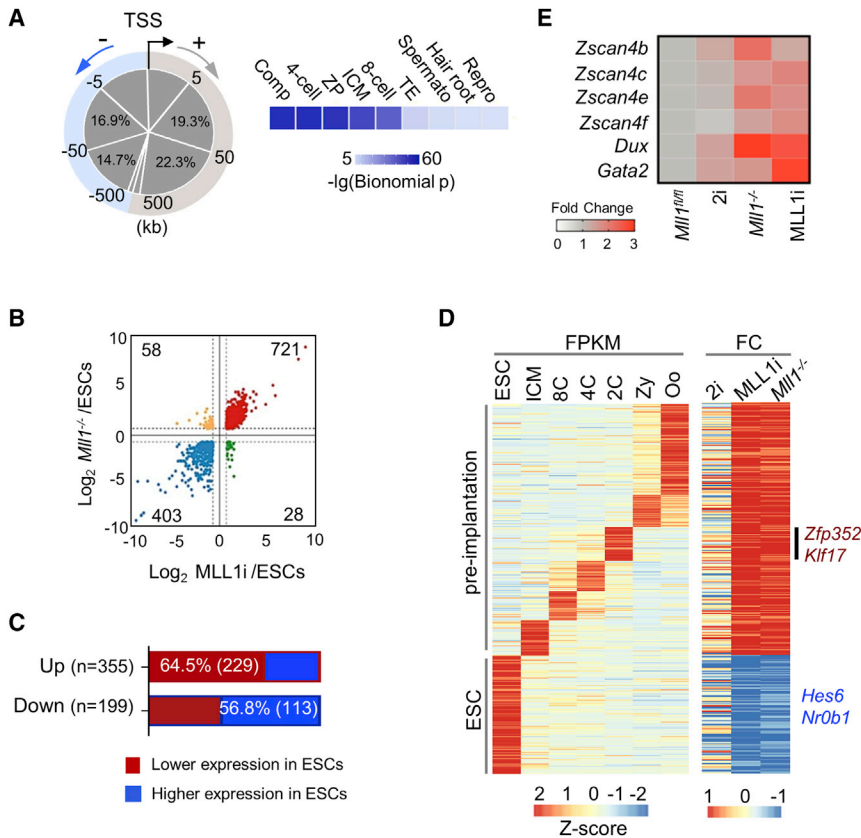
Figure 1. MLL1i Promotes Expression of Early Cleavage Stage Gene Signatures

(A) Left, genome-wide distribution of MLL1 relative to transcription start site (TSS). Right, G.R.E.A.T. pathway enrichment for MLL1 bound genes using −log (binomial p).

(B) Differential gene expression of MLL1i versus mock (x axis) or *Mll1*KO (y axis) versus *Mll1*[f/f] ESCs. Genes with FDR ≤ 0.05 and |FC| ≥ 1.5 were selected. The results were from biological duplicates.

(C) Genes that were up- or downregulated by MLL1i and *Mll1*KO were evaluated for expression dynamics in early embryonic development (Wu et al., 2016). Red, higher expression in pre-blastocyst embryos; blue, higher expression in ESCs.

(D) Heatmap (log₂ fold change) for gene expression changes. Right, gene expression change induced by treatment as indicated on top. Input were genes that were upregulated (67.7%) or downregulated (56.8%) by MLL1i and *Mll1*KO in (C). Heatmap was generated using the defined range from −1 to 1. Left, heatmap (Z score) for the same set of genes on the right at different developmental stages indicated on top. The *in vivo* gene expression dataset was from Wu et al., (2016), Database: GEO: GSE66582.

(E) Heatmap for expression of selective early cleavage stage markers, indicated on the left, in mock, 2i, *Mll1*KO, or MLL1i cells.

See also Figures S1 and S2 and Tables S1 and S2.

2016b). It highlights a causal function of MLL1 and MLL1-mediated H3K4 methylation (H3K4me) in EpiSC reprogramming (Zhang et al., 2016b). Given continuous increase of H3K4me in preimplantation embryos, we raise the question of whether MLL1 and H3K4me play any role in restricting cell fate commitment toward naive pluripotency and whether inhibiting MLL1 activity erases such epigenetic restriction to promote totipotency.

Here we report that MLL1 plays an essential role in cell fate restriction in pluripotent stem cells. MLL1 inhibition increases cell fate plasticity and promotes an expanded pluripotent state that contributes to development of extraembryonic lineages both *in vitro* and *in vivo*. We further find that the function of MLL1 is mediated partly by *Gc* and vitamin D signaling. Combined MLL1 inhibition and vitamin D treatment leads to enhanced EPSC functionality both *in vitro* and *in vivo*.
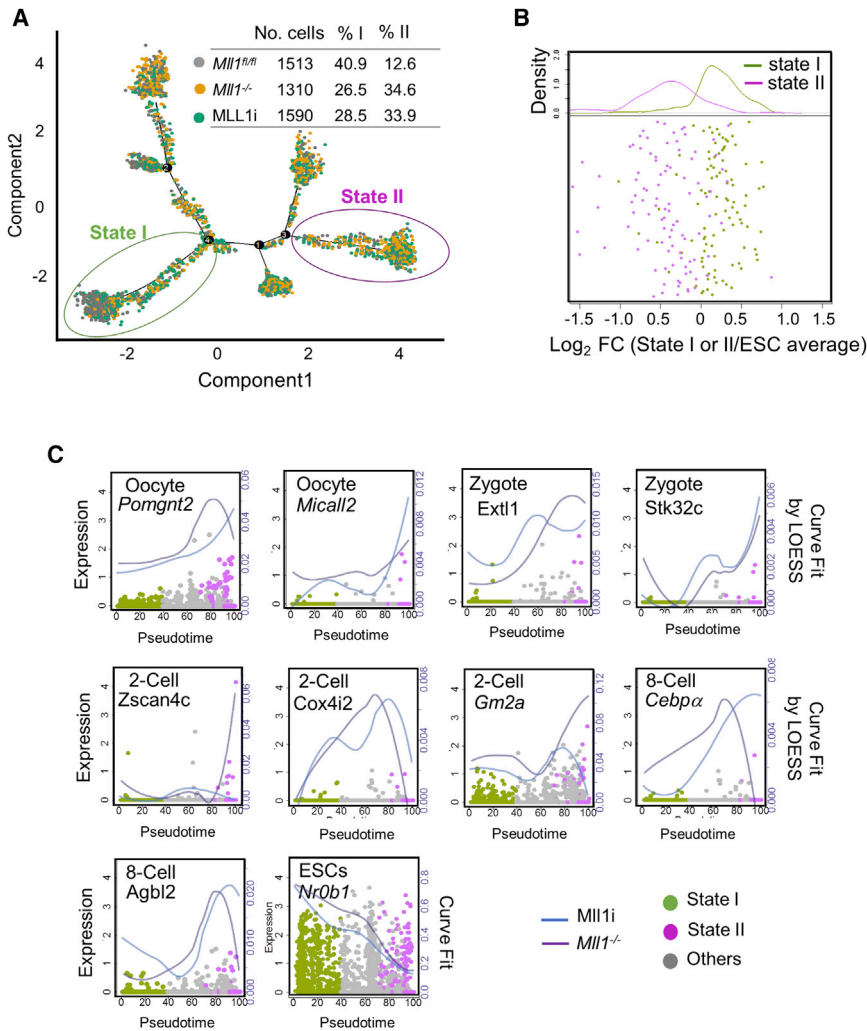
RESULTS

MLL1 Inhibition Reprograms the ESC Transcriptome to Mimic Early Cleavage State Embryos

Our previous study showed that inhibiting histone H3K4 methyltransferase MLL1 by small molecule inhibitor MM-401 reprograms EpiSCs to a stable naive pluripotent state (Zhang et al., 2016b). However, a closer examination of the transcriptome changes showed elevated expression of several cleavage-specific genes, such as Zscan4, *Dux*, and *Gata2*, in the reverted ESCs (Figure S1A). Expression of these genes was higher in cells treated with MM-401

for 30 passages (P30) than in P6 cells. Unlike ESC core transcription factors (e.g., *Pou5f1*) (Zhang et al., 2016b), expression of the cleavage state-specific genes required continuous MM-401 treatment (Figure S1A). Their levels were significantly downregulated at 30 passages (P30) when MM-401 was withdrawn at P6. This result suggests that MLL1 inhibition (MLL1i) can probably promote a transient cellular state beyond naive pluripotency.

To examine this possibility, we treated *Mll1*[fl/fl]*; Cre-ER*[TM] ESCs with either 4-hydroxytamoxifen (4-OHT) for *Mll1* deletion (*Mll1*KO) or MM-401 for MLL1i (Figure S1B). ESCs with *Mll1*KO or MLL1i formed small colonies with strong alkaline phosphatase (AKP) staining (Figures S1C and S1D), similar to 2i (MEK and GSK3 inhibitors)-treated ground-state ESCs (Nichols et al., 2009; Ying et al., 2008). Chromatin immunoprecipitation sequencing (ChIP-seq) analysis showed that most MLL1 bound to promoter distal regions in ESCs (Figure 1A), including intergenic (38%) and intron (42%) regions, and colocalized with H3K4me1 and H3K4me2 peaks (Figure S2A; Table S1). Functional prediction by G.R.E.A.T. (Genomic Regions Enrichment of Annotations Tool) showed enrichment of MLL1 targets in early blastomere development (e.g., 4 cell and 8 cell) (Figure 1A, right). Mouse genome informatics (MGI) analysis also identified pathways such as trophectoderm morphology, as well as organization of extraembryonic tissue for the MLL1 direct targets (Figure S2B), in addition to previously reported MLL1 pathways such as blastocyst hatching, hematopoietic stem cells, and cell adhesion (Figure S2B) (Jude et al., 2007; Zhang et al., 2016b).

A

	No. cells	% I	% II
Mll1[fl/fl]	1513	40.9	12.6
Mll1[-/-]	1310	26.5	34.6
MLL1i	1590	28.5	33.9

State I, State II

B

C

Oocyte *Pomgnt2*; Oocyte *Micall2*; Zygote Extl1; Zygote Stk32c; 2-Cell Zscan4c; 2-Cell Cox4i2; 2-Cell *Gm2a*; 8-Cell *Cebpα*; 8-Cell Agbl2; ESCs *Nr0b1*

Mll1i, Mll1[-/-], State I, State II, Others

Figure 2. Single-Cell RNA-Seq Analyses for *Mll1*[fl/fl], *Mll1*KO, and MLL1i ESCs

(A) Monocle analysis on cell trajectory in *Mll1*[fl/fl], *Mll1*KO, and MLL1i cells. Cell numbers used for the analysis, as well as their distribution in the two most abundant states (I and II), were summarized.

(B) Relative expression of MLL1 direct targets in state I (green) and state II (fuchsia) compared with their average expression in ESCs. Top, histogram of the expression spread in states I and II. Bottom, $\log_2$ fold change of gene expression. Each dot represents an individual gene that was used to define states I and II in (A).

(C) Expression of specific genes in the linear trajectory of cells in state I (green), state II (fuchsia), and other states (gray). Each dot represents a cell. The x axis is pseudo-time with state I as the root state. The left y axis is gene expression. The right y axis is weighted gene expression after curve fit by LOESS (locally weighted scatterplot smoothing) regression. The lines were the LOESS regression pattern in pseudo-time with *Mll1*[-/-] cells (purple) or MLL1i (blue) as independent variables and gene expression as the dependent variable.

We further examined gene expression changes in ESCs upon *Mll1*KO or MLL1i by RNA sequencing (RNA-seq) analysis. There were 1,124 genes (fold change $|FC| \geq 1.5$, false discovery rate [FDR] ≤ 0.05) consistently upregulated (721) or downregulated (403) by *Mll1*KO and MLL1i (Figure 1B; Table S2). About half of these genes (554 of 1,124) exhibited dynamic changes in early embryonic development *in vivo* (Wu et al., 2016). Among the developmentally regulated genes, 355 were upregulated by *Mll1*KO and MLL1i (Figure 1C). A heatmap showed that 64.5% of these genes (229) were highly expressed in early cleavage stage embryos from oocyte to ICM *in vivo* (Figure 1D) (Fujii et al., 2015), including previously reported cleavage markers *Zscan4*, *Gata2*, *Dux*, and MERVL (Figures 1E and S2C) (Evsikov et al., 2004; Falco et al., 2007; Hendrickson et al., 2017; Macfarlan et al., 2012). Among 403 genes downregulated by *Mll1*KO and MLL1i (Figure 1C), 56.8% were typically expressed at high levels in ESCs (e.g., *Sall1* and *Nr0b1*) (Figure 1D). Notably, the transcriptome change induced by *Mll1*KO and MLL1i was largely distinct from that of 2i (Figures 1D and 1E). Unbiased global gene clustering analyses (Short Time-series Expression Miner [STEM]) confirmed that unlike 2i treatment, *Mll1*KO or MLL1i significantly downregu-

lated genes that are continuously upregulated in preimplantation embryos ($|FC| \geq 1.5$, FDR ≤ 0.05, cluster I) (Figure S2D) and significantly upregulated genes that are either continuously downregulated during development (cluster II) or transiently upregulated at the 2C stage before developmentally programmed downregulation (cluster III) (Figure S2D).

Single-cell RNA-seq analysis illustrated heterogeneity of ESCs in culture (Figure 2A), with ~40% of *Mll1*[fl/fl] ESCs in state I defined by the monocle pseudo-time trajectory analysis. Upon *Mll1*KO or MLL1i, there was a reduction of cells in state I and a corresponding increase of cells in state II in culture (Figure 2A). Comparing expression of genes that defined state I versus state II in the monocle analysis, genes (such as Nr0b1) in state I had average or above-average expression in *Mll1*[fl/fl] ESCs (Figures 2B and 2C). In contrast, state II genes, which included multiple early cleavage stage markers (Figure 2C), were expressed at lower-than-average levels (Figure 2B). Representative distribution of selected cleavage state and embryonic stem cell ESC state genes are shown in Figure 2C. These results show that there is a cell state transition in a subpopulation of *Mll1*[fl/fl] ESCs upon *Mll1*KO or MLL1i. The transcription alteration is indicative of a shift toward early blastomere development. State II is still heterogenous, with expression of multiple cleavage stage markers (Figure 2C).

MLL1 Inhibition Promotes a Transient 2C-Reporter+ State *In Vitro*

To functionally determine whether MLL1i and *Mll1*KO cells led to an expanded pluripotency state *in vitro*, we first used a previously established 2C-reporter assay (Macfarlan et al., 2012;
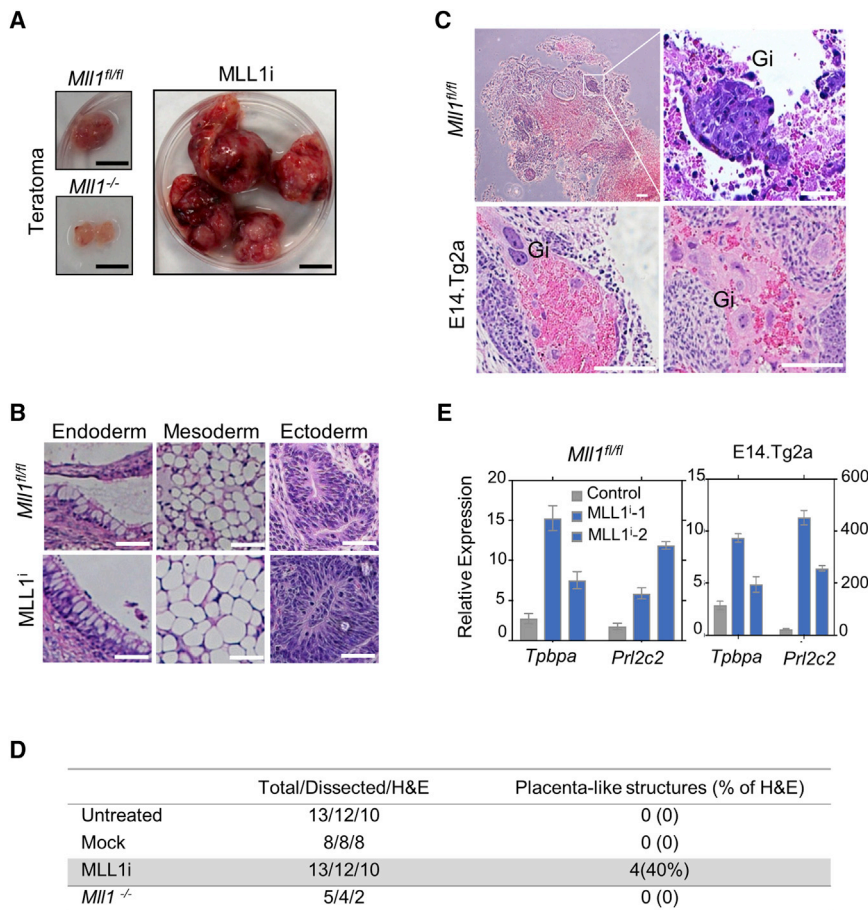
Figure 3. MLL1i Expands Pluripotency Potential *In Vitro* and *In Vivo*

(A) Teratoma generated from engrafted cells (indicated on top) by subcutaneous injection. Scale bar, 1 cm. The experiments were performed for 6 biological duplicates from two independent experiments.

(B) H&E staining for three germ layers in teratomas derived from *Mll1^{f/f}*; ER-Cre (top) or MLL1i ESC (bottom). Scale bar, 50 μm.

(C) Placenta-characteristic hemorrhage and TGCs, indicated by Gi, in MLL1i or E14Tg2a ESC-derived teratoma. Scale bar, 100 μm.

(D) Summary of the teratoma experiments.

(E) qRT-PCR for placenta gene markers *Tpbpa* and *Prl2c2*. Average gene expression is presented after normalization against the gene expression in ESCs. Error bars are SD from triplicate samples.

See also Figures S3–S5.

