LETTERS

The ground state of embryonic stem cell self-renewal

Qi-Long Ying[1], Jason Wray[2], Jennifer Nichols[2], Laura Batlle-Morera[2], Bradley Doble[3], James Woodgett[4], Philip Cohen[5] & Austin Smith[2]

In the three decades since pluripotent mouse embryonic stem (ES) cells were first described[1,2] they have been derived and maintained by using various empirical combinations of feeder cells, conditioned media, cytokines, growth factors, hormones, fetal calf serum, and serum extracts[1–7]. Consequently ES-cell self-renewal is generally considered to be dependent on multifactorial stimulation of dedicated transcriptional circuitries, pre-eminent among which is the activation of STAT3 by cytokines (ref. 8). Here we show, however, that extrinsic stimuli are dispensable for the derivation, propagation and pluripotency of ES cells. Self-renewal is enabled by the elimination of differentiation-inducing signalling from mitogen-activated protein kinase. Additional inhibition of glycogen synthase kinase 3 consolidates biosynthetic capacity and suppresses residual differentiation. Complete bypass of cytokine signalling is confirmed by isolating ES cells genetically devoid of STAT3. These findings reveal that ES cells have an innate programme for self-replication that does not require extrinsic instruction. This property may account for their latent tumorigenicity. The delineation of minimal requirements for self-renewal now provides a defined platform for the precise description and dissection of the pluripotent state.

Mouse ES cells exist in the artificial milieu of cell culture. They are derived and maintained by using a combination of the cytokine leukaemia inhibitory factor (LIF) to activate STAT3 and either serum or bone morphogenetic protein (BMP) to induce inhibitor-of-differentiation proteins[5]. Their differentiation involves autoinductive stimulation of the mitogen-activated protein kinase (ERK1/2) pathway by fibroblast growth factor-4 (FGF4)[9,10]. However, neither LIF nor serum/BMP block the activation of ERK (Supplementary Information and ref. 5). We proposed that the LIF and serum/BMP signals act downstream of phospho-ERK to block ES-cell commitment. To test this idea we used selective small-molecule inhibitors SU5402 (ref. 11) and PD184352 (ref. 12) to inhibit FGF receptor tyrosine kinases and the ERK cascade, respectively.

We found that, in combination with LIF, either inhibitor replaces the requirement for serum/BMP and supports robust long-term ES-cell propagation (Supplementary Information). Lineage commitment does not occur despite a reduced expression of inhibitor-of-differentiation proteins. In contrast, ES cells plated without LIF in either PD184352 or SU5402 progressively degenerate and cannot be maintained even though differentiation is suppressed. To reduce off-target side effects we tried low doses of PD184352 and SU5402 together (PS). In PS we find that undifferentiated ES cells expand through multiple passages (Fig. 1a, b). Differentiation is constrained, although occasional neural rosettes emerge. This result, observed with several independent ES cell lines, suggests that the minimal requirements for ES-cell self-renewal may be to deflect commitment signals emanating from FGF receptor and ERK signalling. However, apoptosis is relatively high in PS, especially immediately after passage, and cells survive poorly at clonal density, which is indicative of collateral compromise to cell growth and viability.

ES-cell propagation has been reported to be enhanced by an indirubin entity, 6-bromo-indirubin-3′-oxime (BIO), that inhibits glycogen synthase kinase-3 (GSK3)[4]. However, indirubins are not highly selective and cross-react with cyclin-dependent kinases and other kinases[13,14]. We found reduced viability of ES cells in BIO with or without PS. Nevertheless we speculated that relief of GSK3-mediated negative regulation of biosynthetic pathways might restore growth to ES cells cultured in PS. We therefore used a more selective inhibitor, CHIR99021 (ref. 14,15). Alone, CHIR99021 enhances survival at low cell density but also induces non-neural differentiation. At higher densities some colonies remain morphologically undifferentiated but are progressively overcome by differentiation on passaging (Fig. 1c). Single blockade of GSK3 therefore has pleiotropic effects, promoting non-neural differentiation, suppressing neural differentiation and enhancing growth capacity. Crucially, however, in a combination of all three inhibitors (3i) the differentiation blocking effect of PS is dominant, resulting in a highly efficient expansion of undifferentiated colonies, even at a low cell density. Multiple ES-cell lines tested all expand continuously for many weeks in 3i. They express Oct4, Nanog and Rex1 with minimal levels of lineage commitment markers, Sox1 or brachyury (Fig. 1d, e). In 3i, ES cells expand with a doubling rate comparable to that in LIF plus serum/BMP (Supplementary Information) with the proportion of Oct4–green fluorescent protein (GFP)-positive undifferentiated cells remaining over 90%. As a rigorous test of the sufficiency of 3i to sustain ES-cell self-renewal, we examined the clonogenicity of isolated cells. After single-cell deposition, undifferentiated Oct4-positive colonies develop at higher frequency than in LIF and serum or BMP (Fig. 1f, g).

The B27 supplement used in serum-free culture contains defined additives, in particular antioxidants and free-radical scavengers. We found that ES cells could be propagated in bulk culture in 3i medium prepared with N2 supplement only, but they did not survive at clonal density. However, in physiological oxygen (5% O_2) clonal propagation was obtained without B27 (Fig. 1g). This excludes an instructive contribution from B27 components to ES-cell self-renewal, while highlighting the damage potential of non-physiological oxygen levels. When insulin was omitted we found ES cells to be more sensitive to FGF receptor (FGFR) and MAP kinase/ERK kinase (MEK) inhibitors. We therefore decreased their concentrations. In these conditions, with only transferrin and albumin additives, ES cells expanded, even from single cells. They remained mostly undifferentiated over four weeks of continuous culture (Fig. 1h), although after the first passage the propagation rate declined steadily. We conclude that insulin promotes long-term growth capability but does not dictate the fate choice between self-renewal and lineage commitment.

[1]Center for Stem Cell and Regenerative Medicine, Department of Cell and Neurobiology, Keck School of Medicine, University of Southern California, 1501 San Pablo Street, ZNI 529, Los Angeles, California 90033, USA. [2]Wellcome Trust Centre for Stem Cell Research, University of Cambridge, Tennis Court Road, Cambridge CB2 1QR, UK. [3]McMaster Stem Cell and Cancer Research Institute, McMaster University, 1200 Main Street West, Hamilton, Ontario L8N 3Z5, Canada. [4]Samuel Lunenfeld Research Institute, Mount Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5, Canada. [5]Division of Signal Transduction Therapy and MRC Protein Phosphorylation Unit, University of Dundee, Dundee DD1 5EH, UK.

Finally, we used recombinant albumin to eliminate serum-derived contaminants. In combination with transferrin and insulin, this supported both bulk passaging and clonal propagation (Fig. 1g).

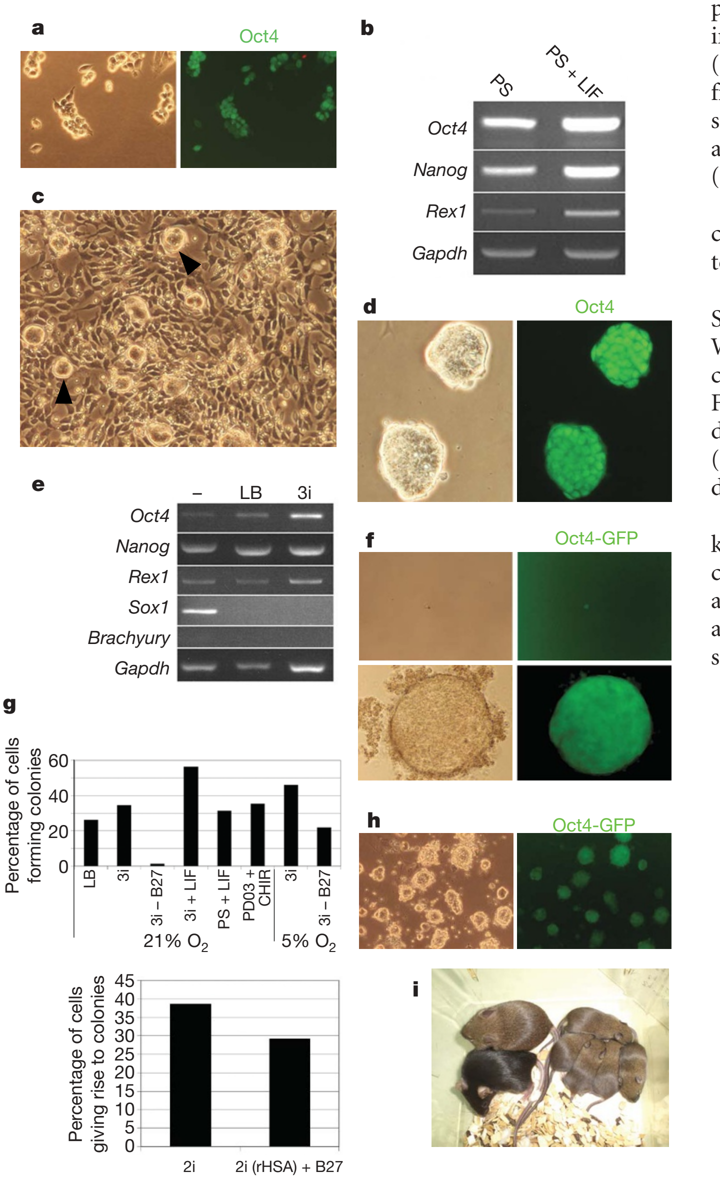


Figure 1 | Three inhibitors (3i) support robust self-renewal and *de novo* derivation of pluripotent ES cells. **a**, Immunostaining of E14Tg2a ES cells with Oct4 after four passages in N2B27 plus PD184352 and SU5402 (PS). **b**, RT–PCR analysis of marker expression in ES cells in N2B27 containing PS with or without LIF. *Gapdh*, gene encoding glyceraldehyde-3-phosphate dehydrogenase. **c**, Low-magnification phase-contrast image of ES cells passaged in N2B27 plus CHIR99021 showing a mixture of differentiated cells with compact undifferentiated colonies (arrowheads). **d**, Immunostaining with Oct4 after several passages in N2B27 plus 3i, showing compact colony morphology. **e**, RT–PCR analysis of marker expression in ES cells cultured in N2B27 alone (−) or with LIF and BMP4 (LB) or 3i. **f**, Phase and fluorescence images of expansion from a single Oct4GiP ES cell in 3i. **g**, Cloning efficiencies of E14Tg2a ES cells after single-cell deposition in the indicated conditions (top), and in CHIR99021 plus PD0325901 (2i, see Fig. 2) with or without B27, or with the replacement of serum albumin with recombinant albumin (rHSA) (bottom; experiment performed in 5% O_2). **h**, Oct4GIP ES cells cultured for five passages (total 28 days) in basal medium supplemented with transferrin and BSA only plus 3 μM CHIR99021, 0.5 μM PD184352 and 1 μM SU5402. **i**, Chimaera and germline offspring produced from CBA ES cells derived in 3i. Chimaera showing extensive contribution of CBA (agouti coat colour) ES cells mated with C57BL/6 (black) produced agouti pups, indicating the transmission of the CBA genome.

To eliminate the possibility that self-renewal in 3i might reflect pre-adaptation to specific culture conditions in our laboratory, we investigated the derivation of ES cells from mouse embryos. ES cells were readily derived from blastocysts of the permissive 129 strain plated directly into 3i on gelatin-coated plastic. Expanded lines injected into blastocysts gave chimaeras and germline transmission (Supplementary Information). ES cell lines were also established from the CBA strain, which is refractory to ES-cell production under standard conditions[16]. Two of these lines were injected into morulae and both yielded high-grade chimaeras and germline transmission (Fig. 1i).

Taken together, the above findings demonstrate that 3i liberates ES cells from requirements for exogenous factors without compromise to developmental potency.

To confirm that blockade of FGF signalling is the critical target of SU5402 we substituted an alternative inhibitor, PD173074 (ref. 17). We found that this could substitute for SU5402 in 3i at 40-fold lower concentrations, which is consistent with its higher affinity for the FGF receptor (Fig. 2a). We then examined *fgf4*-null ES cells[18] and determined that they can expand continuously in CHIR99021 alone (Fig. 2b), providing genetic validation of the significance of autoinductive FGF4.

FGF4 activates the phosphatidylinositol-3-OH kinase/protein kinase B (PKB) and the Ras–MEK–ERK intracellular signalling cascades (Fig. 2c, d). Phosphorylation and activation of PKB is not appreciably altered by the 3i inhibitors. PD184352 or SU5402 applied alone at the low doses used in 3i cause only modest decreases in steady-state phospho-ERK. However, the combination of both

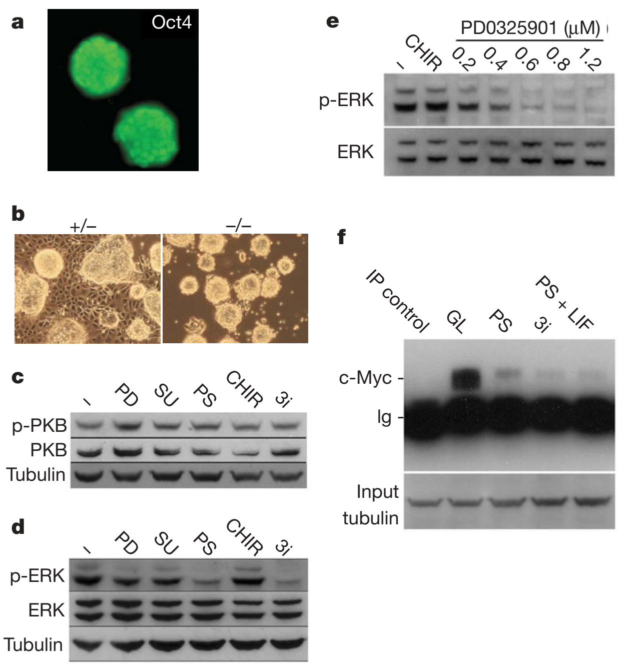


Figure 2 | Effects of 3i components on intracellular signalling cascades. **a**, E14Tg2a ES cells remain undifferentiated and Oct4-positive in alternative 3i with SU5402 replaced by PD173074. **b**, *fgf4*-null ES cells expand without differentiation in N2B27 plus CHIR99021 only, without a requirement for FGFR/MEK inhibition. **c**, **d**, Immunoblot analyses of steady-state levels of phospho(Ser 473)-PKB (p-PKB) (**c**) and phospho(Thr 202, Tyr 204)-ERK (p-ERK) (**d**) in ES cells after 24 h in N2B27 alone (−), plus 0.8 μM PD184352 (PD), 2 μM SU5402 (SU), 3 μM CHIR99021, PS or 3i. **e**, Immunoblot analyses of phospho(Thr 202, Tyr 204)-ERK levels in ES cells after 24 h in N2B27 alone (−), plus 3 μM CHIR99021 (CHIR) or 3 μM CHIR99021 plus PD0325901 at the indicated concentrations. **f**, c-Myc protein in ES cells assayed by sequential immunoprecipitation (IP) and immunoblotting after 24 h in serum plus LIF (GL), PS, 3i, or PS plus LIF. IP control is the GL sample immunoprecipitated with anti-tubulin. Input samples were subjected to SDS PAGE and blotted for tubulin to control for loading.

inhibitors greatly decreases phospho-ERK levels. CHIR99021 does not modulate phospho-ERK (Fig. 2e). We tested *erk2*-null ES cells[19] and found that these can be maintained at high density with CHIR99021 only, although optimal propagation requires supplementation with PD184352; this is consistent with maintained activity of phospho-ERK1 in these mutants. The central role of the ERK cascade was confirmed by using a structurally related, more potent but equally selective MEK inhibitor, PD0325901 (ref. 15), to achieve greater suppression of ERK activation without side effects. This is sufficient to sustain efficient ES-cell self-renewal in combination with CHIR99021 only (Figs 1g and 2e).