	Total/Dissected/H&E	Placenta-like structures (% of H&E)
Untreated	13/12/10	0 (0)
Mock	8/8/8	0 (0)
MLL1i	13/12/10	4(40%)
Mll1 ^{-/-}	5/4/2	0 (0)

(Figures S4B and S4C). Constitutive *Mll1*KO was also able to initiate commitment toward the iTSC fate (Figures S4B and S4C). However, the induced iTSCs could not propagate in subsequent passages (Figures S4B and S4C). This is consistent with a recent report that MLL1 is required for Zfp281-mediated transcription of TSC genes (Ishiuchi et al., 2019). Indeed, TSC growth defects could be rescued by expression of exogenous MLL1 in *Mll1*KO cells (Figure S4D). As the control, ground-state ESCs under the 2i condition were unable to differentiate into iTSCs under the same condition (Figure S4A; data not shown).

Rodriguez-Terrones et al., 2018) to test whether there is an increase in 2C-like cells after MLL1i or *Mll1*KO. We introduced the MERVL-driven 2C reporter into *Mll1^{f/f}*; *Cre-ER^{TM}* ESCs. Upon *Mll1*KO or MLL1i, there was significant increase of spontaneous 2C$^+$ cells, from 0.34% in ESCs to 4.45% or 3.45%, respectively (Figure S3A; Videos S1, S2, and S3). However, the induced 2C$^+$ state remained transient. Only 21%–26% of sorted 2C$^+$ cells maintained 2C-reporter expression at P3, which further decreased to 3%–4% at P5 (Figure S3B).

MLL1 Inhibition Expands Extraembryonic Cell Fate Potential *In Vitro*

A distinguishing feature of EPSCs or totipotent-like cells is the ability to differentiate into extraembryonic lineages such as trophoblast stem cells (TSCs), derived from extraembryonic trophectoderm (TE) (Baker and Pera, 2018; Tanaka et al., 1998). To test whether MLL1i expands stem cell potential toward TSC fate *in vitro*, we treated the *Mll1^{f/f}*; *Cre-ER^{TM}* ESCs with either 4-OHT or MM-401 for five days, followed by FGF4/heparin induction (Figure S4A). About 20%–30% of MLL1i-treated ESCs were able to differentiate into induced trophoblast stem cell-like cells (iTSCs) in two passages (P2), exhibiting typical TSC morphologies (Figure S4A), as well as high expression of TE markers CDX2 (Benchetrit et al., 2015; Kubaczka et al., 2015; Niwa et al., 2005) and CD40 (Kubaczka et al., 2015; Rugg-Gunn et al., 2012)

MLL1 Inhibition Expands ESC Pluripotency Potential in Teratoma *In Vivo*

To examine contributions of MLL1i and *Mll1*KO ESCs to extraembryonic tissues *in vivo*, we engrafted the *Mll1^{fl/fl}*; *ER-Cre* ESCs that were treated for MM-401 or 4-OHT into recipient mice for teratoma development (Figure 3A). Consistent with previous studies (Katada and Sassone-Corsi, 2010; Lim et al., 2009; McMahon et al., 2007), irreversible *Mll1*KO resulted in poorly developed teratomas with differentiation block (Figure 3A). In contrast, MLL1i ESCs, upon MM-401 removal at the time of engraftment, led to well-differentiated teratomas with three distinct embryonic germ layers (Figure 3B). Furthermore, we observed typical placenta-like histological features such as trophoblast giant cells (TGCs), as well as hemorrhage (Abad et al., 2013; Choi et al., 2017), in 40% of teratomas derived from the MLL1i ESCs (Figures 3C and 3D). These features were not observed in teratomas derived from mock or *Mll1*KO ESCs (Figure 3D). Consistently, significant upregulation of *Tpbpa* and *Prl2c2*, two markers of terminally differentiated TGCs (Kubaczka et al., 2015), were detected in MLL1i-ESC-derived teratomas by quantitative real-time PCR(qRT-PCR)
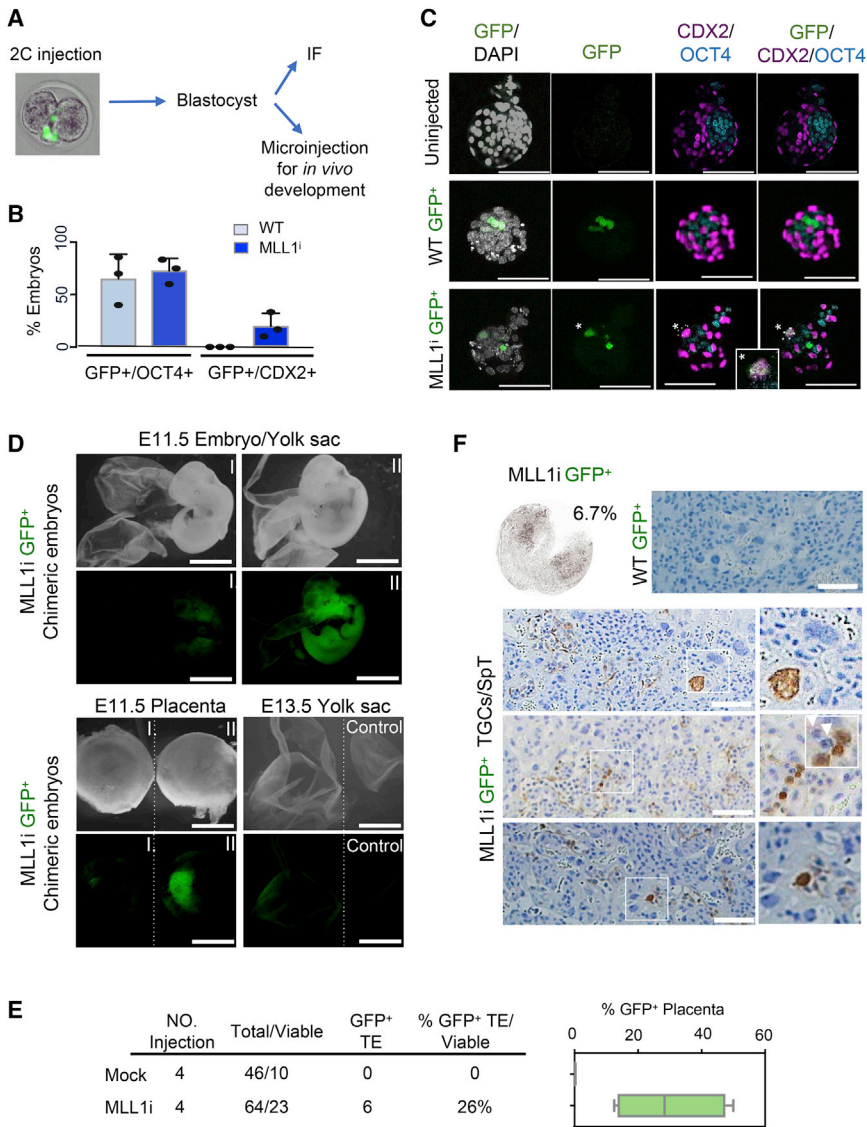
Figure 4. MLL1i ESCs Contribution to Extra-embryonic Lineages Both *In Vitro* and *In Vivo*

(A) Schematic for testing EPSC functions by blastocyst injection.

(B) Summary of the *in vitro* chimeric blastocyst experiments. Average percentage of embryos for dual fluorescent signals is presented. Error bars represent SD from three independent experiments.

(C) Representative images of embryos stained with OCT3/4 (blue) or CDX2 (magenta) that were countered with DAPI. * indicates a GFP⁺/CDX2⁺ cell. Scale bar, 100 μm.

(D) Representative images of placenta and yolk sac from 11.5- or 13.5-day chimeric embryos derived from MLL1i GFP⁺ blastocysts, as indicated on top. Scale bar, 0.25 cm.

(E) Summary of extraembryonic contribution *in vivo*. Average percentage of GFP+ cells in placenta is presented. Range and SD from four independent microinjection experiments are shown in the boxplot on the right.

(F) Immunohistochemistry for GFP in a representative E11.5 placenta of the MLL1i ESC chimera. The percentage of GFP signals was calculated using ImageJ. Immunohistochemistry sections showed trophoblast giant cells (TGCs) and spongiotrophoblasts (Spt) with GFP⁺ signals (white square), which were enlarged on the right. White arrows indcated chimeric GFP⁻ cells. Scale bars, 100 μm.

See also Figures S5 and S6.

cence showed that ∼80%–90% of viable chimeric blastocysts (30 total) had GFP⁺/OCT4⁺ staining, indicating successful integration into EPI (Figure 4B). About 10%–20% of viable chimeric blastocysts had GFP⁺/CDX2⁺ cells (Figure 4B), indicating successful integration into the TE layer. GFP⁺/CDX2⁺ cells were not found in any chimeric blastocyst injected with mock-treated ESCs (Figure 4B). Representative images for the GFP⁺/OCT4⁺ and GFP⁺/CDX2⁺ cells are shown in Figure 4C.

When chimeric blastocysts were injected back into pseudopregnant surrogates for further embryonic development, we obtained 10 and 23 viable embryos (day 11.5 to day 13.5) from four independent microinjection experiments for mock and MLL1i-treated ESCs, respectively. All viable post-gastrulation mouse embryos had GFP⁺ signals in the embryo proper, as expected (Figures 4D and S6A). GFP⁺ cells were also found in extraembryonic tissues, i.e., yolk sac and placenta, of 6 (26% of 23 viable) chimeric embryos microinjected with MLL1i ESCs (Figures 4D and 4E). GFP⁺ cells in TE-derived placenta tissues were confirmed by immunohistochemistry using anti-GFP antibody. It showed GFP⁺ TGCs and spongiotrophoblast (SpT) in the junctional zone of the placenta (Figure 4F). GFP⁺ cells were also found in the labyrinth and placenta septum layer (Figure S6B). No GFP⁺ signals were found in placenta and yolk sac
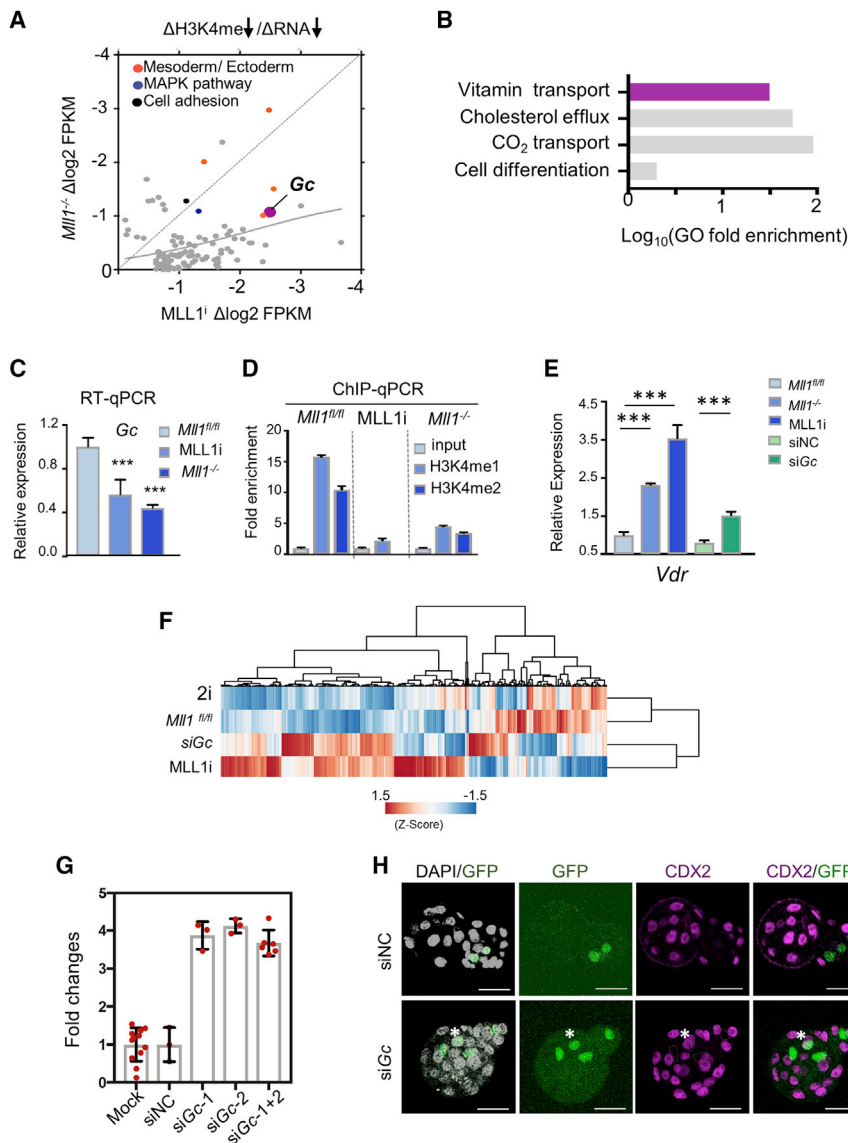
(Figure 3E). Placenta-like characteristics were also observed in teratomas derived from the independent cell line E14Tg2a after MM-401 treatment (Figures 3C and 3E).

MLL1i ESCs Contribute to Extraembryonic Tissues *Ex Vivo* and *In Vivo*

To examine the potential of MLL1i ESCs in extraembryonic development *in vivo*, we microinjected MLL1i- or mock-treated GFP⁺ Nagy R1 cells (Figure S5A) into 2C embryos from C57BL/6N mice and examined chimeric blastocyst development *in vitro* and *in vivo* (Figure 4A). MLL1i ESCs (GFP⁺) colonized both epiblast (EPI) and TE in the *in vitro* cultured chimeric blastocysts. MLL1i GFP⁺ ESCs had strong tendency to colonize the peripheral of a subset of chimeric blastocysts (∼20%), where TE is located (Figures S5B and S5C). Similar peripheral localization was observed for ∼25% of chimeric blastocysts injected with MLLi1-treated E14Tg2a (Tomato⁺) ESCs (Figures S5D–S5F). Immunofluores-

Figure 5. _Gc_ Depletion Recapitulates MLL1i Function in EPSCs

(A) Expression of 131 MLL1 direct targets upon MLL1i (x axis) and _Mll1_KO (y axis). These genes also had concomitant reduction of H3K4me1 (not shown). Representative genes are highlighted. Gray line, nonlinear correlation of RNA-seq signals in MLL1i and _Mll1_$^{-/-}$ cells.

(B) GO term enrichment for downregulated genes upon MLL1i and _Mll1_KO.

(C) qRT-PCR validation for _Gc_ expression as indicated. Average _Gc_ expression from biological triplicates is presented after normalization against its expression in _Mll1_$^{f/f}$ cells, which was arbitrarily set as 1. ***p < 0.001, Mann-Whitney test (unpaired, nonparametric, two tailed).

(D) Fold enrichment of H3K4me1 and H3K4me2 at _Gc_ by ChIP-qPCR.

(E) qRT-PCR for _Vdr_ expression in cells as indicated. Average _Vdr_ expression from triplicate experiments is presented after normalization against its expression in _Mll1_$^{f/f}$ cells, which was arbitrarily set as 1. For (C)–(E), error bars are SD from biological triplicates. ***p < 0.001, Mann-Whitney test (unpaired, nonparametric, two tailed).

(F) Left, principal-component analysis (PCA) plot for expression of MLL1 targets in different cells. Right, clustering analysis for gene expression (same input as PCA) in cells indicated on the left.

(G) Induction of spontaneous 2C$^+$ cells in ESC culture upon _siGc_. Two small interfering RNAs (siRNAs), as well as their equal molar mixture, were used to deplete _Gc_. Average fold increase relative to mock-treated ESCs is presented. Error bars are SD from three independent experiments. For mock, error bar is SD from nine independent experiments.

(H) Immunofluorescence for CDX2 in chimeric blastocyst-derived _siGc_ or siControl GFP$^+$ Nagy R1 ESCs. Scale bars, 50 μm. * indicates GFP$^+$/CDX2$^+$ cells.

See also Figure S7 and Table S3.

of chimeric embryos microinjected with mock-treated ESCs (Figures 4D, 4E, S6A, and S6B), excluding the possibility that embryonic cell lineages, such as extraembryonic mesoderm (e.g., vascular endothelium and blood cells), contributed to extraembryonic tissues. Representative ImageJ and flow cytometry analyses of GFP$^+$ placenta and yolk sac are shown in Figures S6C and S6D. The relatively low contribution of GFP$^+$ to placenta _in vivo_ could be due to instability of the TE cell fate _in vivo_. Nonetheless, our results strongly argue that MLL1i ESCs are bona fide EPSCs, contributing to both embryonic and extraembryonic lineages in chimeric blastocysts _in vivo_.