An unwarranted side effect of suppressing phospho-ERK is to depress *myc* messenger RNA and Myc protein levels (Fig. 2f and Supplementary Information). Upregulation of c-Myc has been suggested to mediate ES-cell self-renewal downstream of LIF and of BIO[20]. However, the low c-Myc levels in cultures in PS are not increased by CHIR99021 or LIF (Fig. 2f). Therefore elevated c-Myc is not necessary for ES-cell propagation, although some requirement for basal Myc activity is not excluded.

STAT3 signalling is central to previous models of ES-cell self-renewal[8,21] and has also been implicated in effects of BIO[20,22,23]. In 3i, however, we do not detect activation of STAT3 or induction of its target SOCS3 (Supplementary Information). To test definitively whether STAT3 is dispensable for ES-cell self-renewal, embryos from intercrosses of *Stat3* heterozygous mice were cultured in 3i. Homozygous mutant ES cells were established (Fig. 3a). *Stat3*-null cells are morphologically indistinguishable from wild-type ES cells. They express Oct4 and Nanog, and initiate multilineage commitment in embryoid bodies (Fig. 3b–d). They show no induction of SOCS3 when stimulated with LIF (Fig. 3e). When transferred to LIF and serum, $stat3^{-/-}$ cells differentiate rapidly, confirming their incompetence to respond to LIF (Fig. 3f). We conclude that the otherwise absolute requirement for STAT3 in the derivation and self-renewal of mouse ES cells is rendered dispensable by 3i.

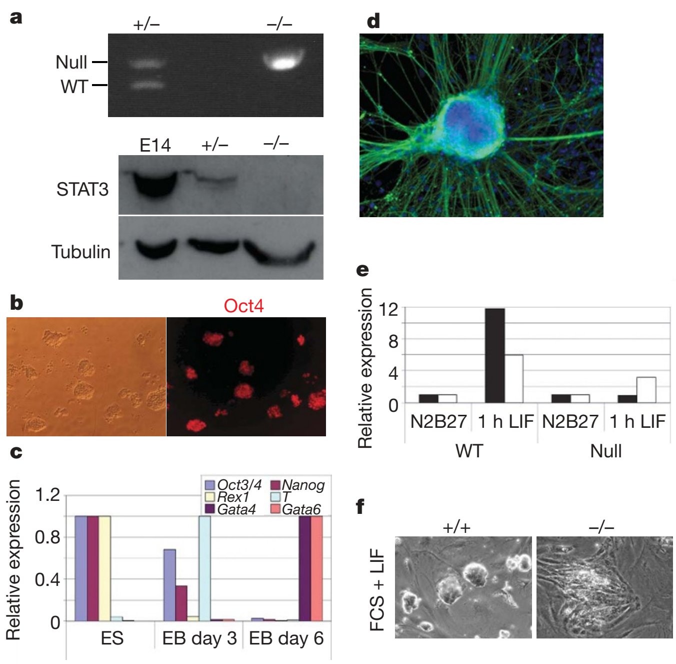


Figure 3 | ES-cell propagation in 3i does not involve STAT3. a, Top: genomic PCR for null and wild-type (WT) *Stat3* alleles in heterozygous and *Stat3* homozygous null ES cells. Bottom: immunoblot analysis of heterozygous and *Stat3* homozygous null ES cells. E14, E74Tg2a ES cells. **b**, Oct4 immunostaining of *Stat3*-null ES cells. **c**, Quantitative RT–PCR analysis of undifferentiated *Stat3*-null ES cells and derivative embryoid bodies (EB) at days 3 and 6. **d**, *Stat3*-null ES cells generate morphologically differentiated cells expressing the neuronal marker βIII-tubulin (TuJ1 immunoreactive). **e**, Quantitative RT–PCR analysis of *Socs3* (Stat3 target gene; filled columns) and *Egr1* (ERK target gene; open columns) expression in undifferentiated *Stat3* wild-type and null ES cells grown in N2B27 alone or stimulated with LIF for 1 h. **f**, *Stat3*-null MF1 ES cells differentiate in the presence of LIF and feeders, in contrast with wild-type MF1 ES cells, which remain undifferentiated.

CHIR99021 induces a decrease in phosphorylation of β-catenin (Supplementary Information) and activation of the T-cell factor (TCF)-responsive TOPFlash reporter (Fig. 4a), simulating canonical Wnt signalling. We investigated whether Wnt could replicate the effect of CHIR99021. Recombinant Wnt3a alone induced non-neural differentiation, as seen with CHIR99021 only. This effect was suppressed by PS and at high concentrations (100 ng ml^{-1}) Wnt3a seemed to eliminate residual neural differentiation and thereby improved ES-cell propagation. However, expansion in PS plus Wnt3a did not match that obtained in 3i. We introduced into ES cells dominant-negative ΔNhLef1, which lacks the β-catenin-binding domain and suppresses TCF-mediated transcriptional activation. As expected, ES cells stably expressing ΔNhLEF1 showed reduced TOPFlash activity (Fig. 4a). Nonetheless they readily formed undifferentiated colonies in 3i. A competitive self-renewal assay was performed after treatment with Cre to excise the floxed ΔNhLEF1 and simultaneously activate GFP. Equivalent numbers of ΔNhLEF1-expressing and revertant GFP-expressing cells were propagated as mixed cultures for four passages. In serum plus LIF the GFP-positive and GFP-negative populations remained equivalent. In 3i the GFP-negative ΔNhLEF1-expressing cells became marginally predominant (Fig. 4b). Decreasing TCF activation therefore does not impede ES-cell self-renewal. Increased β-catenin levels might also enhance cell adhesion. However, E-cadherin-null ES cells that lack adhesion junctions remain undifferentiated and proliferate as rapidly in 3i as in LIF plus serum (Supplementary Information).

To confirm that the effect of CHIR99021 is mediated through the inhibition of GSK3, we interrogated ES cells in which both GSK3α and GSK3β had been deleted[24]. These DKO cells are profoundly deficient in neural differentiation. They can be passaged two or three times in non-supplemented medium but succumb to progressive non-neural differentiation. This short-lived propagation is similar to that of wild-type ES cells cultured in CHIR99021 only (compare Fig. 4c with Fig. 2a). Addition of PS or PD0325901 eliminates differentiation and allows continuous passaging (Fig. 4c). However, expansion is slower than in wild-type cells in 3i. LIF restores normal population doubling (Fig. 4c), but CHIR99021 has no beneficial effect. This confirms that the effect of CHIR99021 is mediated through GSK3 and that LIF operates through a parallel STAT3 (refs 8, 21) pathway independent of GSK3 inhibition. DKO cells show constitutive TOPFlash activation[24], 50-fold higher than CHIR99021-treated wild-type cells (Supplementary Information). This tonic β-catenin/TCF activity, with upregulation of targets such as brachyury and cdx1 (ref. 24), probably underlies their compromised propagation.

ES cells constitutively expressing elevated levels of Nanog are capable of sustained self-renewal in N2B27 alone but expand poorly at clonal density unless LIF is also added[5]. They form only a few small colonies at low density in PS but generate abundant undifferentiated colonies in 3i (Fig. 4e). The key effect of CHIR99021 therefore does not involve the induction of Nanog. Because Nanog-overexpressing ES cells are independently blocked in differentiation, this result further suggests that the contribution of GSK3 inhibition extends beyond limiting differentiation. To probe this further, we evaluated whether CHIR99021 could rescue ES cells subjected to a more profound blockade of phospho-ERK. A higher dose of PD0325901 (2 or 3 μM) almost entirely eliminates phospho-ERK and causes growth arrest and cell death. The addition of CHIR99021 restores viability and allows efficient expansion of undifferentiated ES cells in the near absence of ERK signalling (Fig. 4f). We surmise that as phospho-ERK is diminished, downmodulation of GSK3 becomes increasingly crucial to maintain metabolic activity, biosynthetic capacity and overall viability.

This study reveals that the pathways required to sustain undifferentiated ES cells are dictated by the construction of the culture milieu. In a neutralized environment, ES cells can be efficiently derived and maintained without a requirement for growth factors or cytokines (Fig. 4g). We infer that BMP/Smad/Id and LIF/STAT3 signalling do not instruct self-renewal but act in unrefined culture conditions to shield the pluripotent state from induced phospho-ERK (Fig. 4h). Earlier studies have pointed to a positive effect of inhibiting the ERK cascade on ES-cell propagation in the context of additional signals[25,26]. However, upregulation of c-Myc, Stat3 or anti-apoptotic factors, previously invoked as key effectors of self-renewal, is not relevant in 3i. Our data do not exclude a contribution of stabilized β-catenin through TCF-independent mechanism(s), possibly acting as a noise filter[27]. Wnt3a does enhance neural suppression in PS cultures, but it gives significantly less benefit for overall propagation than CHIR99021 does. We infer that the pivotal contribution of GSK3 inhibition is to restore full growth and viability. This may be achieved by balancing the loss of ERK input into basic cellular processes. We detected no induction of anti-apoptotic factors (Supplementary Information), indicating that reduced GSK3 activity may exert a global modulation of the ES-cell metabolomic and biosynthetic capacity rather than having a direct anti-apoptotic action. Furthermore, restoration of the biosynthetic capacity of ES cells might itself increase the threshold for commitment. This possibility is suggested by the effect of feedback in mitogen-activated protein kinase signalling circuitry on the mating switch decision in yeast[28].

Previous empirical configurations of the culture environment have obscured the critical requirements for maintaining ES-cell pluripotency. We propose that ES cells are a basal cell state that is intrinsically self-maintaining if shielded effectively from inductive differentiation stimuli including autocrine FGF4. This feature may underlie the well-known predisposition of ES cells to generate teratocarcinomas[29,30]. They can dispense with an elementary cell signalling pathway, ERK, and do not seem to require any intercellular stimulation. They have not developed G1 cyclin checkpoint control of cell cycle progression and replicate constitutively[29]. ES cells thus display a self-sufficiency more akin to that of unicellular organisms than the interdependence generally exhibited by metazoan cells.

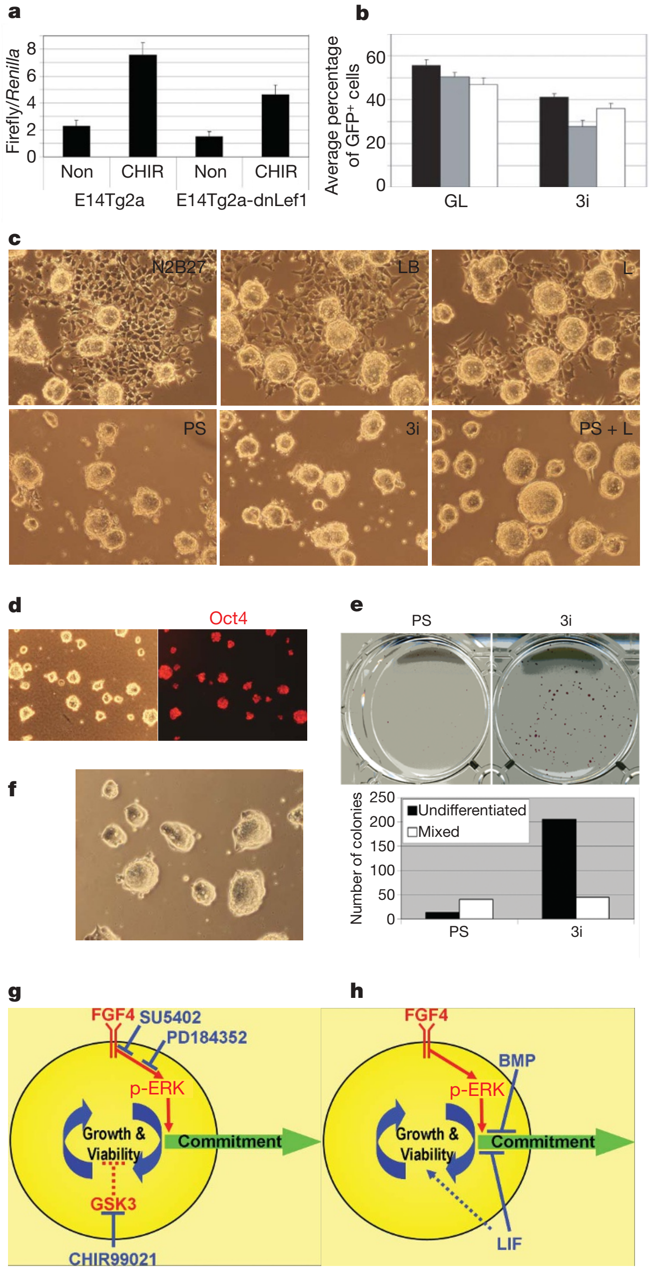


Figure 4 | CHIR99021 acts via inhibition of GSK3 to enhance ES-cell growth capacity and viability. **a**, TOPFlash assay in ΔNLef1 stable transfectants. Results are shown as means and s.d. for three biological replicates. Non, N2B27 alone; CHIR, CHIR99021. **b**, Competitive growth assay for three clones of ΔNLef1 stable transfectants and Cre revertants (GFP^+) in serum and LIF or in 3i. Results are shown as means and s.d. for four biological replicates. **c**, GSK3α/β-deficient DKO cells cultured in the indicated conditions for two passages. L, LIF. **d**, Phase-contrast and Oct4 immunostaining of DKO cells after eight passages in 1 μM PD0325901. Parallel cultures with addition of CHIR99021 were indistinguishable. **e**, ES cells constitutively expressing Nanog respond to CHIR99021 by enhanced self-renewal at low density in 3i compared with PS. **f**, ES cells self-renew in 2 μM PD0325901 plus CHIR99021. **g**, **h**, Diagrams of self-replication of the pluripotent state when inductive phospho-ERK signalling is either inhibited upstream by chemical antagonists (**g**) or counteracted downstream by LIF and BMP (**h**). Inhibition of GSK3 serves a key function in augmenting self-renewal when phospho-ERK (p-ERK) is suppressed by maintaining cellular growth capacity and additionally reinforcing suppression of neural commitment.

METHODS SUMMARY

CHIR99021, PD184352 and PD0325901 were synthesized in the Division of Signal Transduction Therapy, University of Dundee. Inhibitors were used at the following concentrations unless otherwise specified: CHIR99021, 3 μM; PD184352, 0.8 μM; SU5402 (Calbiochem), 2 μM; PD173074 (Sigma), 100 nM; PD0325901, 0.4 μM in 3i or 1 μM with CHIR99021 (2i). Clonal assays were performed by means of single-cell deposition into 96-well plates with a DakoCytomation MoFlo sorter.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 22 November 2007; accepted 2 April 2008.

1. Martin, G. R. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc. Natl Acad. Sci. USA* **78**, 7634–7638 (1981).
2. Evans, M. J. & Kaufman, M. Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154–156 (1981).
3. Rathjen, J. *et al.* Formation of a primitive ectoderm like cell population, EPL cells, from ES cells in response to biologically derived factors. *J. Cell Sci.* **112**, 601–612 (1999).
4. Sato, N., Meijer, L., Skaltsounis, L., Greengard, P. & Brivanlou, A. H. Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3-specific inhibitor. *Nature Med.* **10**, 55–63 (2004).
5. Ying, Q. L., Nichols, J., Chambers, I. & Smith, A. BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. *Cell* **115**, 281–292 (2003).
6. Yoshida, K. *et al.* Maintenance of the pluripotential phenotype of embryonic stem cells through direct activation of gp130 signalling pathways. *Mech. Dev.* **45**, 163–171 (1994).
7. Ogawa, K., Matsui, H., Ohtsuka, S. & Niwa, H. A novel mechanism for regulating clonal propagation of mouse ES cells. *Genes Cells* **9**, 471–477 (2004).

8. Niwa, H., Burdon, T., Chambers, I. & Smith, A. Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. *Genes Dev.* **12,** 2048–2060 (1998).
9. Kunath, T. *et al.* FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. *Development* **134,** 2895–2902 (2007).
10. Stavridis, M. P., Lunn, J. S., Collins, B. J. & Storey, K. G. A discrete period of FGF-induced Erk1/2 signalling is required for vertebrate neural specification. *Development* **134,** 2889–2894 (2007).
11. Mohammadi, M. *et al.* Structures of the tyrosine kinase domain of fibroblast growth factor receptor in complex with inhibitors. *Science* **276,** 955–960 (1997).
12. Davies, S. P., Reddy, H., Caivano, M. & Cohen, P. Specificity and mechanism of action of some commonly used protein kinase inhibitors. *Biochem. J.* **351,** 95–105 (2000).
13. Zhen, Y., Sorensen, V., Jin, Y., Suo, Z. & Wiedlocha, A. Indirubin-3′-monoxime inhibits autophosphorylation of FGFR1 and stimulates ERK1/2 activity via p38 MAPK. *Oncogene* **26,** 6372–6385 (2007).
14. Murray, J. T. *et al.* Exploitation of KESTREL to identify NDRG family members as physiological substrates for SGK1 and GSK3. *Biochem. J.* **384,** 477–488 (2004).
15. Bain, J. *et al.* The selectivity of protein kinase inhibitors; a further update. *Biochem. J.* **408,** 297–315 (2007).
16. Buehr, M. & Smith, A. Genesis of embryonic stem cells. *Phil. Trans. R. Soc. B* **358,** 1397–1402 (2003).
17. Mohammadi, M. *et al.* Crystal structure of an angiogenesis inhibitor bound to the FGF receptor tyrosine kinase domain. *EMBO J.* **17,** 5896–5904 (1998).
18. Wilder, P. J. *et al.* Inactivation of the FGF-4 gene in embryonic stem cells alters the growth and/or the survival of their early differentiated progeny. *Dev. Biol.* **192,** 614–629 (1997).
19. Saba-El-Leil, M. K. *et al.* An essential function of the mitogen-activated protein kinase Erk2 in mouse trophoblast development. *EMBO Rep.* **4,** 964–968 (2003).
20. Cartwright, P. *et al.* LIF/STAT3 controls ES cell self-renewal and pluripotency by a Myc-dependent mechanism. *Development* **132,** 885–896 (2005).
21. Matsuda, T. *et al.* STAT3 activation is sufficient to maintain an undifferentiated state of mouse embryonic stem cells. *EMBO J.* **18,** 4261–4269 (1999).
22. Ogawa, K., Nishinakamura, R., Iwamatsu, Y., Shimosato, D. & Niwa, H. Synergistic action of Wnt and LIF in maintaining pluripotency of mouse ES cells. *Biochem. Biophys. Res. Commun.* **343,** 159–166 (2006).
23. Hao, J., Li, T. G., Qi, X., Zhao, D. F. & Zhao, G. Q. WNT/β-catenin pathway up-regulates Stat3 and converges on LIF to prevent differentiation of mouse embryonic stem cells. *Dev. Biol.* **290,** 81–91 (2006).
24. Doble, B. W., Patel, S., Wood, G. A., Kockeritz, L. K. & Woodgett, J. R. Functional redundancy of GSK-3α and GSK-3β in Wnt/β-catenin signaling shown by using an allelic series of embryonic stem cell lines. *Dev. Cell* **12,** 957–971 (2007).
25. Burdon, T., Stracey, C., Chambers, I., Nichols, J. & Smith, A. Suppression of SHP-2 and ERK signalling promotes self-renewal of mouse embryonic stem cells. *Dev. Biol.* **210,** 30–43 (1999).
26. Chen, S. *et al.* Self-renewal of embryonic stem cells by a small molecule. *Proc. Natl Acad. Sci. USA* **103,** 17266–17271 (2006).
27. Martinez-Arias, A. & Hayward, P. Filtering transcriptional noise during development: concepts and mechanisms. *Nature Rev. Genet.* **7,** 34–44 (2006).
28. Colman-Lerner, A. *et al.* Regulated cell-to-cell variation in a cell-fate decision system. *Nature* **437,** 699–706 (2005).
29. Burdon, T., Smith, A. & Savatier, P. Signalling, cell cycle and pluripotency in embryonic stem cells. *Trends Cell Biol.* **12,** 432–438 (2002).
30. Chambers, I. & Smith, A. Self-renewal of teratocarcinoma and embryonic stem cells. *Oncogene* **23,** 7150–7160 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank D. Alessi for discussion and advice on GSK3 signalling; A. Rizzino, S. Meloche and R. Kemler for *Fgf4*, *Erk2* and *Ecadherin* targeted ES cells, respectively; F. Watt for the ΔNhLef1 construct; N. Shpiro and R. Marquez for synthesizing PD184352, CHIR99021 and PD0325901; B. Amati and G. Faga for advice on Myc immunoblotting; J. Vrana for fluorescence-activated cell sorting support; and C. Manson, K. Savill and colleagues for mouse husbandry. This research was funded by the Medical Research Council and the Biotechnology and Biological Sciences Research Council of the UK, the Canadian Institutes of Health Research, and by the European Commission Framework VI project EuroStemCell. P.C. is a Royal Society Research Professor, and A.S. is a Medical Research Council Professor.

Author Contributions Q.L.Y. and A.S. conceived the study; Q.L.Y., J.W. and J.N. designed, executed and interpreted experiments; L.B.M. generated CBA ES cells; B.D. and J.W. generated and provided GSK3 mutant ES cells; P.C. contributed expert advice and inhibitors; and A.S. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to Q.L.Y. (qying@keck.usc.edu) or A.S. (ags39@cscr.cam.ac.uk).

METHODS

Details of RT–PCR conditions and antibodies are provided in Supplementary Information.

ES-cell culture. N2B27 medium was prepared as described[31,32] or with preformulated NDiff N2B27 base medium (catalogue no. SCS-SF-NB-02; Stem Cell Sciences Ltd.) Where specified, recombinant human albumin (Cellastim; Invitria) was substituted for bovine serum albumin. Cells were routinely propagated by trypsinization and replating every three days, with a split ratio of 1 in 10.

ES-cell derivation. Whole diapause blastocysts (strain 129) or isolated epiblasts (CBA) were plated in pre-equilibrated N2B27 3i medium. After several days, cell masses were disaggregated into small clumps of cells with trypsin and replated. Emerging ES-cell colonies were expanded by replating into successively larger wells. Wild-type MF1 ES cells were derived from whole blastocysts as for strain 129. These cells showed decreased substrate attachment, probably as a result of the outbred MF1 genetic background. They can be passaged on laminin-coated plastic but are more readily maintained on murine embryo fibroblast feeders. Accordingly, for *Stat3* mutant derivations we employed feeders for part of the process. Eight-cell embryos from intercrosses of *Stat3*$^{+/-}$ outbred MF1 mice[33] were cultured in KSOM medium containing 3i until the formation of expanded blastocysts. Trophectoderm layers were removed by immunosurgery and analysed by PCR for genotype determination[34]. Four isolated inner-cell masses (three *Stat3*$^{+/-}$, one *Stat3*$^{-/-}$) were plated individually into four-well plates on feeders in pre-equilibrated N2B27 3i. After five days, cell masses were disaggregated with trypsin and plated into fresh four-well plates. ES cells developed in all four cultures and were expanded in laminin-coated wells without feeders.

Transfections. ΔNhLef1 (ref. 35) was inserted between loxP sites in pPyFloxedMTIPgfp[36]. This construct was transfected by electroporation into E14Tg2a ES cells for stable integration, or into E14T ES cells[36] for episomal propagation. Dual luciferase reporter assays were performed 56 h after lipofection with TOPFlash or FOPFlash constructs.

Competitive self-renewal assay. Three clones of E14Tg2a stable transfectants were transiently transfected with a Cre expression vector to excise ΔNhLef1 and simultaneously activate GFP expression. After 24 h, GFP-positive and GFP-negative cells were fractionated by fluorescence-activated cell sorting (FACS), recombined in equal numbers and plated in six-well plates at 10^5 cells per well in N2B27 with 3i. Cells were expanded for four passages and then analysed by FACS to establish the proportion of GFP$^+$ (ΔNhLef1 excised) and GFP$^-$ (ΔNhLef1-expressing) cells.

31. Brewer, G. J., Torricelli, J. R., Evege, E. K. & Price, P. J. Optimized survival of hippocampal neurons in B27-supplemented Neurobasal, a new serum-free medium combination. *J. Neurosci. Res.* **35**, 567–576 (1993).
32. Ying, Q. L. & Smith, A. G. Defined conditions for neural commitment and differentiation. *Methods Enzymol.* **365**, 327–341 (2003).
33. Takeda, K. *et al.* Targeted disruption of the mouse *Stat3* gene leads to early embryonic lethality. *Proc. Natl Acad. Sci. USA* **94**, 3801–3804 (1997).
34. Nichols, J. *et al.* Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* **95**, 379–391 (1998).
35. Niemann, C., Owens, D. M., Hulsken, J., Birchmeier, W. & Watt, F. M. Expression of DeltaNLef1 in mouse epidermis results in differentiation of hair follicles into squamous epidermal cysts and formation of skin tumours. *Development* **129**, 95–109 (2002).
36. Chambers, I. *et al.* Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* **113**, 643–655 (2003).

The Transcriptional and Epigenomic Foundations of Ground State Pluripotency

Hendrik Marks,[1] Tüzer Kalkan,[2] Roberta Menafra,[1] Sergey Denissov,[1] Kenneth Jones,[2] Helmut Hofemeister,[5] Jennifer Nichols,[2,3] Andrea Kranz,[5] A. Francis Stewart,[5] Austin Smith,[2,4,*] and Hendrik G. Stunnenberg[1,*]
[1]Department of Molecular Biology, Faculty of Science, Nijmegen Centre for Molecular Life Sciences (NCMLS), Radboud University, PO Box 9101, 6500 HB Nijmegen, The Netherlands
[2]Wellcome Trust Centre for Stem Cell Research and Stem Cell Institute
[3]Department of Physiology, Development and Neuroscience
[4]Department of Biochemistry
University of Cambridge, Tennis Court Road, Cambridge CB2 1QR, UK
[5]Genomics, BioInnovationsZentrum, Technische Universität Dresden, Am Tatzberg 47-51, D-01307 Dresden, Germany
*Correspondence: austin.smith@cscr.cam.ac.uk (A.S.), h.stunnenberg@ncmls.ru.nl (H.G.S.)
DOI 10.1016/j.cell.2012.03.026

SUMMARY

Mouse embryonic stem (ES) cells grown in serum exhibit greater heterogeneity in morphology and expression of pluripotency factors than ES cells cultured in defined medium with inhibitors of two kinases (Mek and GSK3), a condition known as "2i" postulated to establish a naive ground state. We show that the transcriptome and epigenome profiles of serum- and 2i-grown ES cells are distinct. 2i-treated cells exhibit lower expression of lineage-affiliated genes, reduced prevalence at promoters of the repressive histone modification H3K27me3, and fewer bivalent domains, which are thought to mark genes poised for either up- or downregulation. Nonetheless, serum- and 2i-grown ES cells have similar differentiation potential. Precocious transcription of developmental genes in 2i is restrained by RNA polymerase II promoter-proximal pausing. These findings suggest that transcriptional potentiation and a permissive chromatin context characterize the ground state and that exit from it may not require a metastable intermediate or multilineage priming.