MLL1i Promotes EPSCs via the Group-Specific Component (Gc) Gene

To determine the mechanism by which MLL1i promotes EPSCs, we examined 111 MLL1 direct targets that had consistent gene expression and H3K4me downregulation upon MLL1i and

_Mll1_KO (Figure 5A; Table S3). The vitamin transport pathway was identified as one of the top Gene Ontology (GO) terms (Figure 5B). The vitamin D binding protein (VDB), encoded by the _Gc_ gene, was one of the top-ranked MLL1 direct targets (Figures 5A and S7A). qRT-PCR and ChIP-qPCR assays validated RNA-seq and ChIP-seq results, respectively, showing reduced expression and H3K4me1/2 after _Mll1_KO or MLL1i (Figures 5C and 5D). Immunofluorescence using anti-VDB antibody confirmed a lower VDB level in the _Mll1_KO and MLL1i ESCs (Figure S7B). Previous studies showed that VDB (_Gc_) functions as a vitamin D sink (Bikle, 2011; Chun et al., 2014) and its depletion sensitizes cellular response to 1α,25-dihydroxyvitamin D$_3$ (1,25-(OH)$_2$D$_3$, or calcitriol) (Zella et al., 2008). _Mll1_KO, MLL1i, and _Gc_ depletion increased both _Vdr_ expression (Figure 5E) and VDR nuclear localization (Figure S7C). Modest change of another vitamin D pathway gene, _Cyp24a1_ (Li et al., 1997;
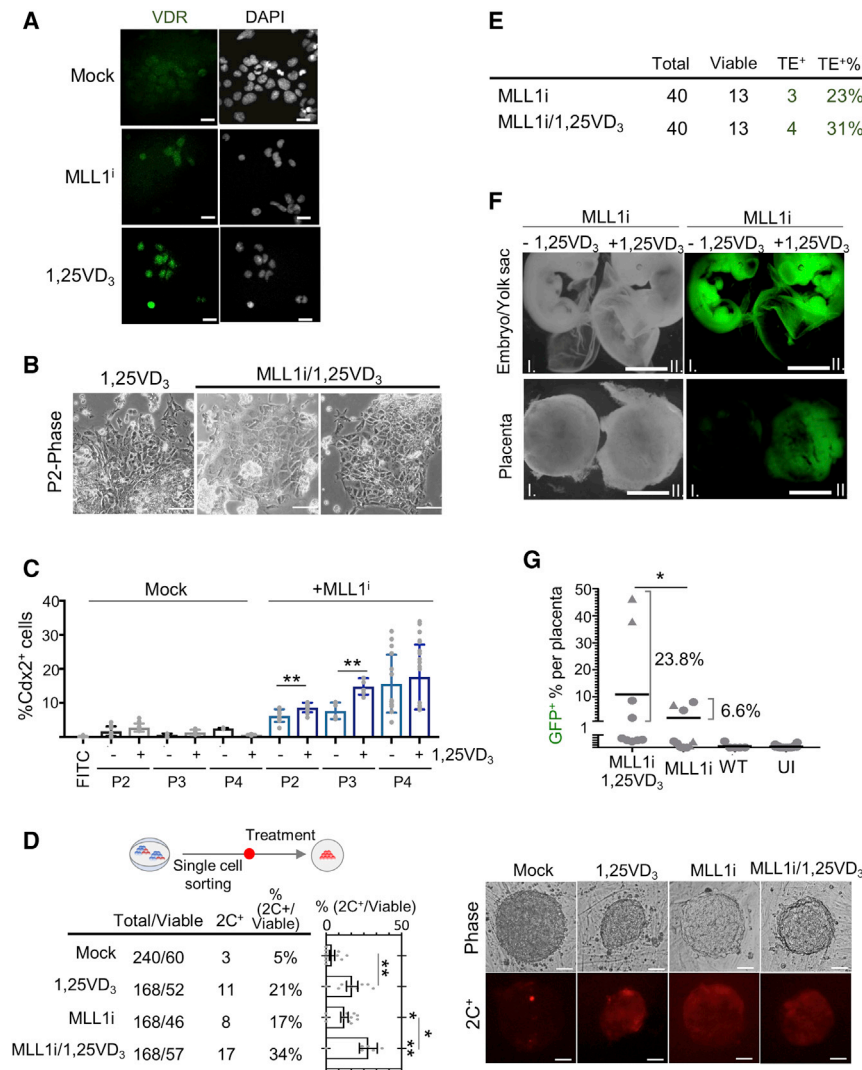
Figure 6. MLL1i and 1,25-(OH)₂D₃ Cooperate to Promote EPSC Functionality

(A) Immunofluorescence for VDR in mock-, 1,25-(OH)₂D₃ (1,25VD3)-, or MLL1i-treated cells. Scale bar, 20 μm.

(B) Representative images for P2 cells after TSC induction. Cells were treated with either 1,25-(OH)₂D₃ or 1,25-(OH)₂D₃/MLL1i, as indicated on top, before Heparin/Fgf4 induction. Scale bar, 100 μm.

(C) Flow cytometry analysis for CDX2⁺ at different passages (P2–P4) after TSC induction. Average percentage of CDX2⁺ cells is presented. Error bars are SD from biological triplicate. **p < 0.01, Mann-Whitney test (unpaired, nonparametric, two tailed).

(D) 2C⁺ maintenance experiment. Spontaneous 2C⁺ cells were sorted and seeded as single cells in a 96-well plate. After 7 days, 2C⁺ colonies were counted and summarized in the table on the left. *p < 0.05, **p < 0.01, Mann-Whitney test (two tailed). Representative images of P1 colonies are shown on the right. Scale bar, 200 μm.

(E) Summary of extraembryonic contribution by MLL1i or MLL1i/1,25VD3 GFP⁺ cells in vivo. The percentage of TE contribution refers to the percentage of viable conceptus. The experiment was repeated four times.

(F) Representative images of embryos/yolk sac (top) and placenta (bottom) of E13.5 chimeric embryos derived from MLL1i ESCs (I, left) or MLL1i/1,25-(OH)₂D₃ ESCs (II, right). Scale bar, 0.25 cm.

(G) Quantification of GFP⁺ cells by flow cytometry in each placenta. Placenta with <1% of GFP⁺ cells was considered negative for TE contribution. <1% was chosen as the cutoff, because we never observed more than 0.5% of GFP⁺ in uninjected or mock-treated ESCs. The comparison between MLL1i and MLL1i/1,25-(OH)₂D₃ was done only for embryos that had TE⁺ contribution. *p < 0.05, Mann-Whitney test (unpaired, nonparametric, two tailed).

Yoshizawa et al., 1997), was also observed in *Mll1*KO or MLL1i ESCs (Figure S7D).

Because little is known about the role of *Gc* or the vitamin D signaling pathway in ESCs or their regulation by MLL1, we decided to examine whether *Gc* is a key MLL1 downstream target in EPSC reprogramming. To this end, we knocked down *Gc* by small RNA interference (*siGc*) in GFP⁺ Nagy R1 ESCs. Effective depletion of *Gc* was confirmed by immunofluorescence (Figure S7B, bottom). Unsupervised gene clustering analyses showed that changes of MLL1 target gene expression upon MLL1i or *siGc* partially overlapped (Figure 5F), which were largely distinct from the 2i-treated or untreated ESCs (Figure 5F). More importantly, *Gc* depletion led to upregulation of early cleavage markers (e.g., *Zscan4c*) and downregulation of ESC-specific markers (e.g., *Sall1* and *Nr0b1*), similar to MLL1i and *Mll1*KO (Figure S7E).

Functional characterization confirmed that *Gc* depletion led to a 5-fold increase of 2C⁺ cells in the 2C-reporter assay (Figure 5G), which was modest compared with the 10-fold increase of 2C⁺ cells upon MLL1i (Figure S3A). *Gc*-depleted ESCs could also

be integrated into TE in the chimeric blastocysts (n = 60) at a lower ratio, i.e., 8.3% and 5.8% from two independent experiments, respectively, *in vitro*, as shown by CDX2 expression in the *Gc*-depleted GFP⁺ ESCs (Figure 5H). These results suggest that *Gc* is probably one of the MLL1 downstream intermediates in EPSC reprogramming.

Vitamin D Treatment Promotes Maintenance of the Expanded Pluripotency State

Given upregulation of VDR by MLL1i and *siGc* (Figure S7D), we next tested whether 1,25-(OH)₂D₃, which activates VDR signaling (Figure 6A), could enable the EPSC state. To our surprise, 1,25-(OH)₂D₃ treatment alone was not able to reprogram ESCs to EPSCs. Cells treated with 1,25-(OH)₂D₃ neither contributed to TE when injected into 2C embryos (data shown) nor promoted TSC induction *in vitro* (Figures 6B and 6C). However, combined treatment of 1,25-(OH)₂D₃ and MLL1i led to increased efficiency of TSC induction compared with MLL1i alone in early passages, as indicated by a higher percentage of CDX2⁺ cells
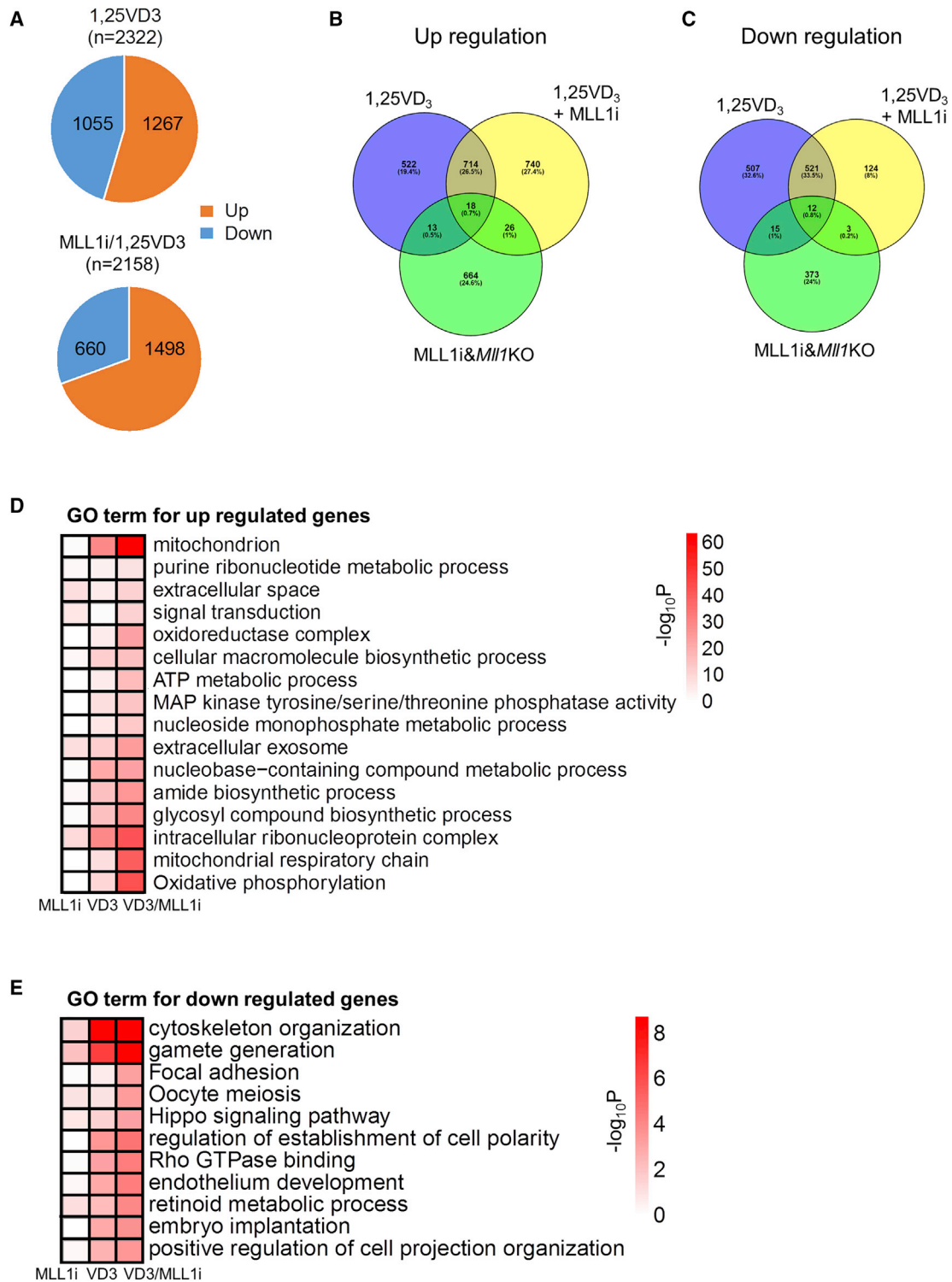
Figure 7. Vitamin D Promotes EPSC Maintenance by Regulating Metabolic Pathways

(A) Venn diagram of genes that were regulated by 1,25-(OH)$_2$D$_3$ treatment (top) or MLL1i/1,25-(OH)$_2$D$_3$ treatment (bottom). Genes were selected by FDR $\leq$ 0.05 and |FC| $\geq$ 1.5 compared with untreated ESCs.

(B and C) Venn diagram for genes (FDR $\leq$ 0.05 and |FC| $\geq$ 1.5) upregulated (B) or downregulated (C) by MLL1i/*Mll1*KO, 1,25-(OH)$_2$D$_3$ or MLL1i/1,25-(OH)$_2$D$_3$. Percentage refers to genes in each category versus all genes that changed expression in three scenarios.

in culture (Figures 6B and 6C). This result suggests that although 1,25-$(OH)_2D_3$ alone was insufficient to induce EPSC (see Discussion), it may facilitate maintenance of EPSCs induced by MLL1i. To test this hypothesis, we sorted single cells expressing the 2C reporter and cultured the single-cell colonies in the presence MLL1i, 1,25-$(OH)_2D_3$, or both (Figure 6D). We found that 1,25-$(OH)_2D_3$ could maintain high $2C^+$ signals in ~20% of P1 colonies derived from the sorted $2C^+$ single cells (Figure 6D). In comparison, only 5% of mock-treated P1 colonies maintained sporadic $2C^+$ expression (Figure 6D, right). Interestingly, 1,25-$(OH)_2D_3$, when added together with MLL1i, significantly increased the stable $2C^+$ P1 colonies (34%) compared with MLL1i alone (17%) (Figure 6D). This result suggests a cooperative function of 1,25-$(OH)_2D_3$ and MLL1 in EPSC maintenance. Consistently, although 1,25-$(OH)_2D_3$ treatment did not increase incidence of ESC contribution to TE *in vivo* (Figure 6E), combined 1,25-$(OH)_2D_3$ and MLL1i treatments led to more robust contribution of GFP^+ ESCs to placenta and yolk sac in embryonic day (E) 13.5 chimeric embryos (Figures 6F and 6G).

MLL1i and Vitamin D Cooperate in Promoting Gene Expression during EPSC Reprogramming

Examining transcriptome changes in 1,25-$(OH)_2D_3$- or 1,25-$(OH)_2D_3$/MLL1i-treated cells showed that 2,332 or 2,158 genes had altered expression after 1,25-$(OH)_2D_3$ or 1,25-$(OH)_2D_3$/MLL1i treatment, respectively (Figure 7A; Table S2). Among them, 1,267 genes were upregulated and 1,055 genes were downregulated by 1,25-$(OH)_2D_3$. Similarly, 1,498 genes were upregulated and 660 genes were downregulated after 1,25-$(OH)_2D_3$/MLL1i treatment. Among genes that were upregulated by 1,25-$(OH)_2D_3$ or 1,25-$(OH)_2D_3$/MLL1i, we found 747 or 810 genes, respectively, that were more highly expressed in preimplantation embryos as compared with ESCs (Table S4). Venn diagrams showed that MLL1i and 1,25-$(OH)_2D_3$ regulated largely distinct sets of genes (Figure 7B). Combined MLL1i and 1,25-$(OH)_2D_3$ treatments led to significant changes in the expression of a unique set of 740 genes (Figure 7B), including 396 genes that have higher expression in preimplantation embryos (Figures S7F and S7G). GO analysis showed that genes upregulated by 1,25-$(OH)_2D_3$/MLL1i treatment were enriched for pathways such as macromolecule biosynthesis, ATP metabolism, and ribonucleotide metabolic process (Figure 7D), consistent with the function of 1,25-$(OH)_2D_3$ in regulating metabolism and homeostasis in cancer cells (Deeb et al., 2007; Sinha et al., 2013). Furthermore, we found that genes upregulated by 1,25-$(OH)_2D_3$/MLL1i treatment had significantly higher enrichment for pathways such as mitochondrion, oxidoreductase complex, and oxidative phosphorylation. Nucleoside and amide biosynthetic pathways, as well as signal transduction, were also modestly more enriched in combined treatment compared with 1,25-$(OH)_2D_3$ or MLL1i treatment alone (Figure 7D). However, genes downregulated by the 1,25-$(OH)_2D_3$/MLL1i treatment

were enriched for pathways including establishment of cellular polarity, focal adhesion, embryo implantation, and Hippo signaling (Figure 7E), which play critical roles in the first cell fate specification (TE versus ICM) in preimplantation embryos (see Discussion) (Frum et al., 2018; Sasaki, 2017). Downregulation of endothelium development and gamete generation were also modestly more enriched after the combined treatment. These results showed that MLL1i and vitamin D have complementary and collaborative functions in shaping the transcription landscape during EPSC reprogramming.

DISCUSSION

Our studies show that histone H3K4 methyltransferase MLL1 plays an important role in restricting cell fate toward the extraembryonic lineages. Inhibiting MLL1 or deletion of *Mll1* in ESCs removes such restriction and allows ESCs to contribute to both embryonic and extraembryonic development *in vitro* and *in vivo*. We also identified an unexpected functional interplay between MLL1 and vitamin D pathway during EPSC reprogramming. We find that *Gc*, which encodes VDB, is a direct MLL1 downstream target. *Gc* knockdown partially mimics MLL1i and promotes the EPSC fate. In serendipity, we find that vitamin D, although it does not induce EPSC by itself, can stabilize the $2C^+$ state *in vitro* and enhance extraembryonic contribution of MLL1i-induced EPSCs *in vivo*. Mechanistic studies highlight the cooperative function of MLL1i and vitamin D in a stepwise EPSC reprogramming process by blocking cellular differentiation and promoting pathways that stabilize preimplantation embryos at the totipotent state *in vitro*.

Our findings, together with our previous observation that MLL1i reprograms EpiSCs to a naive-like pluripotent state (Zhang et al., 2016b), suggest that MLL1 reduces cell fate plasticity and prevents cell fate reversion in early embryonic development. The extraembryonic tissues (i.e., placenta and yolk sac) also contain cells of the extraembryonic mesoderm (e.g., vascular endothelium and blood cells), contributed by the embryonic cell lineages. However, we did not observe extraembryonic contribution from the mock-treated ESCs, suggesting that MLL1i-induced EPSCs have intrinsic bi-directional differentiation potential. To our knowledge, MLL1 is the only chromatin cofactor that has the capacity to govern interconversion of multiple pluripotent states (i.e., EpiSC, ESC, and EPSC). This function is unique among the MLL family histone methyltransferases. Genetic deletion of *Set1a* or *Mll3/4* enzymes in ESCs leads to loss of self-renewal (Fang et al., 2016) and ESC differentiation (Wang et al., 2016). The unique function of MLL1 is consistent with its lower expression in preimplantation embryos, compared with other MLL family histone methyltransferases (data not shown). EPSC reprogramming by MLL1i has several distinct features: (1) the EPSC state remains a transient state *in vitro*, which is in contrast to stable reprogramming of EpiSCs

---

(D) GO term analysis for upregulated gene pathways after MLL1i, 1,25-$(OH)_2D_3$ or MLL1i/1,25-$(OH)_2D_3$ treatment, as indicated on the bottom. The $-\log_{10}$ p value is presented in the heatmap.