INTRODUCTION

Mouse embryonic stem (ES) cells are characterized by the potency to generate all somatic and germline lineages in vitro and in chimaeric embryos (Smith, 2001). The nature of the transcriptional and epigenetic machinery that maintains this potential throughout massive in vitro expansion has been the subject of intense investigation (Young, 2011). Interest is further heightened by appreciation that knowledge of the molecular underpinning of mouse ES cells may enable derivation of equivalent human naive pluripotent stem cells (Hanna et al., 2010).

ES cells are described as transcriptionally hyperactive (Efroni et al., 2008). Promiscuous transcription has been suggested to constitute a platform for lineage specification (Loh and Lim, 2011). When taken together with the observation that several pluripotency factors are expressed heterogeneously (Chambers et al., 2007; Niwa et al., 2009; Toyooka et al., 2008), the notion has arisen that pluripotency may inherently be a metastable condition (Graf and Stadtfeld, 2008; Hanna et al., 2009; Hayashi et al., 2008). Attention has also been drawn to colocalization at many promoters of histone 3 lysine 4 trimethylation (H3K4me3), associated with transcriptional activation, and histone 3 lysine 27 trimethylation (H3K27me3), linked with repression (Azuara et al., 2006; Bernstein et al., 2006; Stock et al., 2007). These bivalent domains are posited to be poised for either up- or downregulation and to provide an epigenetic blueprint for lineage determination.

The preceding views are based on analyses of ES cells cultured in serum and therefore subject to uncontrolled multifactorial perturbations. It is now possible to derive and maintain pluripotent mouse ES cells without serum factors by using two small molecule kinase inhibitors (2i) in combination with the cytokine leukemia inhibitory factor (LIF) (Ying et al., 2008). The 2i inhibitors, PD0325901 and CHIR99021, selectively target mitogen-activated protein kinase kinase (Mek) and glycogen synthase kinase-3 (Gsk3), respectively. The inhibitors shield pluripotent cells from differentiation triggers: fibroblast growth factor-4 stimulation of the Mek-Erk pathway and endogenous repressor activity of Tcf3 (Kunath et al., 2007; Wray et al., 2011). Use of 2i has enabled derivation of germline-competent ES cells from all mouse strains tested and for the first time from rats (Buehr et al., 2008; Hanna et al., 2009; Kiyonari et al., 2010; Li et al., 2008; Nichols et al., 2009). 2i thus provides a better-tuned environment for rodent ES cells. Indeed, the mosaic expression of pluripotency factors observed in serum is effectively eliminated in 2i (Wray et al., 2011). Furthermore, culture in 2i may mimic the environment in the mature mouse inner cell mass (ICM) where the fibroblast growth factor receptor is downregulated in the epiblast (Guo et al., 2010).

Here we applied massively parallel sequencing technology to characterize the global transcriptome and to map selected histone modifications in naive mouse ES cells maintained in 2i compared with heterogeneous cultures in serum.

RESULTS

Transcriptome Analysis

Three ES cell lines derived and maintained in 2i plus LIF ("2i" ES cells) were compared with three ES cell lines established and cultured in serum plus LIF ("serum" ES cells) (Tables S1 and S2). Each cell line is functionally pluripotent as demonstrated by competence to generate high-contribution chimaeras with germline transmission. Expression values from RNA-seq data were calculated by quantifying the number of sequence reads for each gene with standardized RPKM values (reads per kilobase of exon model per million mapped sequence reads). This comparison showed that 1,489 genes have more than 2-fold higher transcript abundance in 2i (p value < 0.2), whereas 1,947 genes exhibit more than 2-fold higher expression in serum (Figure 1A). Moreover, 160 genes expressed in 2i (RPKM > 0.5) were silent in serum (RPKM < 0.2) and 461 genes were expressed only in serum (Figure S1A and Table S3 available online). The majority of categorized stem cell maintenance genes ([SCM] GO:0019827), including validated core pluripotency factors, *Pou5f1*, *Nanog*, *Sox2*, *Esrrb*, *Klf2*, *Klf4*, and *Tbx3*, are transcribed to similar levels in 2i and serum (Figure 1B and Figure S1B). Nine SCM genes are more highly expressed in 2i (Figure 1B and Figure S1C). Of these, only *Tcl1* has been implicated as a regulator of self-renewal (Ivanova et al., 2006), and although this transcript is more abundant in 2i, it is also well-expressed in serum. In serum, 16 SCM genes showed higher expression. Factors in this group, such as c-Myc and the Id proteins, are known to be induced by Erk signaling and by serum. They may confer additional robustness to the pluripotent state to counter differentiation stimuli (Ying et al., 2003a). Interestingly, several of these genes are induced when ICM cells are explanted in medium containing serum (Figure S1D; Tang et al., 2010).

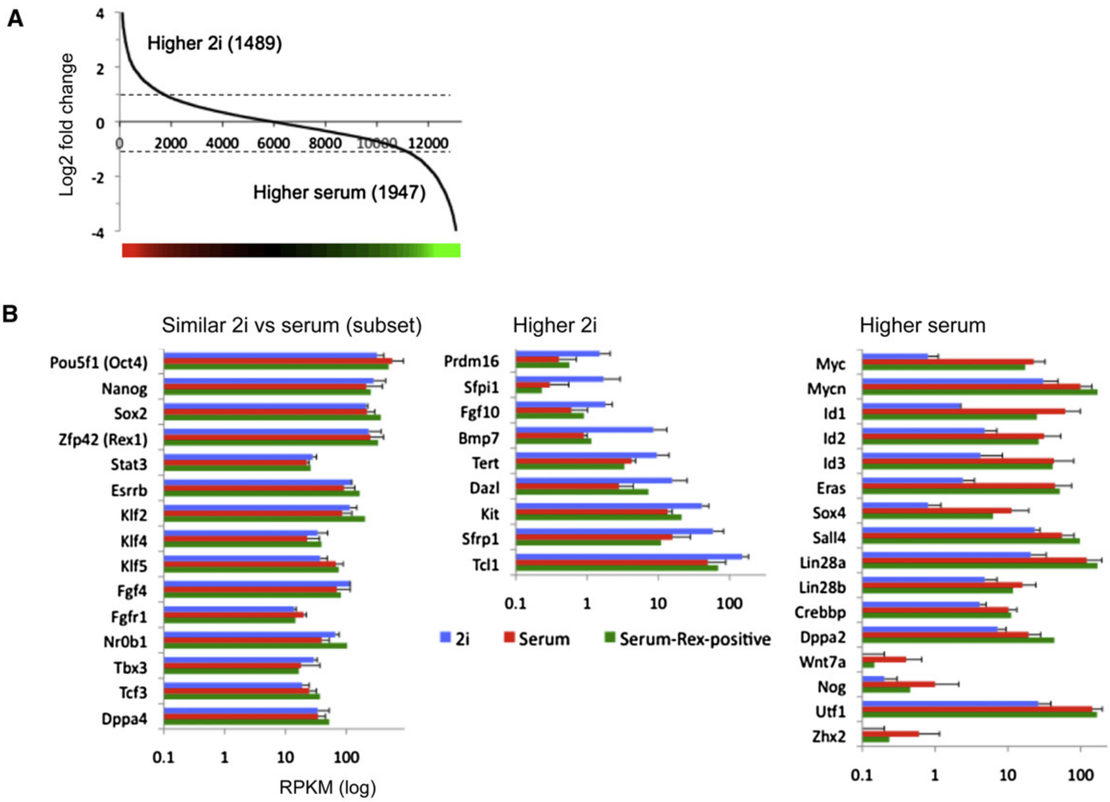
Functional annotation clustering of differentially expressed genes by Gene Ontology ([GO] PANTHER [protein analysis through evolutionary relationships]) and Pathway (KEGG [Kyoto encyclopedia of genes and genomes]) analysis (Figure 1C) revealed that genes upregulated in 2i are highly enriched for terms associated with metabolic processes. This is probably in part a direct consequence of inhibition of Mek and Gsk3 and the absence of serum constituents. Reduced c-Myc may further affect metabolic networks. Major differences are also apparent for genes involved in cell-cycle regulation. Proliferation is similar in the two conditions (Ying et al., 2008), however, reflecting the absence of G1 restriction in ES cells.

Upregulated genes in serum are significantly enriched for GO terms linked to developmental processes, particularly ectoderm and mesoderm germ layer specification (Figure 1C). Genes such as *Pax6*, *T (Brachyury)*, and *Runx1* show very low or undetectable expression in 2i but appreciable transcription in serum (Figure 1D). Other ectoderm and mesoderm marker genes such as *Runx3*, *Sox18*, *Cdx4*, and *Tal1* are also activated in serum, although only to low levels. In contrast, several genes associated with the germline or with endoderm are expressed at similar levels in both conditions.

ES cells maintained in 2i are morphologically uniform and rather homogeneous in expression of pluripotency regulators (Wray et al., 2010). In contrast, serum ES cells are heterogeneous in morphology (Figure S2A) and expression of factors such as Nanog, Rex1, Stella, and Klf4 (Chambers et al., 2007; Hayashi et al., 2008; Toyooka et al., 2008; Figure S2B). In serum, ES cells with a mono-allelic green fluorescent protein (GFP) knockin at the Rex1 (*Zfp42*) locus comprise GFP-positive and -negative populations that can be separated by flow cytometry (Figure S2C; Wray et al., 2011). These populations are functionally distinct. Rex1GFP-positive cells plated in serum generate colonies of undifferentiated cells, whereas Rex1GFP-negative cells produce predominantly small, differentiated colonies (Figure S2D). On plating in 2i, the yield of ES cell colonies in Rex1-positive cells is more than 10 times that in Rex1-negative cells. Rex1-negative cells have therefore largely lost clonogenic self-renewal capacity. Consistent with this, although they express Oct4, they lack Nanog and Klf4 (Figure S2E). All Nanog-positive and almost all Klf4-positive cells are within the Rex1GFP-positive fraction. Expression is still mosaic within this population, reflecting transcriptional fluctuation (Chambers et al., 2007; Kalmar et al., 2009; Figure S2E). In contrast, ES cells in 2i are almost uniformly positive for Rex1, Klf4, and Nanog (Wray et al., 2010; Figure S2B). We examined the transcriptome of the Rex1-positive fraction in serum and found a similar increased expression of a subset of SCM genes as in bulk ES cells (Figure 1B). Some ectoderm- and mesoderm-associated genes were also upregulated compared with expression in 2i ES cells but others showed little induction (Figure 1D). In general, differentiation genes were more highly expressed in Rex1GFP-low cells (Figures S2F–S2H). Nonetheless, transcriptomes of the Rex1-positive compartment in serum show many differences from 2i ES cells (Figure S2I). Notably, many genes that show higher expression in 2i are not upregulated in Rex1-positive serum ES cells (Figure 1B). Therefore it is unlikely that a subpopulation of serum cells persist in an equivalent state to 2i.

Transcriptome Interconversion

To assess whether the distinct gene expression patterns in 2i and serum represent "fixed" transcriptional states, we transferred 2i cells to serum and vice versa. Within two to three passages, cultures adopted morphological characteristics of the new condition. We carried out RNA-seq analyses after eight passages. Most of the SCM genes that showed lower expression in 2i were upregulated in serum, whereas transcripts elevated in 2i were downregulated (Figure 2A red dots and black squares, respectively). Reciprocal behavior was observed when serum cells were transferred to 2i (Figure 2B). Furthermore, genes involved in ectoderm and mesoderm germ layer specification were broadly upregulated after transfer of 2i cells to serum and the reverse transcriptional changes were observed when serum cells were passaged in 2i (Figures 2A and 2B, blue stars). Typical examples are shown in Figure 2C. Irrespective of the direction of exchange between 2i and serum, 818 genes were expressed more highly (>2-fold) in 2i and 1,209 showed higher levels in serum (Figure 2D). GO classification identified developmental genes and cell-cycle control genes as highly enriched upon transfer to serum, whereas genes upregulated in 2i were mainly associated with metabolic categories (Figure 2E). The reciprocity in transcriptome changes demonstrates that the transcriptional profiles are interconvertible.