(E) GO term analysis for downregulated gene pathways after MLL1i, 1,25-$(OH)_2D_3$, or MLL1i/1,25-$(OH)_2D_3$ treatment, as indicated on the bottom. The $-\log_{10}$ p value is presented in the heatmap.

See also Figure S7 and Table S4.

to naive ESCs by MLL1i; (2) expression of early cleavage markers (e.g., *Zscan4b*, *Dux*, and *Gata2*), unlike ESC transcription factors (Zhang et al., 2016b), requires continuous MLL1i treatment *in vitro*; and (3) EPSC reprogramming is unsynchronized and low in efficiency (Figure 2), with a gene signature of multiple pre-implantation stages (Figure 3). These results suggest that MLL1 inhibition is probably not a deterministic factor for promoting the totipotent-like state. Instead, it is likely that MLL1 inhibition promotes expanded pluripotent cell fate by erasing hallmarks of committed cell fate and thus increases cell fate plasticity (described later). In this regard, the function of MLL1 is distinct from other factors, e.g., *Dux*, miR-34a, CAF-1, and EP400-TIP60 (Abad et al., 2013; Choi et al., 2017; De laco et al., 2017; Hendrickson et al., 2017; Ishiuchi et al., 2015), which directly promotes the 2C-specific transcriptome.

Using the well-established MERVL-driven 2C reporter (Macfarlan et al., 2012), we show that MLL1 inhibition not only induces 2C$^+$/EPSCs but also maintains the high expression level of the 2C reporter in the single-cell-derived 2C$^+$ colonies (Figure 6D). Surprisingly, although vitamin D is unable to increase 2C$^+$ cells in ESCs, it is able to maintain single-cell-derived 2C$^+$ colonies (Figure 6D). The function of vitamin D in maintenance, but not initiation, of the EPSC state is supported by enhanced contribution of MLL1i/1,25-(OH)$_2$D$_3$-treated ESCs to extraembryonic lineages *in vivo* (Figure 6). These results indicate that EPSC reprogramming is a stepwise process that may be subject to a different regulation at each step. Transcriptome analysis showed that MLL1 and vitamin D targets are largely distinct (Figure 7B) and cooperate in shaping the transcription landscape in EPSCs. Given the ability of vitamin D in extending 2C$^+$ signals in single-cell-derived colonies, it will be interesting to test whether vitamin D cooperates with other factors (e.g., *Dux*, miR-34a, CAF-1, and EP400-TIP60) in EPSC reprogramming. It will also be interesting to identify the additional signaling molecule or molecules that facilitate maintenance of the EPSC state *in vitro*, which will have tremendous translational potential.

We have identified multiple gene pathways that are more enriched after combined MLL1i and vitamin D treatment. In particular, mitochondrial respiration, oxidative phosphorylation, and biomolecular synthesis pathways are significantly upregulated, whereas several developmental pathways are significantly downregulated (Figures 7D and 7E). These cooperative changes are at least partly because of MLL1i-mediated downregulation of *Gc* and *Cyp24a1*, negative regulators of 1,25-(OH)$_2$D$_3$, which sensitizes cellular response to exogenous vitamin D (Li et al., 1997; Yoshizawa et al., 1997; Zella et al., 2008). It is paradoxical that MLL1i/1,25-(OH)$_2$D$_3$ treatment leads to significant increase in mitochondrial respiration and OXPHOS gene pathways yet enhances EPSC stability and functional robustness. It has been reported that mitochondria respiration activity is important for cell differentiation, and switching mitochondria oxidative respiration to glycolysis is critical for induced pluripotent stem cell (iPSC) reprogramming (Folmes et al., 2011). Recently, mitochondria dynamics has been reported in cell lineage reprogramming independent of its role in metabolism. It can potentially function as a singling hub during cell fate transition (Khacho et al., 2016). Another noncanonical function of mitochondria proteins has been reported in preim-

plantation embryos (Nagaraj et al., 2017). It will be interesting to examine the functional roles of mitochondrial dynamics and function in EPSC reprogramming. More importantly, it is interesting that the Hippo signaling pathway and establishment of the cell polarity and cell adhesion pathways are among those being downregulated by MLL1i/1,25-(OH)$_2$D$_3$ (Figure 7E). Individual MLL1i and 1,25-(OH)$_2$D$_3$ treatment also modestly downregulates Hippo signaling, but to a lesser extent than the combined treatment (Figure 7E). The Hippo signaling pathway, regulated by cell polarity and cell-cell adhesion, plays a central role in the first cell fate specification, i.e., TE versus ICM (Frum et al., 2018; Sasaki, 2017). Starting from 8-cell embryos, the establishment of ICM and TE is tightly regulated by the Hippo-active inner cells and Hippo-inactive outer cells, respectively (Sasaki, 2017). This is due to negative regulation of TEAD-mediated activation of TE-specific genes (e.g., *Cdx2*) by Hippo signaling (Meng et al., 2016; Ota and Sasaki, 2008). Thus, our study reveals a previously unappreciated function of MLL1 and vitamin D in Hippo signaling, which may underlie MLL1i/1,25-(OH)$_2$D$_3$-medicated expansion of fate commitment to TE both *in vitro* and *in vivo* (Figure 6). Furthermore, because TEAD proteins regulate homeostasis of reactive oxygen species (ROS) (Kaneko and DePamphilis, 2013), inactivating Hippo signaling may also lead to stable accommodation of increased mitochondria respiration *in vitro* (described earlier). The detailed molecular mechanisms for MLL1 and vitamin D signaling in regulating Hippo signaling await future studies.

Finally, our study identifies *Gc* as a MLL1 direct target in ESCs and finds a surprising role of *Gc* in promoting EPSCs. *Gc* encodes the glycoprotein VDB, which is a high-affinity, high-capacity carrier that regulates the bioavailability of vitamin D (Bouillon et al., 1981). It has been shown that VDB functions as a sink for vitamin D (Bikle, 2011; Chun et al., 2014), and its depletion sensitizes the cellular response to extracellular vitamin D treatment (Zella et al., 2008). Importantly, our study shows that knockdown of *Gc* partially recapitulates MLL1 inhibition in EPSC reprogramming, albeit at a more modest level. This suggests that *Gc* probably has functions beyond vitamin D regulation. Previous studies showed that VDB is involved in multiple processes, such as actin scavenging and fatty acid binding (Cooke and Haddad, 1989; Deeb et al., 2007; White and Cooke, 2000). The exact function of *Gc* in cell fate restriction remains to be determined. Nonetheless, our study reveals a direct functional interplay between MLL1 and *Gc*, as well as other genes (e.g., *Vdr* and *Cyp24a1*) in the vitamin D pathway. It would be interesting to examine the function of MLL1 in VDR signaling and vitamin D metabolism in other physiological and pathological contexts.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- LEAD CONTACT AND MATERIALS AVAILABILITY
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Mice
  - Mouse ESC lines

- METHOD DETAILS
  - ○ ESC culture and inhibitor treatment
  - ○ Induction of trophoblast-like stem cells (TSCs)
  - ○ Flow cytometry and single-cell sorting experiment
  - ○ Time-lapse imaging of stable 2C$^+$ cells
  - ○ Immunofluorescence
  - ○ RNA interference
  - ○ Microinjection of 2-cell mouse eggs
  - ○ Embryo *in vitro* culture and *in vivo* development
  - ○ Teratoma and histological analysis
  - ○ Chromatin immunoprecipitation (ChIP) and ChIP-seq
  - ○ Gene expression analyses
  - ○ RNA sequencing (RNA-seq)
  - ○ Single cell RNA sequencing
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - ○ RNA-seq analysis
  - ○ Single cell RNA-seq analysis
  - ○ ChIP-seq analysis
  - ○ Time-series clustering analysis on gene expression
  - ○ Functional enrichment and PCA analysis
  - ○ Statistical analysis
- DATA AND CODE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.celrep.2019.10.074.

AUTHOR CONTRIBUTIONS

H.Z. designed and performed experiments and wrote the manuscript. L.T.P.K. performed IF experiments on embryos and contributed to experimental design. F.M. performed bioinformatics analyses. H.X. and L.W. performed single-cell RNA-seq analyses. Y.H. performed the image quantification. B.Z. generated the *Mll1^f/f* ESCs, conducted the teratoma assays, and performed MLL1 and H3K4me ChIP-seq. M.O. and A.N. helped with EPSC characterization. T.L.S. performed microinjection. Y.D. supervised the overall study and wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

Abad, M., Mosteiro, L., Pantoja, C., Cañamero, M., Rayon, T., Ors, I., Graña, O., Megías, D., Domínguez, O., Martínez, D., et al. (2013). Reprogramming *in vivo* produces teratomas and iPS cells with totipotency features. Nature *502*, 340–345.

Baker, C.L., and Pera, M.F. (2018). Capturing Totipotent Stem Cells. Cell Stem Cell *22*, 25–34.

Beddington, R.S., Morgernstern, J., Land, H., and Hogan, A. (1989). An *in situ* transgenic enzyme marker for the midgestation mouse embryo and the visualization of inner cell mass clones during early organogenesis. Development *106*, 37–46.

Benchetrit, H., Herman, S., van Wietmarschen, N., Wu, T., Makedonski, K., Maoz, N., Yom Tov, N., Stave, D., Lasry, R., Zayat, V., et al. (2015). Extensive Nuclear Reprogramming Underlies Lineage Conversion into Functional Trophoblast Stem-like Cells. Cell Stem Cell *17*, 543–556.

Bikle, D.D. (2011). Vitamin D: an ancient hormone. Exp. Dermatol. *20*, 7–13.

Bouillon, R., Van Assche, F.A., Van Baelen, H., Heyns, W., and De Moor, P. (1981). Influence of the vitamin D-binding protein on the serum concentration of 1,25-dihydroxyvitamin D3. Significance of the free 1,25-dihydroxyvitamin D3 concentration. J. Clin. Invest. *67*, 589–596.

Buecker, C., Srinivasan, R., Wu, Z., Calo, E., Acampora, D., Faial, T., Simeone, A., Tan, M., Swigut, T., and Wysocka, J. (2014). Reorganization of enhancer patterns in transition from naive to primed pluripotency. Cell Stem Cell *14*, 838–853.

Cao, F., Townsend, E.C., Karatas, H., Xu, J., Li, L., Lee, S., Liu, L., Chen, Y., Ouillette, P., Zhu, J., et al. (2014). Targeting MLL1 H3K4 methyltransferase activity in mixed-lineage leukemia. Mol. Cell *53*, 247–261.

Choi, Y.J., Lin, C.P., Risso, D., Chen, S., Kim, T.A., Tan, M.H., Li, J.B., Wu, Y., Chen, C., Xuan, Z., et al. (2017). Deficiency of microRNA *miR-34a* expands cell fate potential in pluripotent stem cells. Science *355*, eaag1927.

Chun, R.F., Peercy, B.E., Orwoll, E.S., Nielson, C.M., Adams, J.S., and Hewison, M. (2014). Vitamin D and DBP: the free hormone hypothesis revisited. J. Steroid Biochem. Mol. Biol. *144*, 132–137.

Cooke, N.E., and Haddad, J.G. (1989). Vitamin D binding protein (Gc-globulin). Endocr. Rev. *10*, 294–307.

Dahl, J.A., Jung, I., Aanes, H., Greggains, G.D., Manaf, A., Lerdrup, M., Li, G., Kuan, S., Li, B., Lee, A.Y., et al. (2016). Broad histone H3K4me3 domains in mouse oocytes modulate maternal-to-zygotic transition. Nature *537*, 548–552.

De Iaco, A., Planet, E., Coluccio, A., Verp, S., Duc, J., and Trono, D. (2017). DUX-family transcription factors regulate zygotic genome activation in placental mammals. Nat. Genet. *49*, 941–945.

De Repentigny, Y., and Kothary, R. (2010). Production of mouse chimeras by injection of embryonic stem cells into the perivitelline space of one-cell stage embryos. Transgenic Res. *19*, 1137–1144.

Deeb, K.K., Trump, D.L., and Johnson, C.S. (2007). Vitamin D signalling pathways in cancer: potential for anticancer therapeutics. Nat. Rev. Cancer *7*, 684–700.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.

Dou, Y., Milne, T.A., Tackett, A.J., Smith, E.R., Fukuda, A., Wysocka, J., Allis, C.D., Chait, B.T., Hess, J.L., and Roeder, R.G. (2005). Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF. Cell *121*, 873–885.

Downs, K.M., and Davies, T. (1993). Staging of gastrulating mouse embryos by morphological landmarks in the dissecting microscope. Development *118*, 1255–1266.

Eckersley-Maslin, M.A., Alda-Catalinas, C., and Reik, W. (2018). Dynamics of the epigenetic landscape during the maternal-to-zygotic transition. Nat. Rev. Mol. Cell Biol. *19*, 436–450.

Evans, M.J., and Kaufman, M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. Nature *292*, 154–156.

Evsikov, A.V., de Vries, W.N., Peaston, A.E., Radford, E.E., Fancher, K.S., Chen, F.H., Blake, J.A., Bult, C.J., Latham, K.E., Solter, D., and Knowles, B.B. (2004). Systems biology of the 2-cell mouse embryo. Cytogenet. Genome Res. *105*, 240–250.

Factor, D.C., Corradin, O., Zentner, G.E., Saiakhova, A., Song, L., Chenoweth, J.G., McKay, R.D., Crawford, G.E., Scacheri, P.C., and Tesar, P.J. (2014). Epigenomic comparison reveals activation of "seed" enhancers during transition from naive to primed pluripotency. Cell Stem Cell 14, 854–863.

Falco, G., Lee, S.L., Stanghellini, I., Bassey, U.C., Hamatani, T., and Ko, M.S. (2007). Zscan4: a novel gene expressed exclusively in late 2-cell embryos and embryonic stem cells. Dev. Biol. 307, 539–550.

Fang, L., Zhang, J., Zhang, H., Yang, X., Jin, X., Zhang, L., Skalnik, D.G., Jin, Y., Zhang, Y., Huang, X., et al. (2016). H3K4 Methyltransferase Set1a Is A Key Oct4 Coactivator Essential for Generation of Oct4 Positive Inner Cell Mass. Stem Cells 34, 565–580.

Folmes, C.D., Nelson, T.J., Martinez-Fernandez, A., Arrell, D.K., Lindor, J.Z., Dzeja, P.P., Ikeda, Y., and Perez-Terzic, C. (2011). Somatic oxidative bioenergetics transitions into pluripotency-dependent glycolysis to facilitate nuclear reprogramming. Cell Metab. 14, 264–271.

Frum, T., Murphy, T.M., and Ralston, A. (2018). HIPPO signaling resolves embryonic cell fate conflicts during establishment of pluripotency in vivo. eLife 7, e42298.

Fujii, S., Nishikawa-Torikai, S., Futatsugi, Y., Toyooka, Y., Yamane, M., Ohtsuka, S., and Niwa, H. (2015). Nr0b1 is a negative regulator of Zscan4c in mouse embryonic stem cells. Sci. Rep. 5, 9146.

Gifford, C.A., Ziller, M.J., Gu, H., Trapnell, C., Donaghey, J., Tsankov, A., Shalek, A.K., Kelley, D.R., Shishkin, A.A., Issner, R., et al. (2013). Transcriptional and epigenetic dynamics during specification of human embryonic stem cells. Cell 153, 1149–1163.

Hendrickson, P.G., Doráis, J.A., Grow, E.J., Whiddon, J.L., Lim, J.W., Wike, C.L., Weaver, B.D., Pflueger, C., Emery, B.R., Wilcox, A.L., et al. (2017). Conserved roles of mouse DUX and human DUX4 in activating cleavage-stage genes and MERVL/HERVL retrotransposons. Nat. Genet. 49, 925–934.

Ishiuchi, T., Enriquez-Gasca, R., Mizutani, E., Bošković, A., Ziegler-Birling, C., Rodriguez-Terrones, D., Wakayama, T., Vaquerizas, J.M., and Torres-Padilla, M.E. (2015). Early embryonic-like cells are induced by downregulating replication-dependent chromatin assembly. Nat. Struct. Mol. Biol. 22, 662–671.

Ishiuchi, T., Ohishi, H., Sato, T., Kamimura, S., Yorino, M., Abe, S., Suzuki, A., Wakayama, T., Suyama, M., and Sasaki, H. (2019). Zfp281 Shapes the Transcriptome of Trophoblast Stem Cells and Is Essential for Placental Development. Cell Rep. 27, 1742–1754.

Jude, C.D., Climer, L., Xu, D., Artinger, E., Fisher, J.K., and Ernst, P. (2007). Unique and independent roles for MLL in adult hematopoietic stem cells and progenitors. Cell Stem Cell 1, 324–337.

Kaneko, K.J., and DePamphilis, M.L. (2013). TEAD4 establishes the energy homeostasis essential for blastocoel formation. Development 140, 3680–3690.

Karatas, H., Townsend, E.C., Cao, F., Chen, Y., Bernard, D., Liu, L., Lei, M., Dou, Y., and Wang, S. (2013). High-affinity, small-molecule peptidomimetic inhibitors of MLL1/WDR5 protein-protein interaction. J. Am. Chem. Soc. 135, 669–682.

Katada, S., and Sassone-Corsi, P. (2010). The histone methyltransferase MLL1 permits the oscillation of circadian gene expression. Nat. Struct. Mol. Biol. 17, 1414–1421.

Khacho, M., Clark, A., Svoboda, D.S., Azzi, J., MacLaurin, J.G., Meghaizel, C., Sesaki, H., Lagace, D.C., Germain, M., Harper, M.E., et al. (2016). Mitochondrial Dynamics Impacts Stem Cell Identity and Fate Decisions by Regulating a Nuclear Transcriptional Program. Cell Stem Cell 19, 232–247.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 14, R36.