C

Higher 2i

PANTHER biological processes

Term	Count	Total category	P-Value
Lipid, fatty acid and steroid metabolism	86	780	8.7E-08
Fatty acid metabolism	28	192	4.0E-05
Carbohydrate metabolism	56	552	2.0E-04
Immunity and defense	131	1594	2.5E-04
Other carbon metabolism	14	79	9.0E-04
Other metabolism	54	562	9.6E-04
Detoxification	14	83	1.4E-03
Phospholipid metabolism	19	140	2.1E-03
Cation transport	43	464	6.3E-03

KEGG pathway

Term	Count	Total category	P-Value
Lysosome	32	119	9.5E-12
Propanoate metabolism	10	30	8.1E-05
Tryptophan metabolism	10	40	8.6E-04
Hematopoietic cell lineage	15	84	9.4E-04
Glycosphingolipid biosynthesis	6	14	1.4E-03
Glutathione metabolism	11	52	1.6E-03

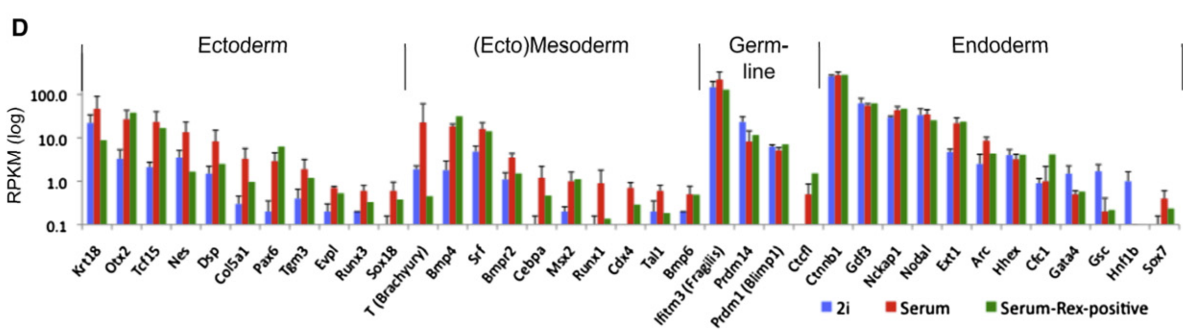
Higher serum

PANTHER biological processes

Term	Count	Total category	P-Value
Developmental processes	293	2106	8.8E-14
Mesoderm development	95	533	7.1E-10
Ectoderm development	105	659	4.3E-08
Cell structure and motility	145	1015	1.1E-07
Neurogenesis	91	579	7.1E-07
Synaptic transmission	49	261	3.4E-06
Cell proliferation and differentiation	117	881	6.7E-05
Protein phosphorylation	89	657	2.9E-04
JNK cascade	16	60	3.0E-04
Cell cycle control	57	390	7.0E-04
Muscle development	24	133	2.6E-03
MAPKKK cascade	28	181	9.6E-03
Cell cycle	98	877	3.9E-02

KEGG pathway

Term	Count	Total category	P-Value
Pathways in cancer	73	323	4.8E-12
Focal adhesion	51	198	1.1E-10
ECM-receptor interaction	25	83	6.4E-07
MAPK signaling pathway	50	265	6.2E-06
TGF-beta signaling pathway	21	87	2.0E-04
Wnt signaling pathway	29	149	4.8E-04
Hedgehog signaling pathway	15	54	5.0E-04



Histone Modification Profiles

We performed chromatin immunoprecipitation and deep sequencing (ChIP-seq) to analyze posttranslational histone modifications: H3K4me3 and H3K36me3 associated with active promoters and transcribed genes, respectively; H3K27me3 linked to silencing, and H3K9me3 associated with constitutive heterochromatin and imprinted genes (Table S1). We also analyzed the polycomb repressor complex 2 (PRC2) component Ezh2 that methylates H3K27 (Cao et al., 2002).

Determination of average profiles over 2,000 genes that are most highly expressed in both conditions (Figure 3A and Figure S3A) revealed conventional distribution of H3K4me3 on active promoters and of H3K36me3 extending over the coding body. The H3K9me3 ChIP-seq state maps of 2i and serum cells were nearly identical. In both conditions deposits were most prominent at satellites and imprinted genes (Figures S3B–S3E). As expected, the H3K27me3 mark is absent from actively transcribed loci (Figure 3B). It appears to be mutually exclusive with H3K36me3 (Figures 3A–3C and Figure S3F), in line with recent biochemical data showing that PRC2 activity is inhibited by active marks including H3K36me3 (Schmitges et al., 2011). However, H3K27me3 is widely deposited over intergenic regions and inactive genes at levels appreciably higher than random distribution. This lawn of H3K27me3 is qualitatively and quantitatively similar in 2i and serum ES cells (Figure 3C). A pronounced difference is apparent only at promoters of lowly expressed genes (Figure 3B). The averaged profile of these promoters showed markedly less H3K27me3 in 2i than in serum. This was not reflected in any overall increase in expression (Figure S3H). Three independent 2i ES cell lines exhibited a significantly reduced level of H3K27me3 at the promoters of poorly or nonexpressed genes compared to the level in serum cultures (Figure 3D). To investigate whether differences in H3K27me3 deposition reflected heterogeneity in serum, we performed H3K27me3 ChIP-seq on Rex1-positive and Rex1-negative serum subpopulations. Intriguingly, the H3K27me3 signals were very similar, each resembling the total serum ES cell population (Figure 3E).

Intensity plots covering a region of 5 kb up- and downstream of all promoters that are decorated with H3K27me3 in serum demonstrate the major reduction in H3K27me3 in 2i cells (Figure 3F). Ezh2 levels were also diminished at these locations in 2i. In either condition, the H3K27me3 pattern follows a camelback profile with a depleted region around the transcriptional start site. Ezh2 appears as a single peak centered on this trough. Representative examples of differential H3K27me3 profiles are shown in Figure 3G (Figure S3I shows the PCR validation). The *Gata6*, *Pax9*, and *Lhx1* genes are barely expressed in either 2i or serum, but in all cases the H3K27me3 signal around the promoter is selectively and greatly reduced in 2i. For the *Lhx1* locus, adjacent *Aatf* provides a contrasting example of a gene that is productively transcribed in both 2i and serum and remains devoid of H3K27me3 in either condition.

Given the interchangeable transcriptome profiles between 2i and serum, we examined the epigenomic landscape in ES cells transferred between the two conditions. Cells taken from 2i into serum acquired substantially elevated H3K27me3 at H3K27me3-associated promoters (Figure 3H), the *Hox* clusters (Figure S4A), and many other loci (Figure S4B). Conversely, serum ES cells transferred into 2i displayed diminished H3K27me3 at these loci. Ezh2 and Suz12 localization similarly switched between culture conditions (Figure 3H and Figures S4A and S4B). Therefore these epigenomic states are interconvertible.

H3K27me3 was reduced by between 63%–75% over all *Hox* clusters in 2i (Figure 3G and Figure S4A). The *Hoxc* locus follows this pattern but with a distinctive variation; in the *Hoxc13-c12* region H3K27me3 deposition is lost entirely. This region (boxed in Figure 3G) is transcribed only in 2i. Strand-specific RNA-seq profiling after rRNA depletion revealed two nonoverlapping transcripts on the reverse and forward strand (both boxed in Figure S4C, left). These ncRNAs are distinct from the HOTAIR ncRNA located between *HOXC11-12* in human (Rinn et al., 2007). Consistent with recent findings (Guttman et al., 2010), we detected known as well as multiple previously unidentified ncRNAs. Many of these, such as H19, showed differential expression between 2i and serum (Figure S4C and Table S4).

Global Redistribution of H3K27me3

We computed the number of H3K27me3 reads over nonrepetitive regions and plotted the frequency of occurrence and the genomic location. In 2i, high H3K27me3 deposition is scarce with very little enrichment at promoters. In contrast, in serum H3K27me3 is elevated at many genomic locations, 60%–65% of which are promoters (Figures 4A and 4B). In 2i, H3K27me3 is somewhat reduced over long interspersed nuclear element (LINE) repeats (Figure 4C). This is more than offset, however, by much higher levels of H3K27me3 present at satellites. Immunoblotting showed that the total cellular level of H3K27me3 is comparable in 2i and serum (Figure 4D), confirming that the differences at promoters are not secondary to a general reduction in H3K27me3 deposition in 2i.

H3K27me3 is deposited by PRC2 and facilitates recruitment of the PRC1 complex. Transcripts of PRC2 and PRC1 subunits were present at similar levels in 2i and serum (Figure S5A). Transcripts for the H3K27me3 demethylases Kdm6a and

Figure 1. Transcriptome Profiling of ES Cells in 2i and Serum

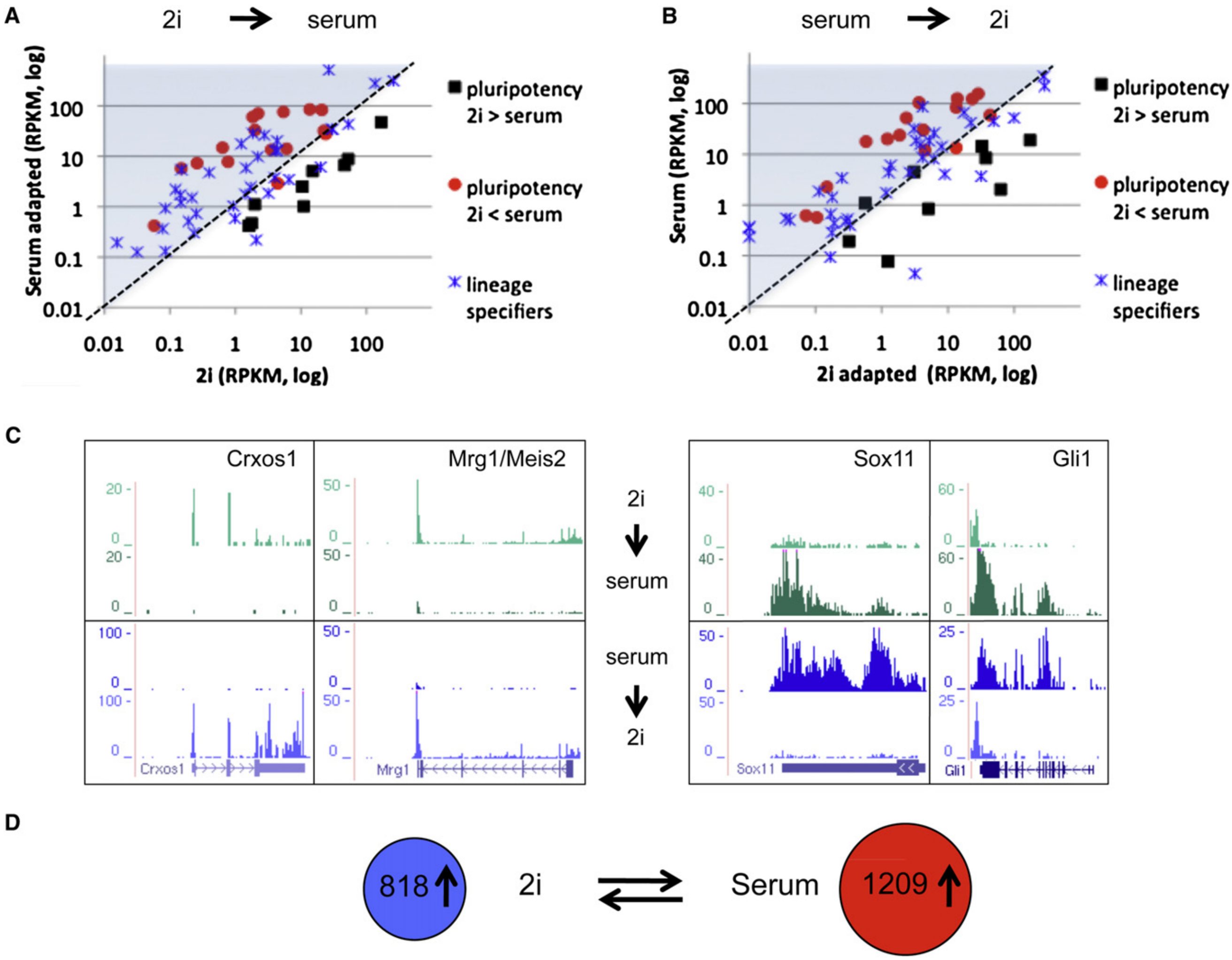
(A) Fold change ($\log_2$ values) in transcript level of all genes in 2i versus serum. Gene expression values of three ES cells lines derived and maintained in either 2i (TNGA, NOD_male, and NOD_female) or serum (E14, XT67E1, and RGD2) were averaged, after which ratios were calculated. A 2-fold change is indicated by the dotted line. The corresponding heatmap is shown at the bottom. For the remaining analysis, an extra constraint for differential gene expression was set by a p value < 0.2 (Student t test).

(B) RNA-seq levels of a selection of pluripotency, self-renewal, and stem cell markers for 2i and serum ES cells as shown in (A). Expression values for the Rex1-positive serum ES cell population as collected by FACS (Figure S2C) are included.

(C) Functional annotation analysis of the differentially expressed genes between 2i and serum ES cells.

(D) Transcript level of genes associated with the various germ layers.

See also Figure S1, Figure S2, Table S1, Table S2, and Table S3.



E

Higher 2i

PANTHER biological processes

Term	Count	Total category	P-Value
Other carbon metabolism	11	79	4.9E-05
Other metabolism	32	562	1.3E-04
Lipid, fatty acid and steroid metabolism	38	780	5.8E-04
Fatty acid metabolism	15	192	7.1E-04
Carbohydrate metabolism	28	552	2.1E-03
Coenzyme and prosthetic group metabolism	11	153	8.5E-03
Ion transport	27	601	1.2E-02
Amino acid metabolism	13	219	1.6E-02

KEGG pathway

Term	Count	Total category	P-Value
Glycolysis / Gluconeogenesis	10	68	1.7E-04
Propanoate metabolism	7	30	2.2E-04
Metabolism of xenobiotics by cytochrome P4	9	66	7.2E-04
Lysosome	12	119	8.1E-04

Higher serum

PANTHER biological processes

Term	Count	Total category	P-Value
Developmental processes	176	2106	1.3E-12
Mesoderm development	61	533	2.7E-09
Cell communication	99	1115	1.9E-08
Cell structure and motility	92	1015	2.7E-08
Ectoderm development	67	659	4.6E-08
Neurogenesis	59	579	3.1E-07
Cell adhesion	55	575	5.7E-06
Cell proliferation and differentiation	68	881	3.7E-04
Cell cycle control	29	390	3.4E-02

KEGG pathway

Term	Count	Total category	P-Value
ECM-receptor interaction	19	83	1.8E-07
Pathways in cancer	38	323	4.1E-06
Focal adhesion	27	198	1.1E-05
MAPK signaling pathway	30	265	1.1E-04
TGF-beta signaling pathway	15	87	1.5E-04
Hedgehog signaling pathway	11	54	4.1E-04

Figure 2. Expression Profiles of 2i and Serum ES Cells Are Interconvertible

(A and B) Comparison of expression of pluripotency and lineage-specific genes (as shown in Figures 1B and 1D) for TNGA cells adapted from 2i to serum (A) and E14 adapted from serum to 2i (B).

(C) Typical examples of genes that show transcriptional interconvertibility.

(D) Number of genes that show interconvertibility (>2-fold difference between both 2i and serum conditions).

(E) Functional analysis of the differential genes shown in (D), the genes consistently higher in either 2i or serum.

Kdm6b (also known as Utx and Jmjd3) were also comparable. Ezh2 immunoblotting indicates slightly lower protein in 2i than in serum (Figure S5B). However, phosphorylation of Ezh2 at Thr345, reported to be important for PRC2 recruitment (Kaneko et al., 2010), is similar (Figure S5C). Collectively, these data suggest that the difference in H3K27me3 occupancy at silent promoters in 2i is not primarily attributable to reduced expression of polycomb nor to altered demethylase expression.

Bivalency

Promoters that are marked by H3K27me3 may also display H3K4me3. Such bivalent genes are thought to be poised for activation (Azuara et al., 2006; Bernstein et al., 2006; Mikkelsen et al., 2007). We binned and ranked promoters according to the read density for H3K27me3 (Figure 5A) measured in serum and assessed whether bivalency is preserved in naive ES cells. Applying similar filters and thresholds as used by Mikkelsen et al. (2007), we classified almost 3,000 genes as bivalent in serum (Figure 5B, upper, and Table S5). In 2i, due to the reduced deposition of H3K27me3, many of these genes fall below the threshold, resulting in less than 1,000 genes that qualify as bivalent (Figure 5B, lower). Intensity plots show the general and pronounced diminution in H3K27me3 deposition, whereas H3K4me3 is only slightly altered (Figure 5C). Figure S6A documents the levels of mRNA, H3K27me3, and H3K4me3 in 2i versus serum. Notably, the profiles interconvert upon switching cells between serum and 2i (Figures S6B and S6C).

In both serum and 2i, the bivalent genes are enriched for involvement in developmental processes. Representative examples are the mesoderm specification marker *Hey2* and ectodermal *Metrnl* (Figure 5D). Transcripts are barely detectable in 2i, although H3K4me3 is present and H3K27me3 is low. In serum, transcription is slightly upregulated even though the promoters show a broad gain of H3K27me3. A significant proportion of genes with bivalent promoter marking (31%) exhibit only background transcription in either condition (RPKM < 0.2). However, 14% of the bivalent genes are serum specific (RPKM > 0.5 in serum; RPKM < 0.2 in 2i), whereas a minor fraction (4%) are expressed only in 2i.

In either serum or 2i, H3K27me3 does not colocalize precisely with H3K4me3 but accumulates on either side of the H3K4me3 peak at the transcription start site (Figure 5C). This is consistent with observations that targets of TrxG proteins, which methylate H3K4, show reduced H3K27 methylation (Papp and Müller, 2006; Srinivasan et al., 2008) and that PRC2 activity is inhibited by active marks, including H3K4me3 (Schmitges et al., 2011). Strikingly RNA polymerase II (Pol II) is evident over transcription start sites at higher levels in 2i than in serum (Figure 5C), suggestive of promoter proximal pausing.

Influence of c-Myc

c-Myc is implicated in Pol II pause release (Rahl et al., 2010). We previously noted a very low level of c-Myc protein in 2i (Ying et al., 2008). The RNA-seq data show that *c-myc* mRNA is 40- to 50-fold lower in 2i than in serum and *n-myc* and *l-myc* are also reduced (Figure 1B). We analyzed c-Myc targets that are upregulated in serum (Figure 6A and Figure S7A). Averaged profiles show that promoters of these genes are loaded with H3K4me3 and Pol II in 2i. In serum, Pol II is reduced at the promoters and increased over coding bodies. The Pol II traveling ratio is accordingly increased (Figure 6B), consistent with c-Myc acting as a pause release factor in serum. Typical examples are *Npm1* and *Ncl* (Figure 6C).

We assessed to what extent global differences in transcriptome between 2i and serum might be related to c-Myc. Several differentially expressed genes are c-Myc targets, notably cell-cycle regulators (Figures S7B and S7C). These include cdk/cyclinD components that are positively regulated by Myc and are increased in serum, and conversely cell-cycle inhibitors, p16(Ink4A), p19(Arf), and p21 that are repressed by Myc and upregulated in 2i (Figure S7D). Overall, however, direct c-Myc targets as determined by Chen et al. (2008) represent less than 15% of genes differentially expressed between 2i and serum (Figure 6D). Furthermore, gene ontology classification of c-Myc targets upregulated in serum did not identify categories associated with developmental processes. Therefore Myc is unlikely to be a major determinant of differential expression and metastability.

RNA Polymerase II Pausing

We evaluated the average histone modification profile of all genes that change expression more than 2-fold in 2i versus serum. Genes upregulated in 2i show the expected increased H3K4me3 deposition at the promoter and higher levels of H3K36me3 in the coding body than in serum (Figure 6E and Figure S7E). The repressive mark H3K27me3 is correspondingly reduced. Upregulated genes in serum also show an increase in H3K36me3 over the coding body but in general do not exhibit a significant change in H3K4me3 deposition. More remarkably, on average they show increased H3K27me3.

We then examined Pol II occupancy at these two groups of genes. This showed that upregulation in 2i is reflected in elevated Pol II over the transcriptional start site as well as the coding body (Figure 6F, left). In contrast, genes upregulated in serum showed increased Pol II over the coding body but also on average a reduced signal at the start site (Figure 6F, right). These features indicate that transcriptional elongation at genes already loaded with Pol II is a widespread mechanism of upregulation in serum. RNA-seq data reveal no overt differences in expression of pTEFb components or known regulators of pausing between 2i and serum (Figure S7F). Regulation at the protein level may therefore control differential promoter proximal pause release in naive and metastable ES cells.

Multilineage Differentiation

ES cells maintained in either 2i or serum can colonize the mouse embryo, demonstrating that they are functionally pluripotent. However, they differ markedly in transcription of ectodermal and mesodermal specification genes. The precocious transcription of lineage-associated genes, often termed lineage priming (Hu et al., 1997), is posited to poise stem cells for differentiation. We therefore compared the differentiation behavior of ES cells maintained in 2i or serum. We first used a monolayer neural induction protocol with Sox1GFP reporter ES cells to quantify differentiation (Ying et al., 2003b). Although ES cells maintained

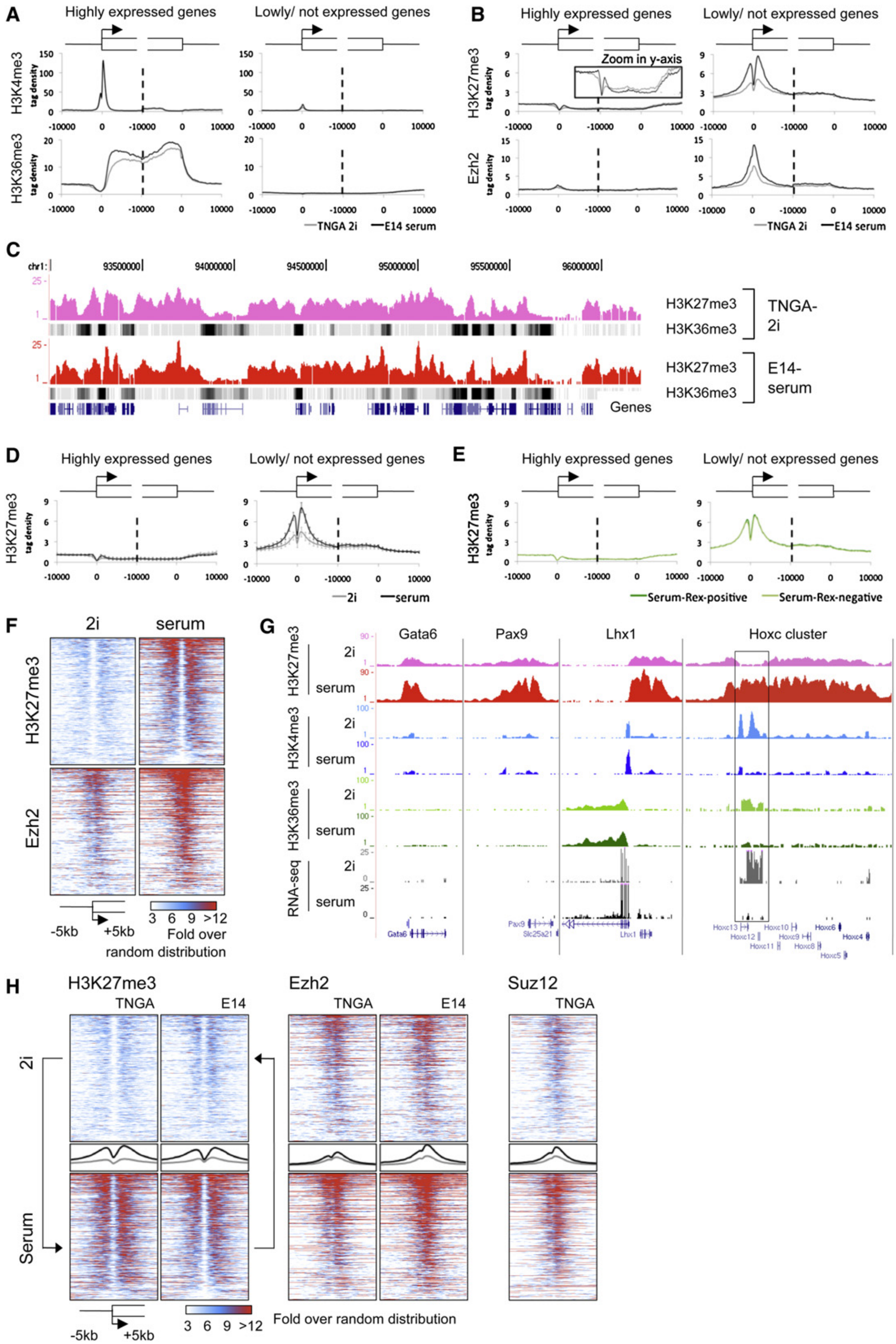


Figure 3. H3K27me3 Is Greatly Diminished at Promoters of Silent Genes and at Hox Clusters in 2i

(A) Average H3K4me3 and H3K36me3 profiles of the 2,000 most active (left plots) and 2,000 silent genes (right plots) from −10kb to +10kb at the transcription start site and the transcription stop. The negative control (genomic DNA) is shown in Figure S3B.

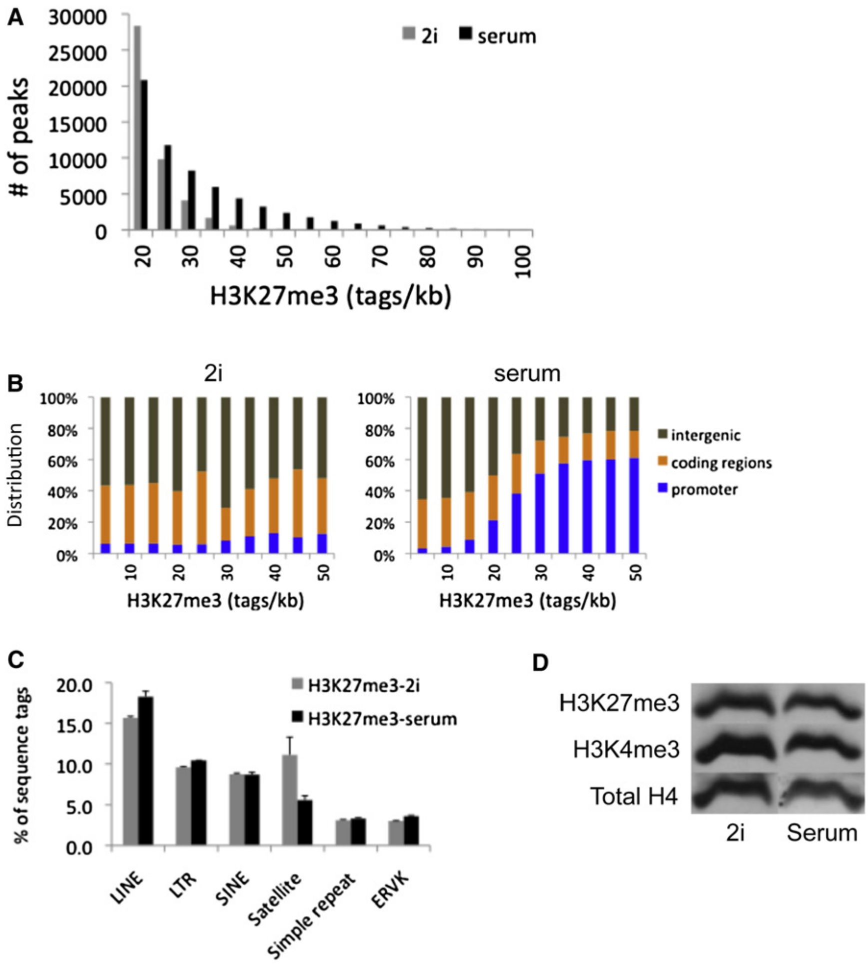


Figure 4. Localization and Quantification of H3K27me3 in 2i and Serum

(A) Binning of H3K27me3 -enriched regions in 2i and serum according to tag densities per peak (in reads/kb). The calculated genomic background is 5 reads/kb per 10 million mapped sequence reads (see Experimental Procedures).

(B) Genomic distribution of H3K27me3 peaks in 2i (left) or serum (right) per bin as shown in (A).

(C) Percentage of H3K27me3 reads present in the major repeat categories.

(D) Immunoblot analysis of histone modifications in total cell extracts.

See also Figure S5.

in serum express several neuroectodermal genes, they were less efficient in generating Sox1GFP-positive neural precursors than ES cells taken from 2i (Figure 7A). This could be due to the presence of cells already biased toward a mesodermal fate in serum. Clearly pre-expression of neural genes in serum does not predispose to this fate. We then used Rex1GFP fractionation to compare ES cell subpopulations in serum with 2i ES cells in embryoid body (EB) differentiation. Rex1GFP-positive cells from serum showed similar behavior to 2i cells (Figure 7B). Downregulation of *Nanog* and *Rex1* was followed by upregulation of the postimplantation epiblast marker *Fgf5*. From 3 days onward, *Fgf5* was downregulated and first *T* (*Brachyury*; although very minor as compared to the Rex1-negative cells), then mesoderm and endoderm lineage markers *Tbx6*, *Cxcr4*, *Sox17*, and *Gata4* appeared. This order is consistent with the developmental progression from blastocyst to gastrulation. In contrast, Rex1GFP-negative cells exhibit accelerated upregulation of *T* and *Tbx6* consistent with their partial differentiation and the loss of self-renewal (Figure S2D).