Kubaczka, C., Senner, C.E., Cierlitza, M., Araúzo-Bravo, M.J., Kuckenberg, P., Peitz, M., Hemberger, M., and Schorle, H. (2015). Direct Induction of Trophoblast Stem Cells from Murine Fibroblasts. Cell Stem Cell 17, 557–568.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359.

Li, Y.C., Pirro, A.E., Amling, M., Delling, G., Baron, R., Bronson, R., and Demay, M.B. (1997). Targeted ablation of the vitamin D receptor: an animal model of vitamin D-dependent rickets type II with alopecia. Proc. Natl. Acad. Sci. USA 94, 9831–9835.

Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30, 923–930.

Liao, Y., Smyth, G.K., and Shi, W. (2019). The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. Nucleic Acids Res. 47, e47.

Lim, D.A., Huang, Y.C., Swigut, T., Mirick, A.L., Garcia-Verdugo, J.M., Wysocka, J., Ernst, P., and Alvarez-Buylla, A. (2009). Chromatin remodelling factor Mll1 is essential for neurogenesis from postnatal neural stem cells. Nature 458, 529–533.

Liu, X., Wang, C., Liu, W., Li, J., Li, C., Kou, X., Chen, J., Zhao, Y., Gao, H., Wang, H., et al. (2016). Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation embryos. Nature 537, 558–562.

Macfarlan, T.S., Gifford, W.D., Driscoll, S., Lettieri, K., Rowe, H.M., Bonanomi, D., Firth, A., Singer, O., Trono, D., and Pfaff, S.L. (2012). Embryonic stem cell potency fluctuates with endogenous retrovirus activity. Nature 487, 57–63.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17, 10–12.

McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. Nat. Biotechnol. 28, 495–501.

McMahon, K.A., Hiew, S.Y., Hadjur, S., Veiga-Fernandes, H., Menzel, U., Price, A.J., Kioussis, D., Williams, O., and Brady, H.J. (2007). Mll has a critical role in fetal and adult hematopoietic stem cell self-renewal. Cell Stem Cell 1, 338–345.

Meng, Z., Moroishi, T., and Guan, K.-L. (2016). Mechanisms of Hippo pathway regulation. Genes Dev. 30, 1–17.

Morgani, S.M., Canham, M.A., Nichols, J., Sharov, A.A., Migueles, R.P., Ko, M.S., and Brickman, J.M. (2013). Totipotent embryonic stem cells arise in ground-state culture conditions. Cell Rep. 3, 1945–1957.

Nagaraj, R., Sharpley, M.S., Chi, F., Braas, D., Zhou, Y., Kim, R., Clark, A.T., and Banerjee, U. (2017). Nuclear Localization of Mitochondrial TCA Cycle Enzymes as a Critical Step in Mammalian Zygotic Genome Activation. Cell 168, 210–223.

Nichols, J., Silva, J., Roode, M., and Smith, A. (2009). Suppression of Erk signalling promotes ground state pluripotency in the mouse embryo. Development 136, 3215–3222.

Niwa, H., Toyooka, Y., Shimosato, D., Strumpf, D., Takahashi, K., Yagi, R., and Rossant, J. (2005). Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. Cell 123, 917–929.

Ota, M., and Sasaki, H. (2008). Mammalian Tead proteins regulate cell proliferation and contact inhibition as transcriptional mediators of Hippo signaling. Development 135, 4059–4069.

Papaioannou, V.E., Mkandawire, J., and Biggers, J.D. (1989). Development and phenotypic variability of genetically identical half mouse embryos. Development 106, 817–827.

Poueymirou, W.T., Auerbach, W., Frendewey, D., Hickey, J.F., Escaravage, J.M., Esau, L., Doré, A.T., Stevens, S., Adams, N.C., Dominguez, M.G., et al. (2007). F0 generation mice fully derived from gene-targeted embryonic stem cells allowing immediate phenotypic analyses. Nat. Biotechnol. 25, 91–99.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842.

Ramírez, F., Dündar, F., Diehl, S., Grüning, B.A., and Manke, T. (2014). deepTools: a flexible platform for exploring deep-sequencing data. Nucleic Acids Res. 42, W187–W191.

Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. 44 (W1), W160–W165.

Rodriguez-Terrones, D., Gaume, X., Ishiuchi, T., Weiss, A., Kopp, A., Kruse, K., Penning, A., Vaquerizas, J.M., Brino, L., and Torres-Padilla, M.E. (2018). A molecular roadmap for the emergence of early-embryonic-like cells in culture. Nat. Genet. *50*, 106–119.

Rugg-Gunn, P.J., Cox, B.J., Lanner, F., Sharma, P., Ignatchenko, V., McDonald, A.C., Garner, J., Gramolini, A.O., Rossant, J., and Kislinger, T. (2012). Cell-surface proteomics identifies lineage-specific markers of embryo-derived stem cells. Dev. Cell *22*, 887–901.

Sasaki, H. (2017). Roles and regulations of Hippo signaling during preimplantation mouse development. Dev. Growth Differ. *59*, 12–20.

Satija, R., Farrell, J.A., Gennert, D., Schier, A.F., and Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. Nat. Biotechnol. *33*, 495–502.

Schwartzberg, P.L., Goff, S.P., and Robertson, E.J. (1989). Germ-line transmission of a c-abl mutation produced by targeted gene disruption in ES cells. Science *246*, 799–803.

Sinha, A., Hollingsworth, K.G., Ball, S., and Cheetham, T. (2013). Improving the vitamin D status of vitamin D deficient adults is associated with improved mitochondrial oxidative function in skeletal muscle. J. Clin. Endocrinol. Metab. *98*, E509–E513.

Tanaka, S., Kunath, T., Hadjantonakis, A.K., Nagy, A., and Rossant, J. (1998). Promotion of trophoblast stem cell proliferation by FGF4. Science *282*, 2072–2075.

Tarkowski, A.K. (1959). Experiments on the development of isolated blastomers of mouse eggs. Nature *184*, 1286–1287.

Thorvaldsdóttir, H., Robinson, J.T., and Mesirov, J.P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief. Bioinform. *14*, 178–192.

Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N.J., Livak, K.J., Mikkelsen, T.S., and Rinn, J.L. (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat. Biotechnol. *32*, 381–386.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat. Protoc. *7*, 562–578.

Van Keuren, M.L., Gavrilina, G.B., Filipiak, W.E., Zeidler, M.G., and Saunders, T.L. (2009). Generating transgenic mice from bacterial artificial chromosomes: transgenesis efficiency, integration and expression outcomes. Transgenic Res. *18*, 769–785.

Vergara, G.J., Irwin, M.H., Moffatt, R.J., and Pinkert, C.A. (1997). *In vitro* fertilization in mice: Strain differences in response to superovulation protocols and effect of cumulus cell removal. Theriogenology *47*, 1245–1252.

Wang, C., Lee, J.E., Lai, B., Macfarlan, T.S., Xu, S., Zhuang, L., Liu, C., Peng, W., and Ge, K. (2016). Enhancer priming by H3K4 methyltransferase MLL4 controls cell fate transition. Proc. Natl. Acad. Sci. USA *113*, 11871–11876.

White, P., and Cooke, N. (2000). The multifunctional properties and characteristics of vitamin D-binding protein. Trends Endocrinol. Metab. *11*, 320–327.

Wu, J., Huang, B., Chen, H., Yin, Q., Liu, Y., Xiang, Y., Zhang, B., Liu, B., Wang, Q., Xia, W., et al. (2016). The landscape of accessible chromatin in mammalian preimplantation embryos. Nature *534*, 652–657.

Wu, G., Lei, L., and Schöler, H.R. (2017). Totipotency in the mouse. J. Mol. Med. (Berl.) *95*, 687–694.

Yang, J., Ryan, D.J., Wang, W., Tsang, J.C., Lan, G., Masaki, H., Gao, X., Antunes, L., Yu, Y., Zhu, Z., et al. (2017a). Establishment of mouse expanded potential stem cells. Nature *550*, 393–397.

Yang, Y., Liu, B., Xu, J., Wang, J., Wu, J., Shi, C., Xu, Y., Dong, J., Wang, C., Lai, W., et al. (2017b). Derivation of Pluripotent Stem Cells with *In Vivo* Embryonic and Extraembryonic Potency. Cell *169*, 243–257.

Ying, Q.L., 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.

Yoshizawa, T., Handa, Y., Uematsu, Y., Takeda, S., Sekine, K., Yoshihara, Y., Kawakami, T., Arioka, K., Sato, H., Uchiyama, Y., et al. (1997). Mice lacking the vitamin D receptor exhibit impaired bone formation, uterine hypoplasia and growth retardation after weaning. Nat. Genet. *16*, 391–396.

Yu, G., Wang, L.G., Han, Y., and He, Q.Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS *16*, 284–287.

Zella, L.A., Shevde, N.K., Hollis, B.W., Cooke, N.E., and Pike, J.W. (2008). Vitamin D-binding protein influences total circulating levels of 1,25-dihydroxyvitamin D3 but does not directly modulate the bioactive levels of the hormone *in vivo*. Endocrinology *149*, 3656–3667.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137.

Zhang, B., Zheng, H., Huang, B., Li, W., Xiang, Y., Peng, X., Ming, J., Wu, X., Zhang, Y., Xu, Q., et al. (2016a). Allelic reprogramming of the histone modification H3K4me3 in early mammalian development. Nature *537*, 553–557.

Zerbino, D.R., Achuthan, P., Akanni, W., Amode, M.R., Barrell, D., Bhai, J., Billis, K., Cummins, C., Gall, A., Girón, C.G., et al. (2018). Ensembl 2018. Nucleic Acids Res. D754–D761.

Zhang, H., Gayen, S., Xiong, J., Zhou, B., Shanmugam, A.K., Sun, Y., Karatas, H., Liu, L., Rao, R.C., Wang, S., et al. (2016b). MLL1 Inhibition Reprograms Epiblast Stem Cells to Naive Pluripotency. Cell Stem Cell *18*, 481–494.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
anti-H3K4me1	Abcam	Cat# ab8895; RRID: AB_306847
anti-H3K4me2	Millipore	Cat# 07-030; RRID: AB_310342
anti-H3K4me3	Millipore	Cat# 07-473; RRID: AB_1977252
PE-conjugated anti-mouse CD40 Antibody	Biolegend	Cat# 124609; RRID: AB_1134084
Anti-CDX2 antibody	Biogenex	Cat# NU777-5UC; RRID: N/A
Anti-MLL/HRX antibody, CT., clone 9-12	Millipore	Cat# 05-765; RRID: AB_309977
Anti-Oct-3/4 antibody (C-10) PE	Santa Cruz	Cat# sc-5279; RRID: AB_628051
Anti-GC antibody	Thermo Fisher	Cat# PA5-29082; RRID: AB_2546558
Anti-VDR (D-6) FITC-conjugated antibody	Santa Cruz	Cat# sc-13133 FITC
FITC-conjugated goat anti-rabbit IgG (H+L) cross adsorbed secondary antibody	Thermo Fisher	Cat# 31583; RRID: AB_228239
Alexa Fluor® 555-conjugated goat anti-mouse IgG	Abcam	Cat# ab150114; RRID: AB_2687594
Alexa Fluor Plus 647-conjugated goat anti-rabbit IgG highly cross-adsorbed secondary antibody,	Invitrogen	Cat# A32733; RRID: AB_2633282
Alexa Fluor 488 Tyramide SuperBoost Kit, goat anti-mouse IgG	Thermo Fisher	Cat# B40941; RRID: N/A
10-min. Tissue (Tumor) Dissociation / Single Cell Isolation Kit	101Bio	Cat# P712-50; RRID: N/A
Chemicals, Peptides, and Recombinant Proteins		
ESGRO® leukemia inhibitory factor (LIF)	EMD Millipore	Cat# ESG1107
Recombinant Human FGF-4 Protein	R&D Systems Inc.	Cat# 235-F4-025
Heparin	Sigma-Aldrich	Cat# H3149
KnockOut DMEM	Thermo Fisher	Cat# 10829018
RPMI Media 1640	Thermo Fisher	Cat# 11875-093
Fetal bovine serum (FBS)	Atlas Biologicals	Cat# F-0500-D
KnockOut Serum Replacement	Thermo Fisher	Cat# 10828-028
PD 0325901	Sigma Aldrich	Cat# PZ0162
CHIR99021	Sigma Aldrich	Cat# SML1046
1α,25-Dihydroxyvitamin D3	Sigma Aldrich	Cat# D1530
Hygromycin	Sigma Aldrich	Cat# H3274
Geltrex® Matrix Products	Thermo Fisher	Cat# A1413202
KSOMaa Evolve^R medium	Zenith Biotech	Cat# ZEKS-050
Bovine Serum Albumin solution	SIGMA	Cat# A8412
Mineral oil	SIGMA-ALDRICH	Cat# M8410
EmbryoMax® 0.1% Gelatin Solution	Millipore	Cat# ES-006-B
VECTOR Red Alkaline Phosphatase (Red AP) Substrate Kit	Vector Laboratories	Cat# SK-5100
FuGENE® 6	Promega	Cat# E2691
Normal Goat Serum	Cell Signaling Technology	Cat# 5425
Deposited Data		
RNA-seq, ChIP–seq	This paper	GSE107408
Single cell RNA-seq	This paper	GSE127943
Flow cytometry	http://flowrepository.org/	FR-FCM-Z29P

(*Continued on next page*)

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental Models: Organisms/Strains		
Nagy ES cell line R1 with EGFP (D4)	MMRRC	Cat# 011980-MU
Mll1^{fl/fl};Cre-ER^{TM} ESCs	Y. Dou	Zhang et al., 2016b
REAGENT or RESOURCE	SOURCE	IDENTIFIER
H2B-Tomato E14Tg2a ESCs	J.M. Brickman	Morgani et al., 2013
C57BL/6NCrl female mice	Charles River Laboratory	Cat# 493
(C57BL/6J X SJL/J)F1 males	Jackson Laboratory	Cat# 100012
Severe combined immunodeficiency (SCID) mice	Charles River Laboratory	Cat# 236-CB17SCID
Oligonucleotides		
GC siRNAs	Thermo Fisher	Cat# s66510
GC siRNAs	Thermo Fisher	Cat# s66511
Silencer® Negative Control No. 4 siRNA	Fisher	Cat# AM4641
See STAR Methods for Primers for Real-time PCR	Integrated DNA Technologies	N/A
Plasmids		
2C::tdTomato Reporter	Macfarlan et al., 2012	Addgene #40281
Flag-MLL1(full length)	Y. Dou	Dou et al., 2005
Software and Algorithms		
Tophat2, v2.1.1	Kim et al., 2013	https://ccb.jhu.edu/software/tophat/manual.shtml
STEM, v1.3.11	Wu et al., 2016	http://www.cs.cmu.edu/~jernst/stem
Integrative Genomics Viewer (IGV), v2.3	Thorvaldsdóttir et al., 2013	http://software.broadinstitute.org/software/igv/
Cufflinks, v2.2.1	Trapnell et al., 2012	http://cole-trapnell-lab.github.io/cufflinks/
CummeRbund, v2.1	N/A	http://compbio.mit.edu/cummeRbund/
Bowtie2, v2-2.2.4	Langmead and Salzberg, 2012	
MACS, v1.4.2	Zhang et al., 2008	http://liulab.dfci.harvard.edu/MACS/
HOMER	Salk Institute	http://homer.salk.edu
GREAT, v3	McLean et al., 2010	http://great.stanford.edu/public/html/
DeepTools2, v2.5.0	Ramírez et al., 2016	
Panther Classification System, v13.0	N/A	http://pantherdb.org/
GraphPad Prism, v7	GraphPad	https://www.graphpad.com/scientific-software/prism/
FlowJo, v10.0.7r2	FlowJo LLC	http://docs.flowjo.com/d2
R, v3.4.3	R foundation	https://cran.r-project.org/
Cutadapt, V1.8.1	Martin, 2011	https://cutadapt.readthedocs.io/en/stable/
Fastq_screen, v0.5.1.4	Babraham institute	http://www.bioinformatics.babraham.ac.uk/projects/fastq_screen/
STAR, v2.5.2b	Dobin et al., 2013	https://code.google.com/archive/p/rna-star/
Subread, v1.4.6-p1	Liao et al., 2019	http://subread.sourceforge.net/
Seurat, V2.3.4	Satija et al., 2015	https://cran.r-project.org/web/packages/Seurat/index.html
Monocle, V2.6.1	Trapnell et al., 2014	http://cole-trapnell-lab.github.io/monocle-release/
Ggplot2, V3.0.0	Salk Institute	https://cran.r-project.org/web/packages/ggplot2/
Other		
RNA-seq for pre-implantation embyos	Wu et al., 2016	GEO: GSE66582
RNA-seq for EpiSC reprograming	Zhang et al., 2016b	GEO: GSE66112

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Yali Dou (yalid@med.umich.edu). This study did not generate new unique reagents.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mice

All animal procedures were reviewed and approved by the University of Michigan Institutional Animal Care and Use Committee. Animals were housed by the Unit for Laboratory Animal Medicine. The animal care and use program conforms to the standards of "The Guide for the Care and Use of Laboratory Animals," Revised 2011. ULAM is accredited by the Association for Assessment and Accreditation of Laboratory Animal Care, International (AAALAC, Intl).

Mouse ESC lines

The derivation and culture of $Mll1^{fl/fl};Cre-ER^{TM}$ ESCs were described previously (Zhang et al., 2016b). E14Tg2a ESCs express the H2B-driven Tomato reporter as previously described (Morgani et al., 2013). The CMV-EGFG$^+$ R1 ESCs were purchased from MMRRC (Mutant Mouse Resource& Research Centers). Strain name is Nagy ES cell line R1 with EGFP (D4). This is a mouse ESC line that constitutively expresses EGFP in all lineages throughout development. The $Mll1^{fl/fl};Cre-ER^{TM}$ ESCs was derived from C57BL/6 mouse strain. H2B-Tomato E14 ESCs was derived from 129/Ola mouse strain. CMV-EGFG$^+$ R1 ESCs were derived from 129X1/S1 mouse strain. For the 2C::tdTomato reporter ESC line, $Mll1^{fl/fl};Cre-ER^{TM}$ ESCs were transfected with the plasmid (Addgene, #40281) (Macfarlan et al., 2012) using FuGENE® 6 (Promega, #E2691) and selected with 150 μg ml^{-1} hygromycin (Sigma, H3274) for two weeks. The cells were maintained in media containing 50 μg ml^{-1} hygromycin. Spontaneous activation of the 2C::tdTomato reporter in $Mll1^{fl/fl};Cre-ER^{TM}$ ESCs was around 0.5%, consistent with previous reports (Ishiuchi et al., 2015; Macfarlan et al., 2012).

METHOD DETAILS

ESC culture and inhibitor treatment

All ESC lines in this study were adapted to the KnockOut D-MEM-based medium (Thermo Fisher Scientific) supplemented with 2 mM Glutamax (Thermo Fisher Scientific), 1X nonessential amino acids (Thermo Fisher Scientific), 1X sodium pyruvate (Thermo Fisher Scientific), 0.1 mM 2-mercaptoethanol (Sigma-Aldrich), 10^3 U ml^{-1} LIF (Millipore, ESG1107). Cells were cultured with either 20% FBS (Atlas Biologicals, catalog number: F-0500-D) if grown on top of MEF feeder cells or 15% KO-KSR (Thermo Fisher Scientific 10828028) and 5% FBS if grown on Matrigel (Thermo Fisher Scientific, A1413202) coated plates, or 20% KO-KSR if grown under 2i condition. 2i consists of 1μM PD0325901 (Sigma Aldrich, PZ0162) and 3μM CHIR99021 (Sigma Aldrich, SML1046). 1α,25-Dihydroxyvitamin D$_3$ (Sigma Aldrich, D1530) is used at a final concentration of 1 μM. $Mll1$ was deleted by treating $Mll1^{fl/fl}$-Cre-ER^{TM} ESCs with 5μM 4-OHT for 48hrs. MLL1 inhibition was done by MM-401 treatment at a final concentration of 100 μM (Zhang et al., 2016b).

Induction of trophoblast-like stem cells (TSCs)

For TSC induction, ESCs were trypsinized and plated at density of 1-2x10^4 cells/cm^2 for each passage and continuously cultured in medium containing RPMI1640, 70% feeder cell conditioned medium (CM), 15% KSR (Thermo Fisher Scientific), 5% Serum, 25 ng/ml human recombinant FGF4 (R&D, 235-F4-025) and 1 μg/ml heparin (Sigma-Aldrich, H3149) for four passages.

Flow cytometry and single-cell sorting experiment

Flow cytometry was performed using FACS Calibur (BD Biosciences). The antibodies were CD40-PE (Biolegend, 124609), CDX2 (Biogenex, NU777-5UC) and Goat anti-Rabbit IgG (H+L) Cross-Adsorbed Secondary Antibody FITC (Thermo Fisher Scientific, 31583). Analyses were performed using FlowJo software. Gating methodology was previously described and included in Supplemental Information (Macfarlan et al., 2012). SY3200 flow cytometer (Sony Biotechnology) was used to sort 2C$^+$ cells, which were seeded either as single cell or 1000 cells per well in a 96-well plate. Single cell clones were examined using IX73 microscope system (Olympus) with the CyC5.5 filter-set.

Time-lapse imaging of stable 2C$^+$ cells

The 2C$^+$ cells were cultured in the Matrigel-coated 24-well plate 12hr before imaging. Time-lapse videos and images were acquired using the Axio Observer.Z1microscope with a cage incubator (37 °C and 5% CO$_2$. Time-lapses images (Rhodamine and phase contrast) were sequentially captured under LD Plan-Neofluar 20 × /0.4 Korr differential interference contrast (DIC) objective for all positions. Images were automatically acquired at 10 min intervals for 24 hr. The image sequences were exported via separated or merged channel as an .avi files and processed by ZEN Pro Digital Imaging Software.

Immunofluorescence

Alexa Fluor 488 Tyramide SuperBoost Kit and goat anti-mouse IgG (ThermoFisher, B40941) were used according to manufacturer's protocol with modifications. Briefly, cells were fixed in 4% PFA/PBS for 30 min at 4°C, washed twice for 10 min in phosphate-buffered saline (PBS) and permeabilized in 0.25% Triton X-100/PBS (PBST) for 30 min at 4°C. Cells were incubated with H_2O_2 for 60 min at room temperature to quench the endogenous peroxidase activity, followed by PBS wash for three times and blocking with 10% Goat serum for 60 min at room temperature. The primary antibody incubation was performed overnight at 4°C for the MLL1 antibody (Millipore, 05-765). After washing three times with PBS, cells were incubated with poly-HRP-conjugated secondary antibody overnight at 4°C. Nuclei were co-stained with DAPI. Immunofluorescence for Gc (Thermo Fisher, PA5-29082; 1:200 dilution) and VDR (Santa Cruz, sc-13133 FITC; 1:100 dilution) were performed as described (Zhang et al., 2016b). The images were captured using the IX73 microscope system (Olympus), and further processed by ImageJ.

RNA interference

ESCs were plated at a density of about 1×10^5 ESCs per 12-well culture plate, and were transfected with siRNA at final concentration of 10 nM. Two independent siRNA sequences for *Gc* were purchased from Silencer® Select (Thermo Fisher, s66510, s66511) with a Silencer® negative control siRNA (Thermo Fisher, AM4641). siRNA resuspension and transfection were performed by FuGENE® 6 transfection reagents according to manufacturer's instruction, and cells were harvested for RNA extraction approximately 96h post transfection.

Microinjection of 2-cell mouse eggs

Two-cell mouse eggs were collected from super-ovulated C57BL/6NCrl female mice (Charles River Laboratory Strain Code 493) that were treated with PMSG and HCG as described (Van Keuren et al., 2009) and mated with (C57BL/6J X SJL/J)F1 males (Jackson Laboratory Stock Number 100012). C57BL/6 egg donors were selected because they can be efficiently colonized by ESCs(Schwartzberg et al., 1989). Hybrid males were used because of their superior fertility(Vergara et al., 1997). On embryonic day 1.5 of development, two-cell eggs were collected by flushing oviducts with M2 medium supplemented with penicillin and streptomycin (Sigma Aldrich). Eggs were placed in culture in KSOMaa Evolve^R medium (Zenith Biotech, ZEKS-050) supplemented with 4 mg/ml bovine serum albumin in a 37°C, 5% CO_2, humidified incubator. Embryos were transferred to M2 medium in a hanging drop microinjection chamber on a Leica AM6000 Advanced Micromanipulation system equipped with a laser for microinjection. The zona pellucida sliced with an XY Clone laser system (Hamilton-Thorne). Laser settings were adjusted to 100% power with a 312 micro-second pulse. An ES Transfer Tip (Eppendorf) containing 8 ES cells was inserted into the egg. ESCs were expelled under the zona pellucida with a CellTram Oil piston syringe (Eppendorf) so that they were placed in direct contact with the egg's cells. Following microinjection, the eggs were washed in four 75μl drops of KSOMaa under mineral oil (SIGMA ALDRICH, M8410) and incubated until they developed into blastocysts. Mouse eggs were surgically transferred to the oviducts of pseudopregnant (C57BL/6J X DBA/2) F1 female mice as described(Van Keuren et al., 2009). The method described here is a variation of microinjection of ESCs into single-cell fertilized eggs (De Repentigny and Kothary, 2010) and the laser-assisted microinjection of ES cells into 8-cell mouse eggs (Poueymirou et al., 2007).

Embryo *in vitro* culture and *in vivo* development

After microinjection, 2-cell mouse eggs injected with ESCs were cultured up to the blastocyst stage. Blastocysts were washed twice in PBS solution containing 0.2% bovine serum albumin (BSA) and fixed in 4% paraformaldehyde (PFA, Ted Pella Inc.18505)/PBS for 2 hours at 4°C, washed in 0.1% Triton X-100 (SIGMA ALDRICH, 93443)/PBS twice, total of 10 min. They were permeabilized in 0.25% Triton X-100/PBS (PBST) for 30 min at 4°C and blocked in blocking solution containing 10% Goat serum (Cell Signaling Technology, 5425) and PBS solution with 0.1% BSA and 0.01% Tween20. The processed blastocysts were incubated with primary antibodies overnight at 4°C. Nuclei were counter-stained with DAPI (ThermoFisher Scientific, 62248) in 0.25% PBST for 15 min at room temperature before mounting onto glass slides in Vectashield solution (Vector Laboratories, H-1200) for imaging. Antibodies were Oct3/4 (Santa Cruz Biotechology, sc-5279), Cdx2 (Biogenex, NU777-5UC), goat anti-rabbit IgG Alexa Fluor Plus 647 (Invitrogen, A32733) and goat anti-mouse IgG Alexa Fluor^R555 (Abcam, ab150114). Blastocyst images were captured by a Leica SP5 upright 2-photon confocal microscope (Leica Microsystems). Fluorescence was excited with a 405-nm UV laser (DAPI), a 488-nm laser (GFP), a 561-nm laser (Alexa Fluor555) and a 633-nm laser (Alexa Fluor647). All confocal images were acquired using HC PL APO CS2 40X/1.25 oil immersion objective lens, with an optical section thickness of 0.5 μm. Image data was further processed by either Leica software or ImageJ. For chimeric experiments, embryos at embryonic day E11.5 and E13.5 were staged and dissected in accordance with Downs and Davies (1993). Chimeric embryos were detected by Stereo fluorescent microscope (Leica Microsystems).

Teratoma and histological analysis

Teratoma generation was described previously (Zhang et al., 2016b). Briefly, 4-weeks old female severe combined immunodeficiency (SCID) mice (Charles River Laboratory Strain Code 236-CB17SCID) were subcutaneously injected with ESCs for teratoma formation. Approximately 9 weeks after injection, tumors were dissected and fixed in 4% paraformaldehyde. Tissues were then Paraffin-embedded and subjected for either Hematoxylin and Eosin (H&E) staining or immunohistochemistry using anti-GFP antibody (Rockland, 600-406-215) using standard procedures at University of Michigan ULAM *in vivo* animal core.

Chromatin immunoprecipitation (ChIP) and ChIP-seq

ChIP assays were carried out as previously described (Zhang et al., 2016b). ChIP antibodies are: anti-H3K4me1 (Abcam, ab8895), anti-H3K4me2 and H3K4me3 (Millipore, 07-030 and 07-473) and custom anti-MLL1 antibody as previously described (Dou et al., 2005; Zhang et al., 2016b). About 1E10 of ESCs were used for MLL1 and 2E7 cells were used for H3K4me ChIP-seq. All cells were cultured on Matrigel-coated plates. ChIP-seq library preparation and sequencing (HiSeq 2000) were performed at DNA Sequencing Core Facility, University of Michigan.

Gene expression analyses

For quantitative reverse transcription (qRT–PCR) analysis, first-strand cDNA was generated from 2 μg total RNA using Superscript III (Invitrogen) and random hexamer priming (Invitrogen). For qPCR with ChIP (qChIP–PCR) analysis, ChIP- and input- DNA were quantitated by Qubit and normalized to 2.5ng per qPCR reaction. qPCR was performed using SYBR green master mix (Applied Biosystems) or Radiant Green Lo-ROX qPCR Kit (Alkali Scientific Inc.) in triplicate. At least two biological replicates were performed. qPCR Primers were designed by PrimerBank: *Gc* (ID# 12846768a1), *Vdr* (ID# 31543944a1), *Cyp24a1* (ID# 6753572a1) and *Actb* (ID# 6671509a1). ChIP-qPCR primers for MLL1 binding site at *Gc* are F: TTTGCTAGGGAGAGCAGTGT; R: CGCATGCTGCTTCAATCCTG.

RNA sequencing (RNA-seq)

Total RNAs were extracted by the RNeasy mini kit (QIAGEN) in accordance with manufacture instructions. To eliminate residual genomic DNA, total RNAs were treated on-column with RNA-free DNase I (Thermo Fisher Scientific). 10 ng of total RNAs were used for library preparation. In all cases, two replicates per condition were sequenced on Illumina HiSeq2500 High-Output with v4 single end 50 cycle at the University of Michigan core facility.

Single cell RNA sequencing

Single-cell sequencing library was generated by using the iCell8 platform (Takara). Specifically, single cell suspension of ESCs was stained with a mixture of Hoechst 33342 and Propidium Iodide (R37610, Thermo Scientific) according to the manufacturer's instruction. After staining, cells were washed by PBS, and counted by MoxiTM Automated Cell Counter. Cell suspension (20,000 cells/ml) was submitted to the MultiSample NanoDispenser (MSND, Wafergen Biosystems) for single cell preparation. The dispensed cells were imaged and single live cells, defined by Hoechst-positive and Propidium Iodide-negative staining, were selected. Selected cells were subjected to reverse transcription and first-step amplification in a Chip Cycler (Bio-Rad), and the resulting cDNA was purified and size-selected with Agencourt AMPure XP beads (A63880, Beckman Coulter). One ng of purified cDNA was used to generate a sequencing library using the Nextera XT DNA sample preparation Kit (FC-131-1024, Illumina). Libraries were sequenced on the NextSeq500 sequencer (Illumina).

QUANTIFICATION AND STATISTICAL ANALYSIS

RNA-seq analysis

All RNA-seq data were mapped to the mm10 mouse genome using Tophat2 (2.1.1) (Kim et al., 2013) with default parameters. The bigwig files for visualization in Integrative Genomics Viewer (IGV) (Thorvaldsdóttir et al., 2013) were generated from BAM files by using "bamCoverage" from deepTools (Ramírez et al., 2014) with parameters "–outFileFormat bigwig–normalizeUsingRPKM–minMappingQuality 30–binSize 50." For differential gene expression, aligned reads were assembled using Cufflinks (v2.2.1) (Trapnell et al., 2012) and assembled transcriptome catalog was used as input for Cuffdiff2 (2.2.1) (Trapnell et al., 2012), with quartile normalization method. CummeRbund v2.1 (http://compbio.mit.edu/cummeRbund/) was used to process, index, and visualize the output of the Cuffdiff2 results. Genes with FDR < 0.05 and fold change (FC) > 1.5 were considered to be differentially expressed genes. For Pearson correlation coefficients, log10 transformed FPKM values of all the selected genes were used as input. Pearson correlation coefficients between each two samples were calculated using function cor() in R. The matrix of Pearson correlation coefficients was used to generate heatmap using R package corrplot. For gene clustering analyses, Z-score-normalized FPKM signals from developmental stage-specific dataset (GEO GSE66582) (Wu et al., 2016) or $\log_2$ FC(FPKM) from ESCs dataset were used in the clustering analysis by hcluster. Heatmap were visualized by R package heatmap.

Single cell RNA-seq analysis

Perl pipeline developed by WaferGen was used to process single cell RNA-seq data. Paired raw reads with designed barcode (10nt) and UMI (11-14nt) in read1 were kept. Cutadapt (v1.8.1) (Martin, 2011) was used to remove adaptor sequences as well as short and low quality reads with parameters -m 20–trim-n–max-n 0.7 –q 20. The remaining reads were aligned to genome and non-mouse reads were removed by bowtie2 (v2.2.4) (Langmead and Salzberg, 2012) and fastq_screen (v0.5.1.4). The clean reads were mapped to UCSC mm10 genome by STAR (v2.5.2b) (Dobin et al., 2013) and matched to Ensemble genes (GRCm38) (Zerbino et al., 2018) by featureCounts (subread-1.4.6-p1 command line) (Liao et al., 2014). We used software R and its package Seurat (v2.3.4) (Satija et al., 2015) for gene /cell filtering. Specifically, we removed genes that were detected in less than 10 cells. We removed cells that had less than 1000 transcribed genes or more than 20% reads from mitochondrial genes. In the end, 4413 out of 4610 cells and

19739 genes were used for further analysis. Gene expression was normalized using $\log_2(TPM)$. R package monocle (v2.6.0) (Trapnell et al., 2014) was used to analyze pseudotime trajectory. UMI count matrix and good quality cells were used as raw input to build cell trace. The function orderCells and BEAM were used to identify genes that contribute to cell branch. We used state1(enriched for *Mll1*$^{f/f}$ cells) as the root state and re-ordered cells for the linear trajectory. We treated pseudotime as the independent variable and the gene expression scaled by TPM as the dependent variable for LOESS regression in R software. The regression pattern showed the trend of gene expression over pseudotime. The R package clusterProfiler (v3.2.14) (Yu et al., 2012) was used to do KEGG pathway and GO enrichment analysis.

ChIP-seq analysis
All ChIP-seq data were mapped to the mm10 mouse genome by Bowtie2 (v2-2.2.4) (Langmead and Salzberg, 2012). Uniquely mapped reads (with XS tag in SAM file) were kept and transformed to BAM files for peak identification using MACS (Zhang et al., 2008). BAM files of mapping results were merged for the same sample using SAMtools and converted to BED format by using BEDTools (Quinlan and Hall, 2010). Peaks were called for each sample by using MACS (v 1.4.2) (Zhang et al., 2008) with parameters "-w -S -p 10^{-5}." The input signal was used as the control for peak calling. HOMER (http://homer.salk.edu) was used for peak annotation and for calculating the distribution of peaks in gene elements (annotatePeaks.pl). Gene annotation of peaks was performed by using GREAT (McLean et al., 2010) with default parameters. For ChIP-seq fold enrichment visualization in IGV (Thorvaldsdóttir et al., 2013), MACS2 subcommands bdgopt and bdgcmp were used to determine ratio of tags from experiment and input control after normalization. The heatmap plot of signals centered on peaks and gene promoters was generated by deepTools2 (v2.5.0) (Ramírez et al., 2016).