DISCUSSION

High-resolution genome-wide analyses have revealed that culture environments impose distinctive transcriptional and epigenomic properties on mouse ES cells. In total some 13,000 genes are transcribed at above background levels (>0.2 RPKM). The corollary of this is that around half of genes are effectively inactive. Therefore, undifferentiated ES cells do not show promiscuous gene expression or global transcriptional hyperactivity (Efroni et al., 2008). Nonetheless, the pluripotent transcriptome displays a broad bandwidth; more than 25% of active genes show 2-fold or greater differences between 2i and serum. Around 1,400 genes, predominantly associated with metabolic processes, are upregulated in 2i. In contrast, KEGG analysis points to decreased expression in 2i

(B) Average profiles of H3K27me3 and Ezh2 associated with silencing as in (A). Figure S3G shows a biological replicate analysis for the H3K27me3 profiles of TNGA 2i ES cells.

(C) Typical example of H3K27me3 and H3K36me3 (dense setting) profiles.

(D) H3K27me3 profiling for three different cell lines maintained and derived in either 2i or serum, as in (B). H3K27me3 profiles were generated for TNGA, NOD_male, and NOD_female ES cells in 2i and E14, HM1, and RGD2 ES cells in serum.

(E) As in (B); H3K27me3 profiling in Rex1-positive and Rex1-negative serum ES cells.

(F) H3K27me3 and Ezh2 intensity plots of all promoters containing H3K27me3 at >3-fold over random distribution in at least one of the conditions, TNGA-2i or E14-serum; 3,870 promoters are depicted in rows on the y axis.

(G) Typical examples showing H3K27me3 reduction in 2i as compared to serum.

(H) As in (F); promoter profiles for 2i ES cells adapted to serum and vice versa. Average profiles are plotted in the middle (gray: 2i; black: serum).

See also Figure S3, Figure S4, and Table S4.

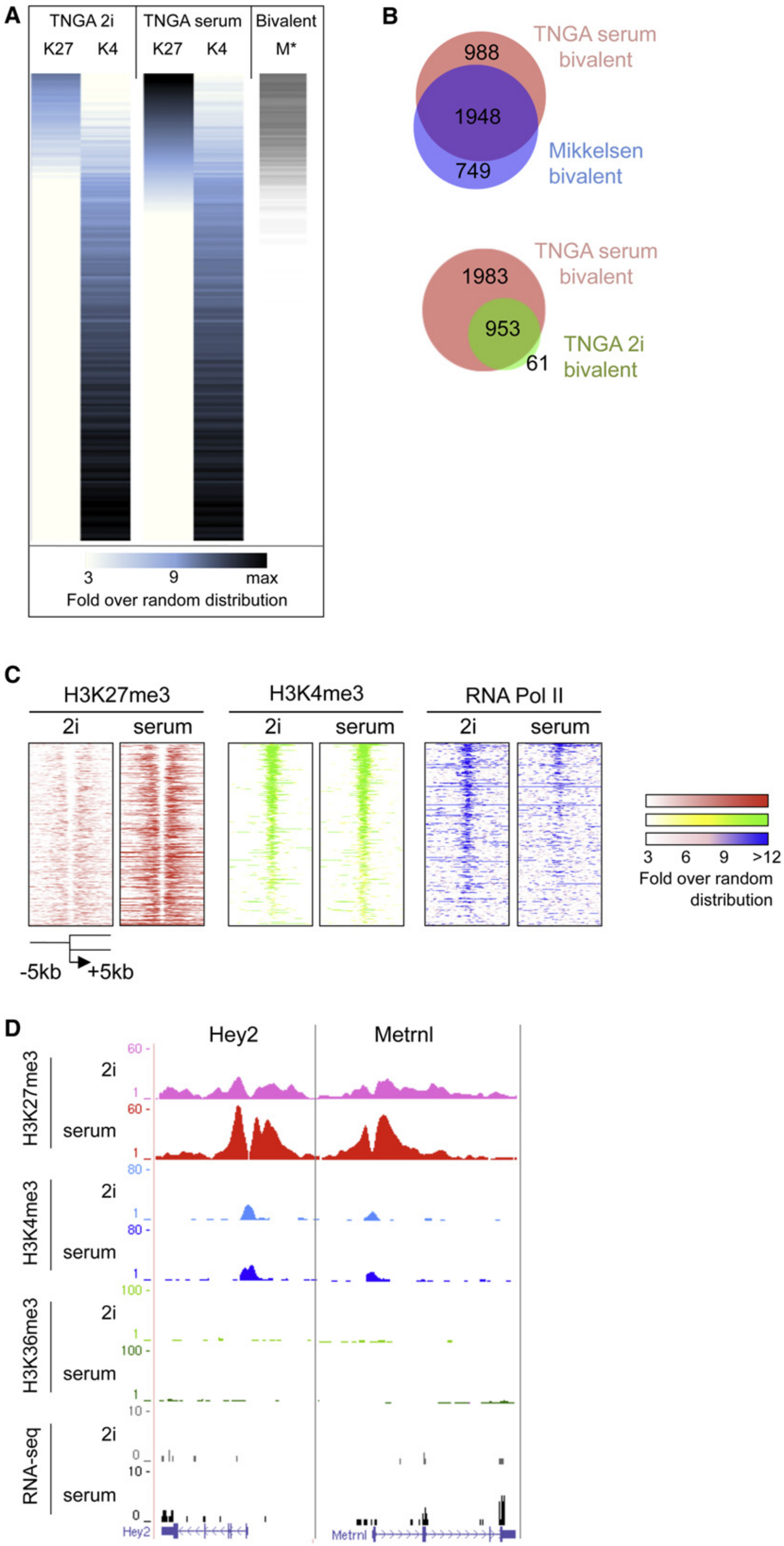


Figure 5. Bivalency in 2i and Serum

(A) H3K27me3 and H3K4me3 intensity plots at the promoters of all genes (sum of reads), ranked on highest to lowest H3K27me3 values in TNGA-serum (blue). Density plot of bivalent genes identified by Mikkelsen et al. (2007; "M") (gray).

(B) Overlap between the bivalent genes as described by Mikkelsen et al. (2007), and the bivalent genes in serum determined in this study (top) and between 2i and serum (bottom).

(C) H3K27me3, H3K4me3, and Pol II intensity plots of all promoters that are bivalent in TNGA serum (this study).

(D) Typical examples of bivalent genes.

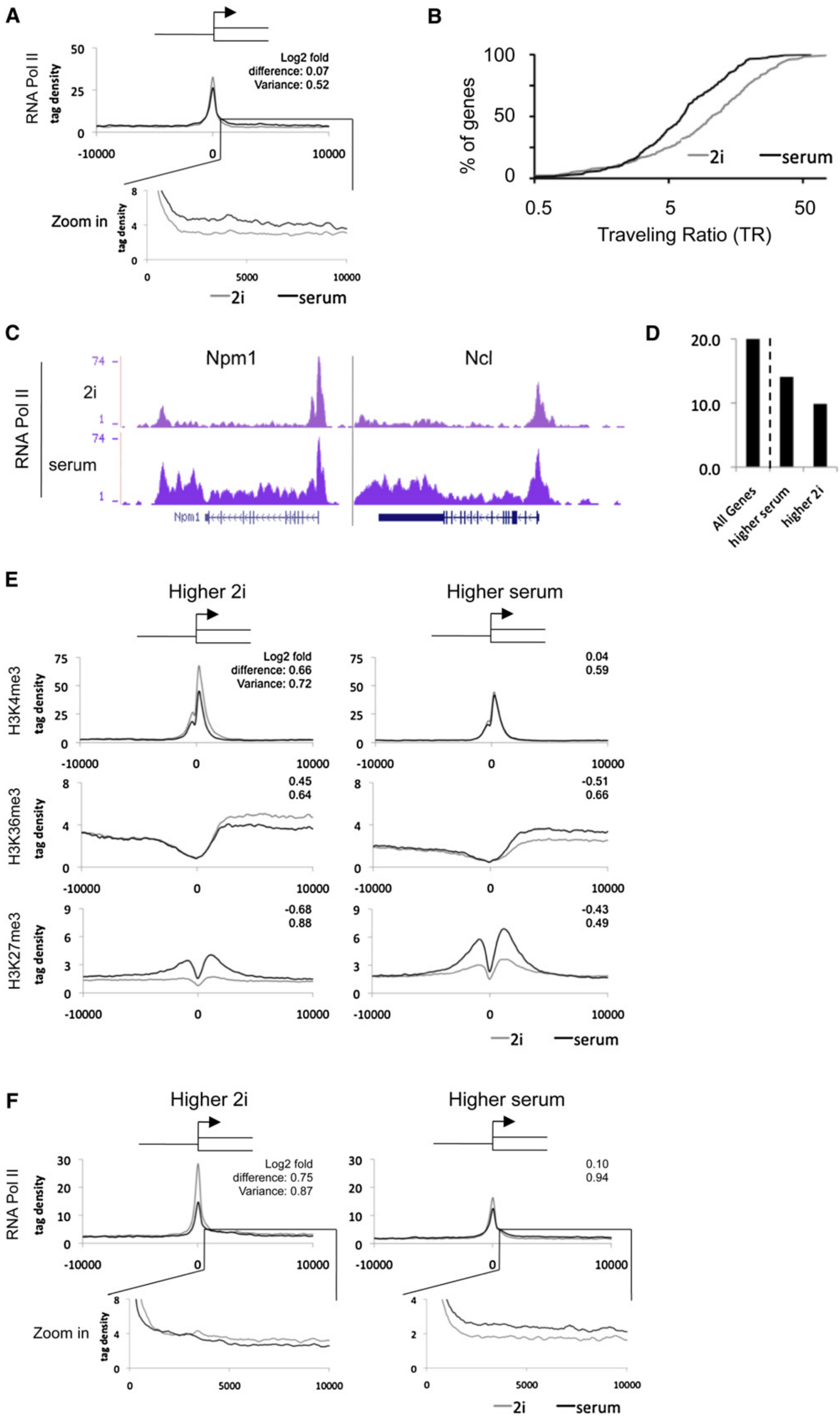
See also Figure S6 and Table S5.

of components that might drive differentiation, such as cell communication, mitogen-activated protein kinase (MAPK), and transforming growth factor β (TGFβ and Wnt) pathways. Most strikingly, many ectodermal and mesodermal specification genes that exhibit significant expression in serum are repressed in 2i. Low to absent lineage-affiliated gene expression indicates that multilineage priming is not an intrinsic feature of self-renewing ES cells. Upregulation of such genes in serum suggests that metastability may be an induced condition rather than an inherent property of pluripotent cells.

Some endodermal genes such as *Hex* retain low-level expression in 2i. This may reflect the potential to generate extraembryonic endoderm (Canham et al., 2010). High levels of Prdm14, which has been reported to repress extraembryonic endoderm transcription factors (Ma et al., 2011), may prevent full activation of this program.

Importantly, ES cells transferred between 2i and serum switch their transcriptional profile. Thus a significant component of previously described ES cell signatures reflects an induced serum response. However, critical pluripotency factors are transcribed at similar or only slightly higher levels in 2i. The pluripotency repressors Tcf3 (Wray et al., 2011) and components of the NuRD complex (Kaji et al., 2006) are also expressed at comparable levels. A subset of SCM factors are specifically upregulated in serum, including the *Id* genes that are induced by BMPs or fibronectin and are thought to directly counter the effects of Erk activation (Ying et al., 2003a). Increased Eras, shown to be important for ES cell propagation (Takahashi et al., 2003), and factors such as Sall4, Lin28, and Utf1, may also contribute to reinforcing self-renewal in the face of differentiation stimuli.

The conflict between pluripotency factors and lineage specifiers results in metastability and incipient differentiation in serum. It is suggested that this "precarious balance" (Loh and Lim, 2011) may reflect the circumstance in egg cylinder epiblast cells. However, serum stimulation is an artifactual scenario that may be far from representative of the spatiotemporal precision of


A
RNA Pol II
tag density
Log2 fold
difference: 0.07
Variance: 0.52
50
25
0
-10000
0
10000
Zoom in
tag density
8
4
0
5000
10000
2i
serum
B
% of genes
100
50
0
0.5
5
50
Traveling Ratio (TR)
2i
serum
C
RNA Pol II
Npm1
Ncl
2i
serum
74
1
Npm1
D
20.0
10.0
0.0
All Genes
higher serum
higher 2i
E
Higher 2i
Higher serum
H3K4me3
tag density
Log2 fold
difference: 0.66
Variance: 0.72
75
50
25
0
-10000
0
10000
0.04
0.59
H3K36me3
tag density
8
4
0.45
0.64
-0.51
0.66
H3K27me3
tag density
9
6
3
-0.68
0.88
-0.43
0.49
2i
serum
F
Higher 2i
Higher serum
RNA Pol II
tag density
30
20
10
0
Log2 fold
difference: 0.75
Variance: 0.87
0.10
0.94
-10000
0
10000
Zoom in
tag density
8
4
2
0
5000
10000
2i
serum

inductive stimuli in the embryo. To access postimplantation definitive lineages, ES cells should pass through a phase equivalent to egg cylinder epiblast (Rossant, 2008). Consistent with this, *Fgf5* is upregulated in EBs prior to definitive germ layer markers. From 2i ES cells and the Rex1GFP-positive fraction of serum ES cells, this process follows similar kinetics. Therefore although serum induces transcriptional and epigenetic changes and associated metastability, developmental potential within the Rex1-positive compartment is not fundamentally altered. This is substantiated by the capacity of ES cells from either condition to contribute extensively to chimaeras. However, a significant proportion of cells in serum lose expression of *Rex1* and of core pluripotency factors such as *Nanog* and *Klf4*. They are developmentally more advanced and should be considered distinct from ES cells even though they retain Oct4 (Smith, 2010).

Expression of many genes associated with metabolic and biosynthetic processes is enriched in 2i. This is likely in large part a response to absence of serum constituents, loss of MAPK signaling, and inhibition of GSK3 and indicates that ES cells have adaptable metabolomic capacity. Probably as a consequence of low c-Myc, the cell-cycle inhibitors p16, p19, and p21 are upregulated in 2i, even at the protein level (Figure S7D). Nonetheless, ES cells continue to proliferate rapidly, reflecting their freedom from cyclin checkpoint control (Burdon et al., 2002; Stead et al., 2002). These features can explain the robust expansion of ES cells independent of serum factors and likely underlie their latent tumorigenicity (Chambers and Smith, 2004).