Time-series clustering analysis on gene expression
Short Time-series Expression Miner (STEM, v1.3.11) were used to model significant expression profiles in all stages of preimplantation development (e.g., Oocyte, 2-cell, 4-cell, 8-cell, ICM) (GSE66582) (Wu et al., 2016). STEM clustering was performed by setting of a maximum of model profiles of 25 and a maximum unit change between time points of 2. STEM first generated model expression profiles with five time points to fit the data. After assigning each transcript to a profile, enrichment of genes in each profile was calculated to determine profile significance. A false discovery rate of 0.05 was used to identify significant expression profiles. Our dataset was fitted to 8 model expression profiles by STEM that were significantly enriched for transcripts expressed longitudinally in each development stages. We then performed ggplot2 in R for the violin-plot to demonstrate log2 fold change (log2FC) of differential gene expression between early embryos and ESC in each STEM cluster. Similarly, we performed violin-plot of $\log_2$FC of differential gene expression between wild-type ESC and ESC after treatments in the same STEM cluster for our own RNA-seq data.

Functional enrichment and PCA analysis
GO (Gene Ontology) analysis was performed by Panther Classification System (http://pantherdb.org/). Statistical overrepresentation test was performed with "default settings." The annotation dataset is "GO biological process complete" and Test type is "Fisher's Exact with FDR multiple test correction" for reducing false discovery rate. GO term with FDR < 0.05 and fold change (FC) > 2 was considered as enriched. Principal component Analysis (PCA) was performed by ClustVis: a web tool for visualization clustering of multivariate data(BETA) (https://biit.cs.ut.ee/clustvis/). Unit variance scaling is applied to each dataset; principal components were calculated by SVD with imputation. Principal component 1 and 2 were shown as X and y axis.

Statistical analysis
Most experimental data were shown as Mean with SD. All statistical analyses were performed with GraphPad Prism 7 software. Details of numbers and types of replicates were indicated in each figure legend. All statistics with * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ were calculated by Mann-Whitney test (unpaired, nonparametric, two tailed), which does not require assumption of normal distribution. Information on numbers of biological replicates can be found in main text and figure legends.

DATA AND CODE AVAILABILITY

The accession number for RNA-seq and ChIP-seq data reported in this paper is: GEO: GSE107408 (for RNA-seq and ChIP-seq) The accession number for single cell RNA-seq data is GEO: GSE127943. The flow cytometry data is available at FlowRepository (http://flowrepository.org/) with identifier FR-FCM-Z29P. The accession numbers of the other deposited datasets are shown in the Key Resources Table.

Supplemental Information

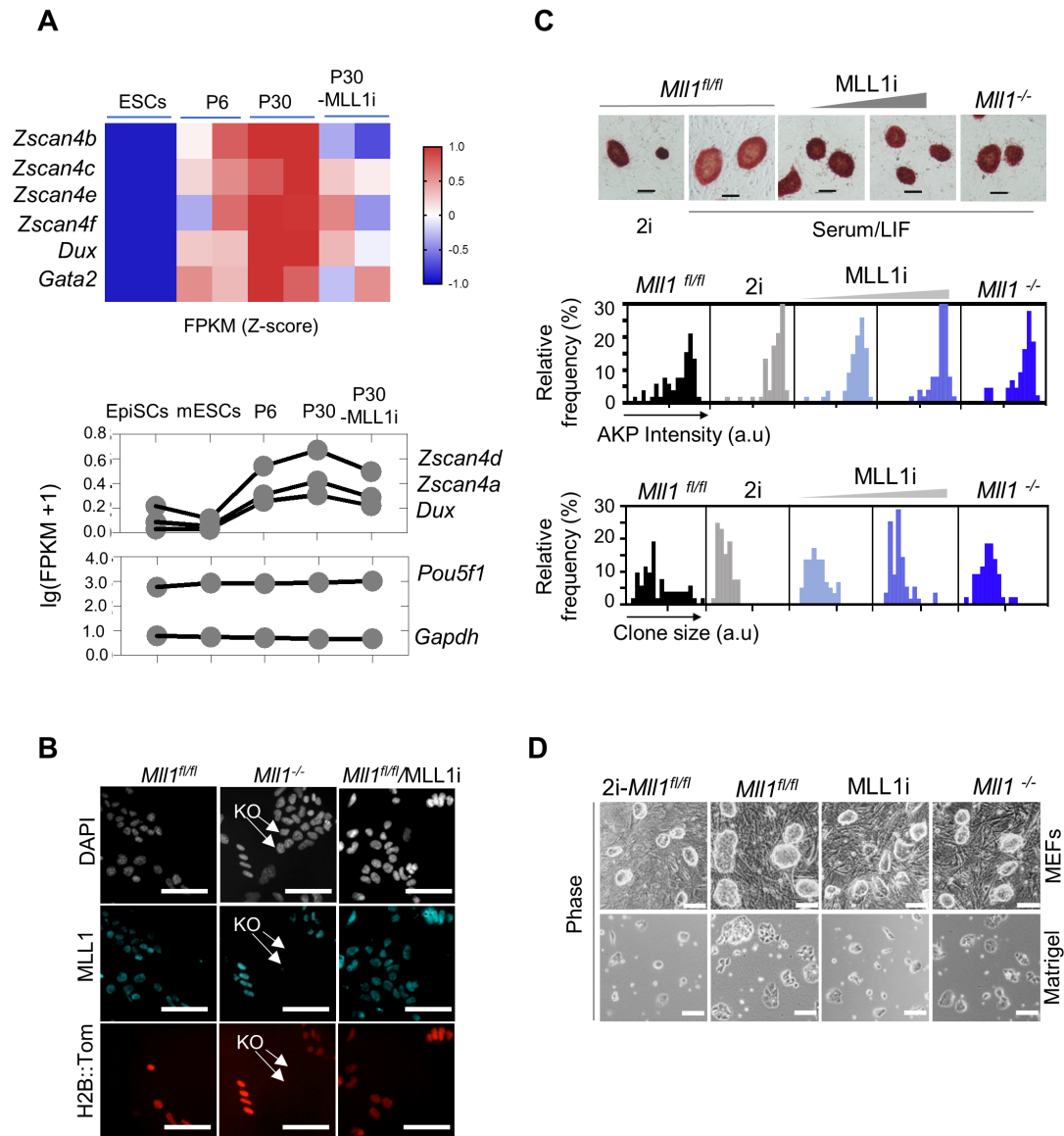**MLL1 Inhibition and Vitamin D**

Signaling Cooperate to Facilitate

the Expanded Pluripotency State

Hui Zhang, Le Tran Phuc Khoa, Fengbiao Mao, Hanshi Xu, Bo Zhou, Yu Han, Monique O'Leary, Asma Nusrat, Li Wang, Thomas L. Saunders, and Yali Dou

A

B

C

D

Figure S1. *Mll1* **deletion or MLL1 inhibition in ESCs. Related to Figure 1.**

(A) Heat map for expression of the selected genes during MLL1i-EpiSCs reprogramming. FPKM Z-score (top) or log (FPKM+1) (Bottom) were shown as indicated. Dataset was from GSE66112 (Zhang et al., 2016). P6 and P30 passage cells were incubated continuously with MM-401. P30-MLL1i, passage 30 cells with MM-401 withdrawal at passage 6. The experiment had biological duplicates.

(B) Immunofluorescence for MLL1 upon *Mll1* deletion. The cells were co-cultured with wild type ESCs that express a Tomato fluorescence gene under the control of the histone H2B promoter (H2B::Tom$^+$). No H2B::Tom$^-$ ESCs express MLL1 in the *Mll1$^{-/-}$* co-culture (middle column), supporting complete deletion of *Mll1* in the knockout cells. Scale bar, 50 µm.

(C) Top, representative images of AKP-stained ESC clones that were cultured under either 2i or serum/LIF condition as indicated on bottom. Scale bars, 50 µm. Quantification of AKP intensity (middle) and clone size (bottom) were analyzed by ImageJ. Data were shown as relative distribution frequency. X-axis, AKP intensity or clone size in arbitrary unit (a.u). At least 40 clones per condition were used for analysis. The experiment was repeated three times.

(D) Representative phase-contrast images for cells, indicated on top, that were cultured on either MEF feeders (top) or Matrigel (bottom). Scale bar, 100 µm.
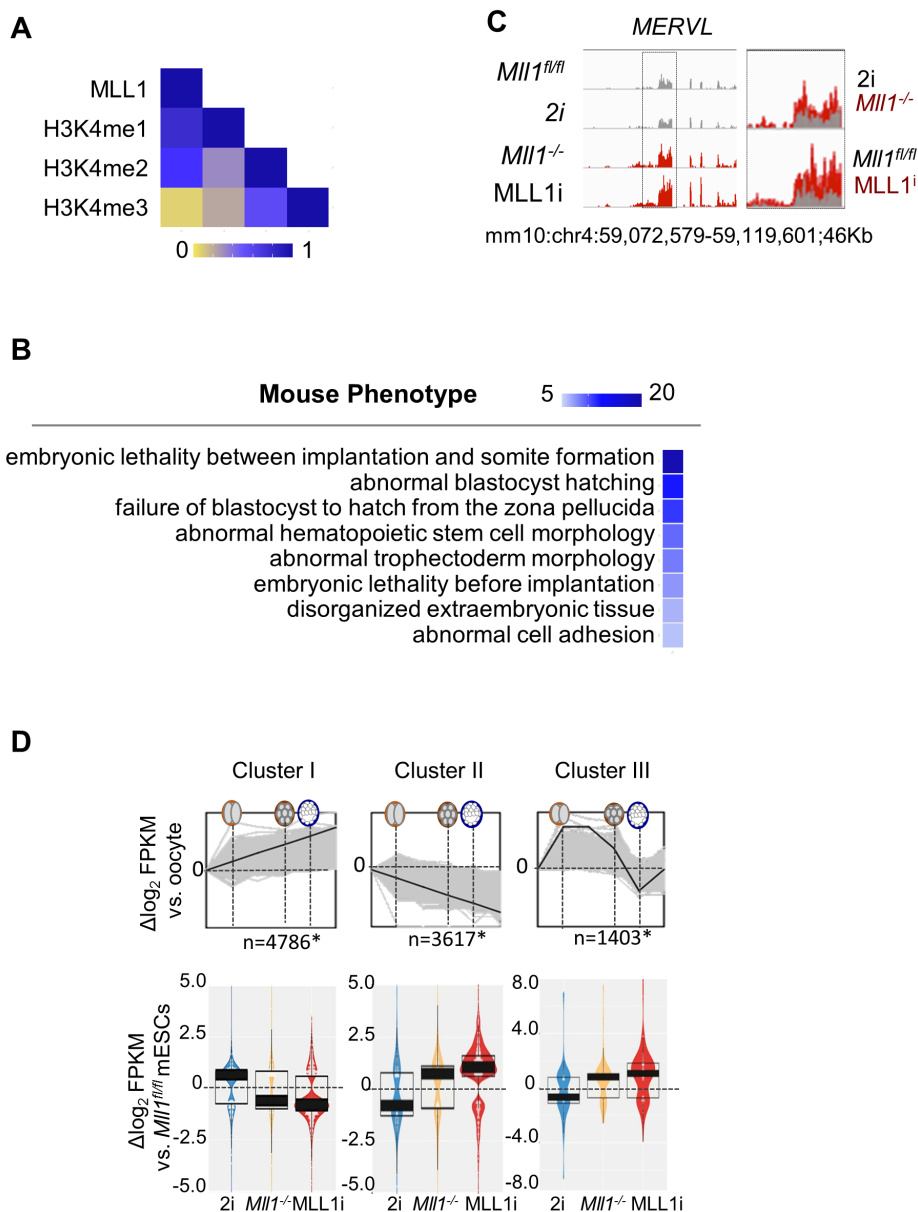
**Figure S2. MLL1i or *Mll1*KO alter ESC transcriptome. Related to Figure 1.**

(A) Heat map for Pearson correlation coefficient of ChIP-seq signals for MLL1, H3K4me1, H3K4me2 and H3K4me3. The experiment had two biological duplicates.

(B) Pathway analysis for MLL1 direct targets using Mouse Genome Informatics (MGI). Heap map showed –log (Bionomial *p*).

(C) The Integrative Genomics Viewer (IGV) view of RNA-seq reads for MERVL in *Mll1^f/f^*, 2i, *Mll1^-/-^* or MLL1i cells as indicated.
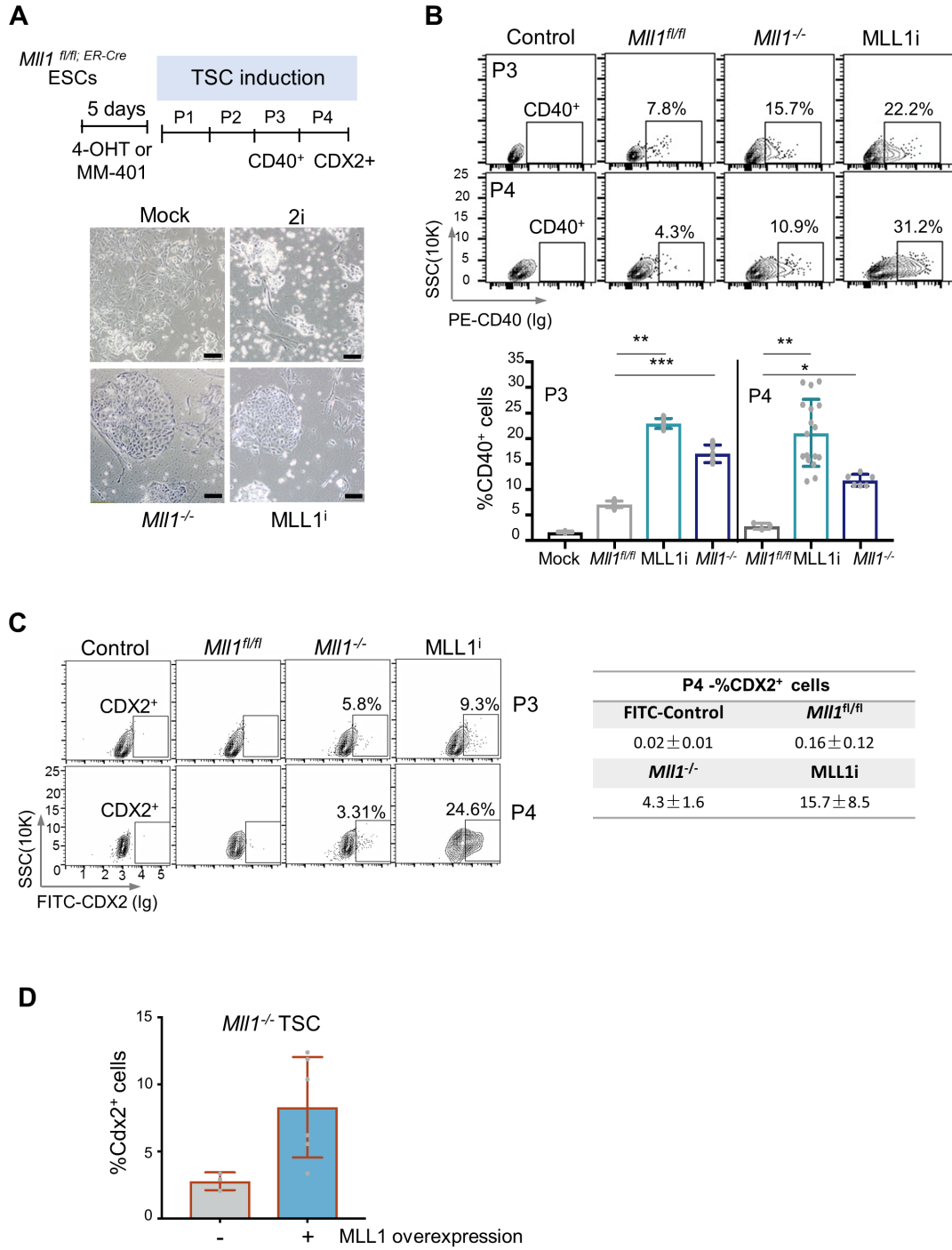
(D) Top, three major dynamic gene expression profiles from oocyte to blastocyst in mouse (dataset from GEO GSE66582) (Wu et al., 2016). They included monotonic up regulation (I), monotonic down regulation (II) as well as variable trajectory with highest expression at the 2-8-cell stages (III). *, FDR<1e-57, calculated by STEM (v1.3.11). Bottom, violin plot showed relative expression changes for genes (|FC|≥1.5, FDR≤0.05, relative to ESCs) in the cluster on top upon 2i, *Mll1* deletion or MLL1i. Dataset was from Wu, J. et al. 2016.

A

B
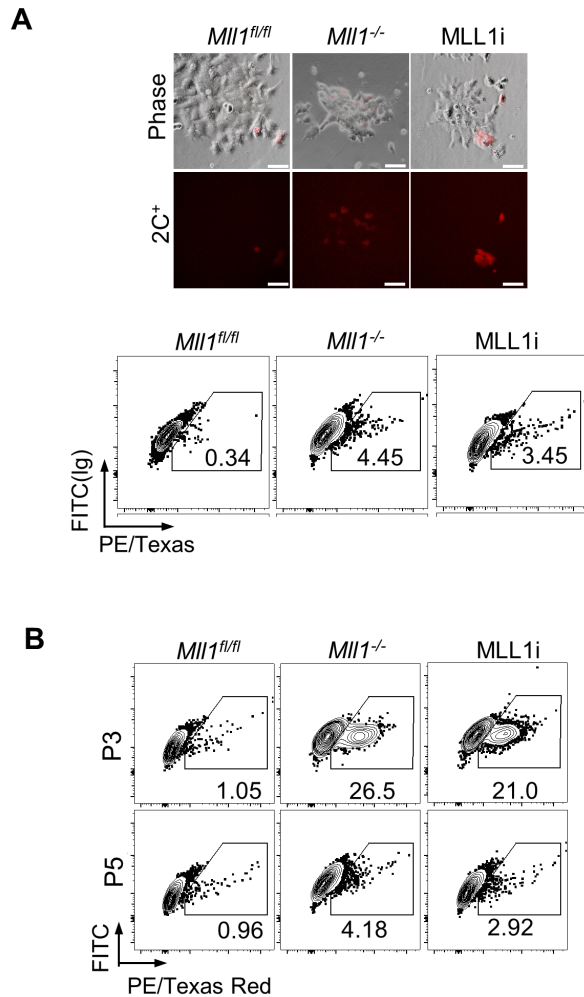
Figure S3. MLL1i or *Mll1*KO promote the 2C⁺ state *in vitro*. Related to Figure 3.

(A) Top, representative images of *Mll1$^{fl/fl}$*; *Mll1$^{-/-}$*, and MLL1i cells with the MERVL-2C-tomato (red) reporter after 5-day treatment. Top panels, overlap of phase-contrast and the fluorescence images from bottom panel. Scale bars, 50 µm. Bottom, representative FACS analysis of 2C⁺ *Mll1$^{fl/fl}$*; *Mll1$^{-/-}$*, or MLL1i cells after 5 day treatment. Percent of 2C⁺ cells were indicated. This experiment was repeated for four times.

(B) FACS analyses of sorted 2C⁺ cells after either three (top) or five (bottom) continuous passaging *in vitro*. This experiment was repeated three times.

A

Mll1 ^fl/fl; ER-Cre^ ESCs

B

C

P4 -%CDX2⁺ cells

| FITC-Control	*Mll1*^fl/fl^
0.02±0.01	0.16±0.12
Mll1^-/-^	MLL1i
4.3±1.6	15.7±8.5

D

Figure S4. MLL1i or *Mll1*KO ESCs have expanded differentiation potential towards TSCs *in vitro*. Related to Figure 3.

(A) Top, Schematics for TSC induction. 4-OHT and MM-401 were withdrawn after switching to TSC culture condition. Bottom, representative images for P2 cells after TSC induction. Scale bar, 200 μm. This experiment was repeated five times.

(B) FACS analyses of TSC marker CD40 for P3 (top) or P4 (bottom) cells. Gating selection was indicated in control. Quantification of CD40$^+$ cells from 6-17 replicates from two independent experiments was shown at bottom. Average % CD40$^+$ cells was presented and standard deviation was presented as the error bar. Mann-Whitney t-test (two-tailed) was used. *, $p<0.05$; **, $p<0.01$ and ***, $p<0.001$.

(C) FACS analyses of TSC marker CDX2 for P3 (top) or P4 (bottom) cells after TSC induction. Gating selection was indicated in control. Average % CDX2$^+$ cells and standard deviation from twp independent experiments were summarized in the table on right. The same dataset were also used in Figure 6C.

(D) Quantification of %CDX2$^+$ cells in the $Mll1^{-/-}$ cells with or without exogenous MLL1 expression as indicated on bottom. Average % CDX2$^+$ cells from 3-6 replicates in two independent experiments (as indicated by grey dots) was presented and standard deviation was presented as error bars.
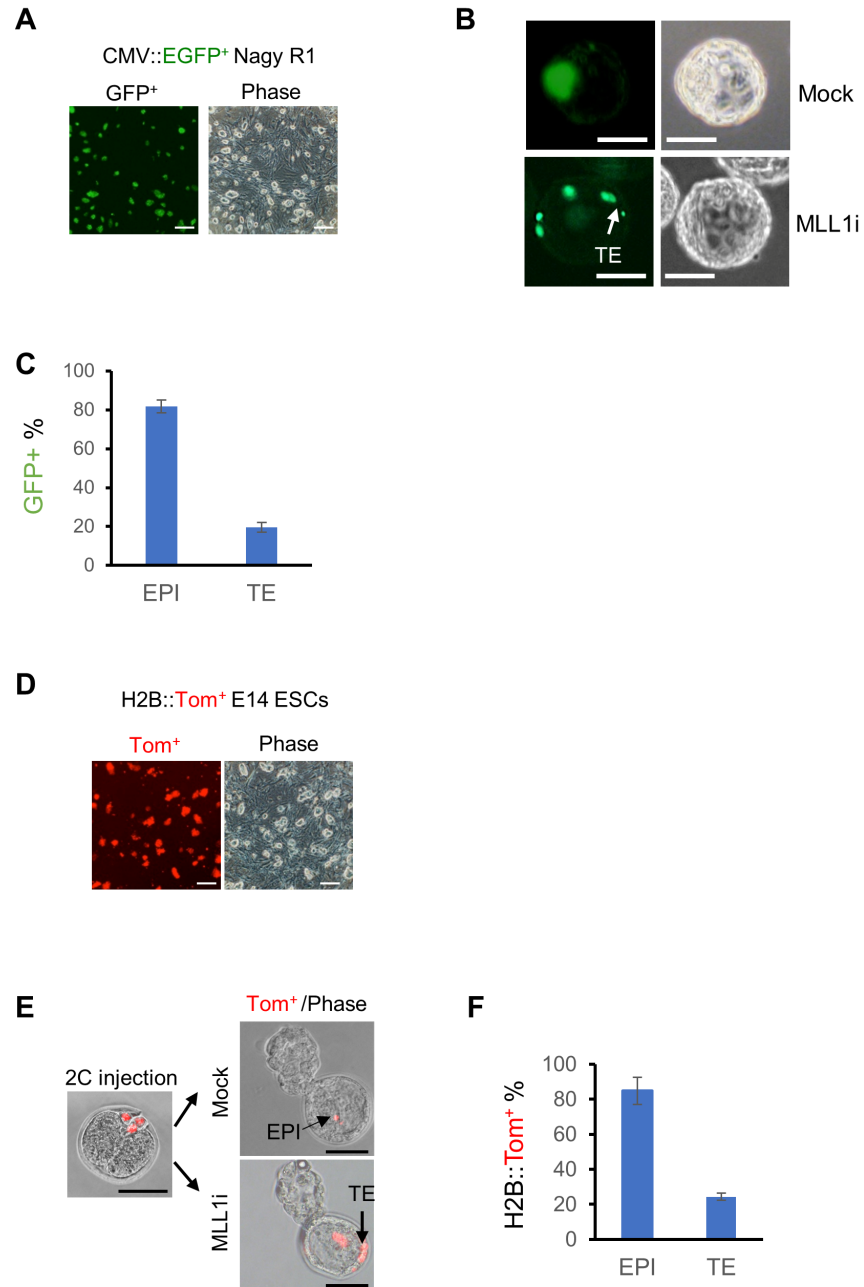
Figure S5. MLL1i promotes the EPSC state *in vitro*. Related to Figures 3 and 4.

(A) The EGFP$^+$ Nagy R1 ESCs were used for testing EPSC *in vivo*. EGFP was constitutively expressed under the control of the CMV promoter. Scale bar, 200 μm.

(B) Representative chimeric blastomere (at 18-24hrs after injecting 3.5d embryos with mock or MLL1i treated EGFP$^+$ Nagy R1 ESCs. Scale bar, 50 μm. TE, trophectoderm.

(C) Quantification of EPI and TE contribution of EGFP$^+$ Nagy R1 ESCs treated with MLL1i. Average results from three independent injection (total n=42) was presented and standard

deviation was shown as error bars. EPI, integration into epiblast. TE, integration into trophectoderm. EPI and TE counts were not mutually exclusive.

(D) The H2B::Tom$^+$ E14Tg2a cells were used for testing EPSC *in vivo*. Scale bar, 200 μm.

(E) Representative chimeric blastomere 72hr after injecting 2-cell embryos with mock or MLL1i treated H2B::Tom$^+$ E14Tg2a ESCs. Scale bar, 50 μm.

(F) Quantification of EPI and TE contribution of H2B::Tom$^+$ E14Tg2a ESCs treated with MLL1i. Average result from three independent injections (n=44) was presented and standard deviation was shown as error bars. EPI and TE counts were not mutually exclusive.
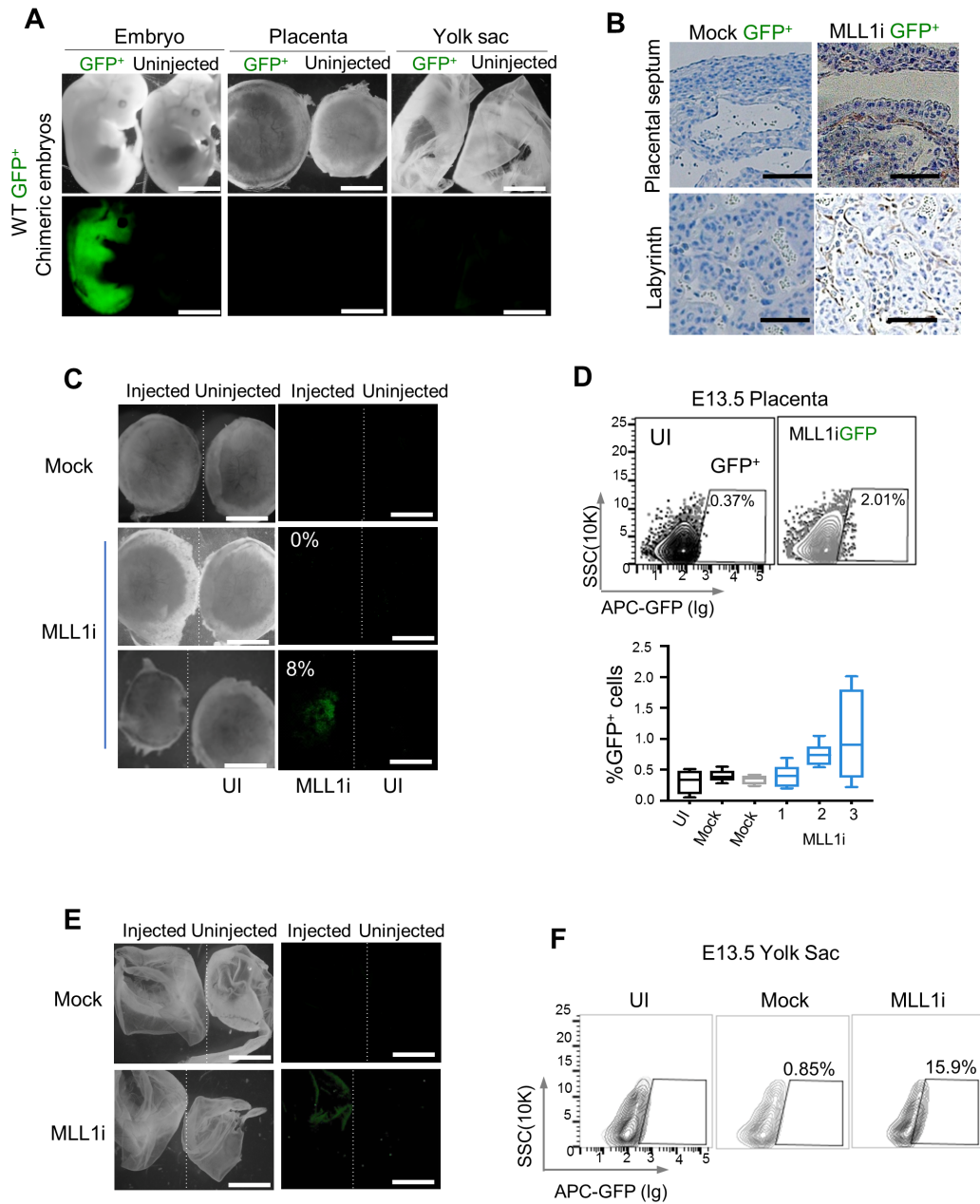
Figure S6. MLL1i promotes the EPSC state *in vivo*. Related to Figure 4.

(A) Representative images of E13.5 chimeric embryos, placenta or yolk sac derived from mock treated ESCs. An un-injected control embryo was shown as the negative control. Scale bar, 0.25 cm.

(B) Immunohistochemistry for GFP in placenta septum and labyrinth structures in chimeric embryos derived from mock or MLL1i treated EGFP⁺ ESCs. Scale bars, 100 μm.

(C) Representative images of E13.5 placenta of chimeric embryos derived from mock or MLL1i treated ESCs. An un-injected E13.5 control placenta was shown side by side as the negative control. GFP$^+$ signals in the imaging field were analyzed by Image J and %GFP signal was normalized to the size of corresponding placenta. Scale bar: 0.25cm.

(D) FACS analyses of GFP$^+$ cells in placenta. Top, representative flow cytometry analysis of GFP$^+$ cells. The un-injected control embryo was used as the negative control for gating selection. Bottom, summary of FACS analyses on at least three placentas. Standard deviation was shown as error bars.

(E) Representative images of E13.5 yolk sac of chimeric embryos derived from mock or MLL1i treated ESCs. An un-injected E13.5 yolk sac was shown side by side as the negative control. Scale bar: 0.25cm.

(F) FACS analyses of GFP$^+$ cells in yolk sac shown in E. The un-injected control embryo was used as the negative control for gate selection. This experiment was repeated three times.
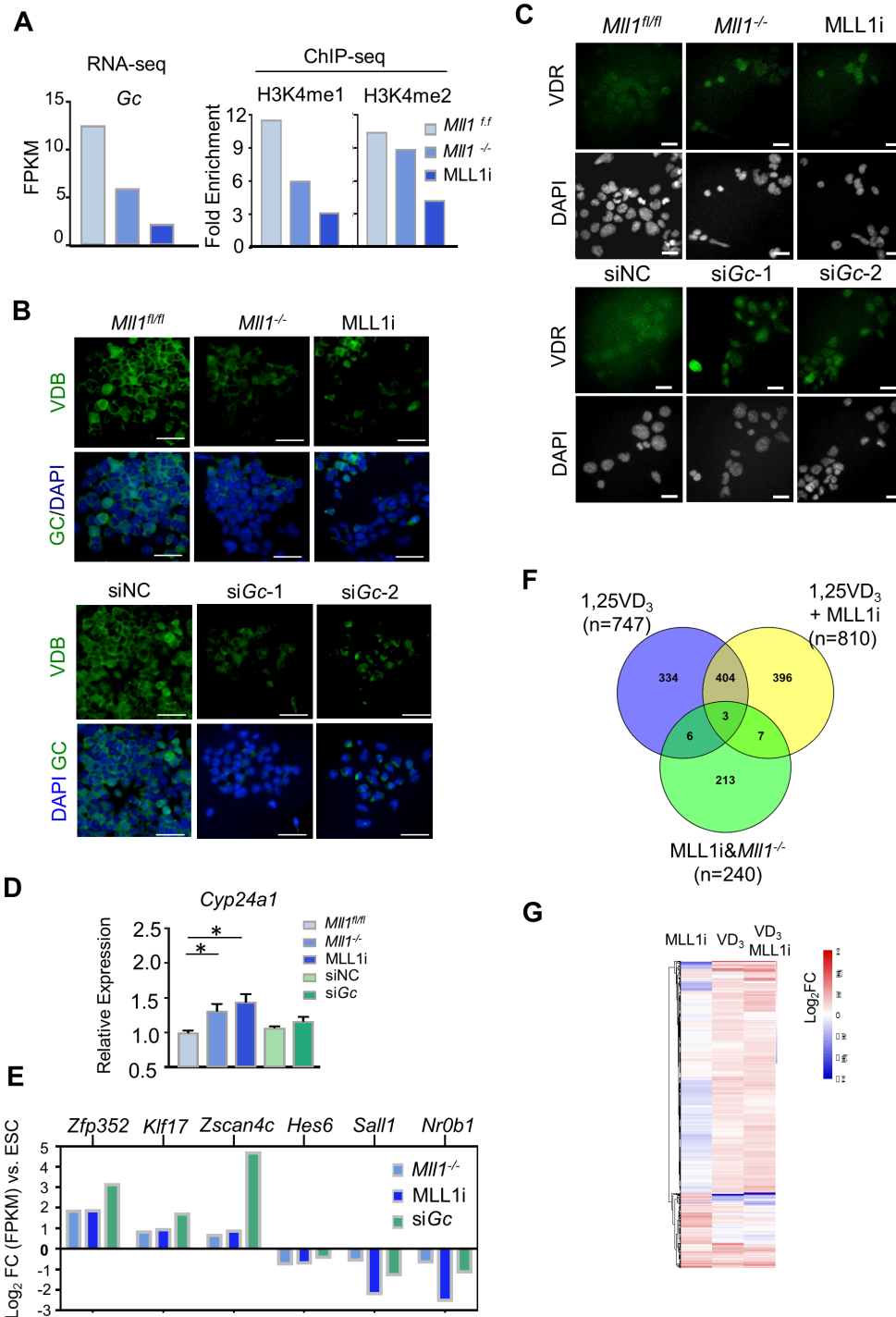
Figure S7. MLL1 regulates *Gc* and Vdr signaling in ESCs. Related to Figures 5 and 7, and Table S4.

(A) RNA-seq data for *Gc* and ChIP-seq data for H3K4me1 and H3K4me2 at *Gc* locus. ChIP-seq data was presented as fold enrichment over input. Each sample has biological duplicates.

(B) Immunofluorescence for VDB (encoded by *Gc*) in *Mll1* ^{f/f}, *Mll1*^{-/-} and MLL1i cells (top) or in control and *Gc*-siRNA treated cells (bottom). Scale bar, 50 µm.

(C) Immunofluorescence for VDR in) in *Mll1* ^{f/f}, *Mll1*^{-/-} and MLL1i cells (top) or in control and *Gc*-siRNA treated cells (bottom). Scale bar, 50 µm.

(D) Relative expression of *Cyp24a1* in MLL1i, *Mll1*^{-/-} or *Gc* knockdown cells versus that of *Mll1*^{f/f} cells, which was arbitrarily set as 1. Average results from three replicates were presented. Standard deviation was presented as error bars. *, $p<0.05$.

(E) Relative fold change of selective 2-cell specific genes and lineage specifiers shown in Figure. 1D. RNA-seq data was normalized and presented as $\log_2$ FPKM change against mock treated ESCs.

(F) Venn diagram for up regulated genes in three different treatments in Figure 7B that have higher expression in preimplantation embryos as compared to ESCs (GEO GSE66582) (Wu et al., 2016).

(G) Heat map for the genes in Figure S7F. $\log_2$ fold change of gene expression (average from two biological duplicates), relative to ESCs, was shown as indicated by heat map key.