In 2i and serum H3K4me3 peaks are globally similar in number and intensity. In contrast, there is a striking difference in the pattern of H3K27me3 deposition. This mark is present as a lawn across intergenic regions and inactive genes (Figure 3C). However, elevated deposits at promoters of repressed genes are greatly diminished in 2i. The majority of these genes show reduced rather than increased transcription in 2i. This promoter-specific diminution in H3K27me3 is common to multiple ES cell lines. The majority of these genes show reduced rather than increased transcription in 2i. Ezh2 is localized less intensely at promoters in 2i, which may underlie the selective reduction in H3K27me3. Global levels of H3K27me3 are similar in 2i and serum. Indeed, H3K27me3 is increased at satellites in 2i, indicating that these may serve as a sink. Notably, there is no change in H3K9me3 over satellites (Figure S3C).

In 2i only around 1,000 genes have both H3K4me3 and H3K27me3 marks, which argues against bivalency as a master epigenetic blueprint. Nonetheless, most of the remaining bivalent genes can be classified as developmental. In serum, more genes are bivalent due to acquisition of H3K27me3. Surprisingly, this is accompanied by a slight overall increase in expression, although the majority remain silent or transcribed at very low levels (Figure S6A). It is conceivable that although the local levels of PRC2 and H3K27me3 are reduced in 2i, they remain sufficient to repress transcription. It should be noted, however, that ES cells lacking PRC2, PRC1, or both are viable and show derepression of lineage-specific markers to only a low level (Leeb et al., 2010). Our findings are thus in line with genetic evidence that polycomb is not a central mechanism for silencing gene expression in the naive state and only becomes critical during differentiation.

RNA polymerase pausing has been identified by GRO-seq analysis (Min et al., 2011) at variable extents at many genes in ES cells cultured in serum. Our findings indicate that pausing is more prevalent in 2i than serum. Induction of c-Myc in serum may facilitate pause release at some loci. This is consistent with recent evidence that Myc function is unnecessary in naive ES cells but required in serum (Hishida et al., 2011). However, many of the genes whose expression is most markedly upregulated in serum, including germ layer specification factors, are not reported Myc targets. Therefore additional mechanisms are likely to control pause release in ES cells.

In mammals pluripotent cells harbor the germline and most pluripotency factors are also key players in germ cell specification and differentiation. It is interesting therefore that in *Caenorhabditis elegans* and *Drosophila*, germline development is dependent on transcriptional pausing mediated at the level of pTEFb antagonism by Pie-1 and Pgc, respectively (Nakamura and Seydoux, 2008). This raises the question of whether naive ES cells might contain an analogous factor that interferes with pTEFb to suppress transcriptional elongation. It will also be revealing to determine whether Erk signaling may cause activation of pTEFb (Fujita et al., 2008; Lee et al., 2010).

Recruitment and pausing of RNA polymerase II with lack of consolidated H3K27me3 silencing may constitute a potentiated template for induction of lineage-specific transcription programs. Pausing may serve to minimize the effects of noise and ensure rapid, coordinated, and synchronous gene induction in response to developmental cues or extrinsic stimuli (Boettiger and Levine, 2009; Nechaev and Adelman, 2011). Recent studies also indicate that Pol II pausing inhibits nucleosome assembly (Gilchrist et al., 2010) and could thereby influence histone modification profile. Interestingly, in *Xenopus* embryos H3K27me3 is not deposited during zygotic gene activation but is acquired later and associated with spatial restriction of gene expression (Akkers et al., 2009). In the mouse ICM, H3K27me localization has not been determined, but various epigenetic silencing

Figure 6. RNA Polymerase II Pausing in Naive ES Cells

(A) Averaged Pol II at promoters of Myc-targets upregulated in serum. The top corner values represent the average $\log_2$-fold difference of the individual data points, and the variance, between 2i and serum.

(B) Pol II traveling ratio (a quantification of pause release) of the Myc targets upregulated in serum.

(C) Typical examples of two Myc target genes showing pause release of Pol II.

(D) Percentage of Myc targets among all genes (left) or among the genes higher (>2-fold change) in serum or 2i (right).

(E) Averaged profiles of the promoter region of genes that are more highly expressed in 2i (left) or serum (right).

(F) Averaged Pol II profiles for the same gene groups as in (E).

See also Figure S7.

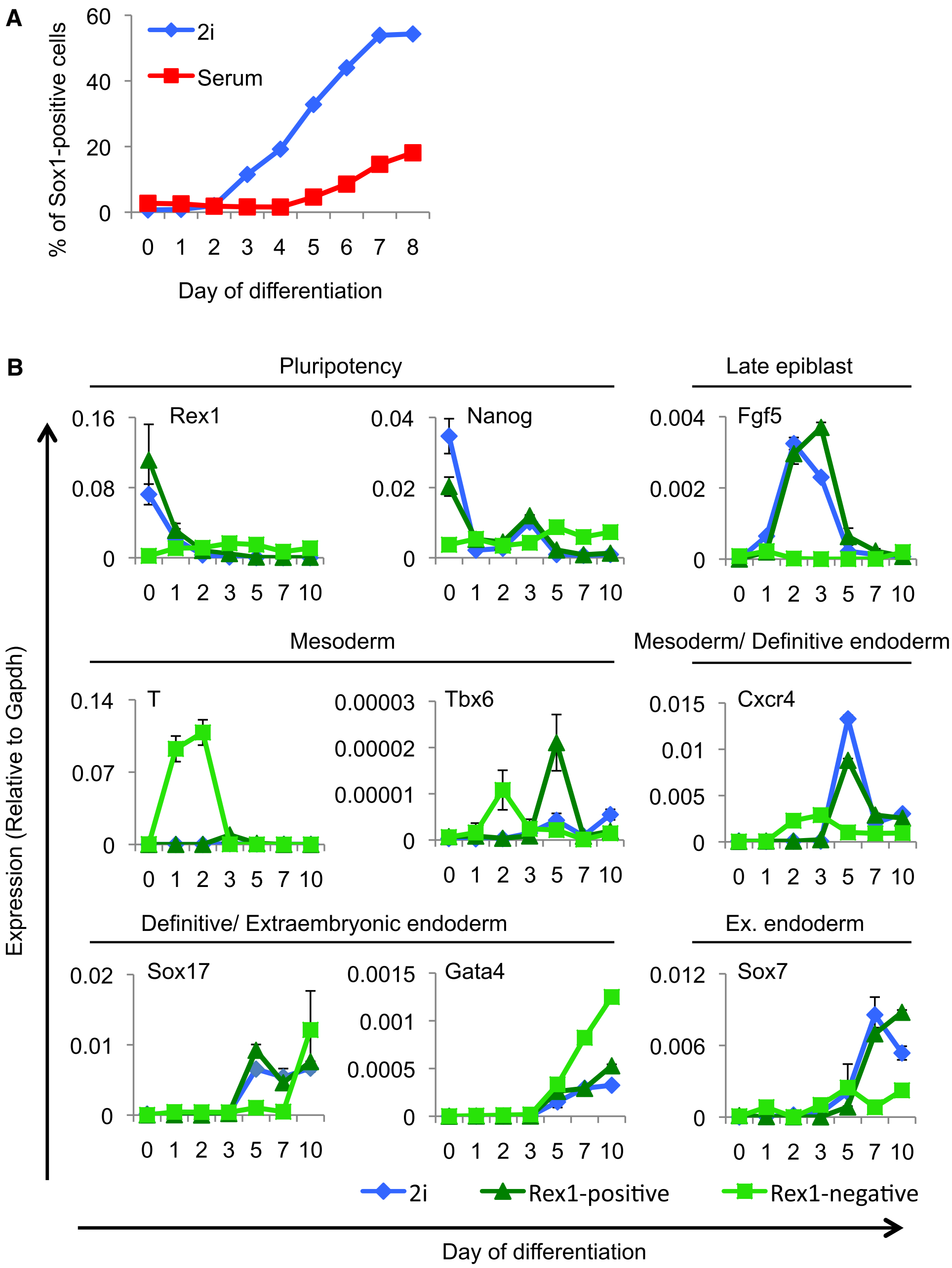


Figure 7. Differentiation Kinetics/Potential of 2i and Serum ES Cells

(A) Monolayer neural differentiation of Sox1-GFP ES cells maintained in either 2i or serum.

(B) Embryoid body differentiation of cells maintained in 2i, and of Rex1-positive and Rex1-negative serum ES cells as sorted by FACS (Figure S2C). Expression levels were determined by RT-qPCR.

components appear to be expressed at low levels (Tang et al., 2010).

Collectively the observations reported here yield insights into the molecular underpinning of naive pluripotency and revise previous assumptions derived from analysis of heterogeneous and metastable serum-treated cultures. The findings provoke questions about the regulation of gene expression in pluripotent cells and the process of lineage specification. Transcriptional potentiation through promoter proximal pausing may play a major role in the establishment and stable maintenance of naive pluripotency. Currently there is great interest in isolating human pluripotent stem cells in a naive state (Hanna et al., 2010; Wang et al., 2011). The distinctive transcriptome and epigenome characteristics of ground state mouse ES cells may provide a valuable criterion against which to measure such claims. In addition, these data sets provide a benchmark resource for analysis and modeling of gene expression control during self-renewal and in the transition from naive pluripotency to lineage commitment.

EXPERIMENTAL PROCEDURES

Cell Culture and Methods

ES cells were cultured without feeders in the presence of leukemia inhibitory factor (LIF) either in Glasgow modification of Eagles medium (GMEM) containing 10% fetal calf serum or in serum-free N2B27 supplemented with MEK inhibitor PD0325901 (1 μM) and GSK3 inhibitor CH99021 (3 μM), together known as 2i (Ying et al., 2008). E14Tg2a (E14), Rex1GFPd2 (RGD2, Rex1GFP), and HM1 are male ES cells of 129 background established and maintained in serum without feeders. The serum-derived female XT67E1 line is from a mixed 129 and PGK/C3H background (Penny et al., 1996). TNGA female ES cells were derived and maintained in 2i from embryos on a mixed strain 129 and C57BL/6 background heterozygous for eGFP knock-in at the *Nanog* gene (Chambers et al., 2007). Female and male ES cells from the nonobese diabetic (NOD) strain were derived and maintained in 2i (Nichols et al., 2009). Chromatin immunoprecipitations (ChIP) were performed as described (Marks et al., 2009). Cell sorting (fluorescence-activated cell sorting [FACS]), immunoblotting, RNA isolation, and cDNA synthesis were performed according to standard protocols described in the Extended Experimental Procedures, which also lists antibodies used.

Sequencing

DNA samples were prepared for sequencing by end repair of 20 ng DNA as measured by Qubit (Invitrogen). Adaptors were ligated to DNA fragments, followed by size selection (~300 bp) and 14 cycles of PCR amplification. Cluster generation and sequencing (36 bp) was performed with the Illumina Genome Analyzer IIx (GAIIx) platform according to standard Illumina protocols. The standard pipeline to generate the sequencing output files are described in the Supplemental Information. All sequencing analyses were conducted based on the *Mus musculus* NCBI m37 genome assembly (MM9; assembly July 2007). Table S1 summarizes the sequencing output. All RNA-seq and ChIP-seq data (FASTQ, BED, and WIG files) are present in the NCBI GEO SuperSeries GSE23943.

RNA-Seq Analysis

To obtain RNA-Seq gene expression values (RPKM), we used Genomatix (www.genomatix.de). Differential genes were called at a 2-fold difference and $p < 0.2$ in a Student t test (among three biological replicates for both 2i and serum). The identification of ncRNAs is described in the Extended Experimental Procedures. GO and KEGG analysis was performed with DAVID (http://david.abcc.ncifcrf.gov/).

ChIP-Seq Analyses

To compensate for differences in sequencing depth and mapping efficiency, the total number of unique reads of each sample was uniformly equalized, allowing quantitative comparisons. Tag densities on the average profiles were determined by calculating tag density over each base pair (using a 40 bp window size) per 10 million total mapped reads. The genes used for the average epigenetic profiles were based on the 2,000 most active/inactive genes in TNGA-2i. The 2,000 lowest/not expressed genes were selected by the additional requirement of H3K27me3 promoter enrichments of > 3-fold over background in either TNGA-2i or E14-serum. Random distribution values were determined by calculating average read densities of the genomic DNA profile (4.962 reads/kb at 10 million sequenced reads equivalent to an average density of 1.489 per bp). Genes were considered bivalent if both H3K4me3 and H3K27me3 were > 3-fold over random distribution, similar to criteria applied by Mikkelsen et al. (2007). The RNA Polymerase II traveling was calculated as described by Rahl et al. (2010). The Chip-seq repeat analysis procedure is described in the Supplemental Information.

Differentiation Assays

For monolayer neural differentiation, we used Sox1-GFP (46C) ES cells, which contain a GFP knock-in at the endogenous Sox1 locus (Ying et al., 2003b). Cells cultured in 2i or serum were plated at a density of 5,000 cells per cm^2. Sixteen hours after plating, media were switched to N2B27 to induce neural differentiation. Percentage of GFP-positive cells was determined by flow cytometry. For EB differentiation, single EBs were formed by sorting 1,500 cells into each well of PrimeSurface96U plates containing 15% serum and no LIF. Sixteen EBs were pooled each day and analyzed by RT-qPCR with TaqMan probes (Applied Biosystems).

ACCESSION NUMBERS

The GEO accession number for the SuperSeries reported in this paper is GSE23943.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and six tables and can be found with this article online at doi:10.1016/j.cell.2012.03.026.

ACKNOWLEDGMENTS

We thank Eva Janssen-Megens, Anita Kaan, and Yan Tan for sequencing; Kees-Jan Françoijs, Hinri Kerstens, Simon van Heeringen, and Arjen Brinkman for bioinformatic assistance; and Rachel Walker for help with flow cytometry. We thank Paul Bertone for comments on the manuscript. Tatyana Nesterova and Neil Brockdorff provided the XT67E1 cells and Olga Ujhelly and Andras Dinnyes the HM1 cells. The pThr345-Ezh2 antibody was a gift from Danny Reinberg. The research leading to these results has received funding from the European Union grants HEROIC (18883; FP6/2005-2010), PluriSys (223485; FP7/2009), EuroSyStem (200720; FP7/2008), and ATLAS (221952; FP7/2009), the Deutsche Forschungsgemeinschaft SPP1356 Pluripotency and Reprogramming, and The Wellcome Trust. A.S. is a Medical Research Council Professor.

Received: July 9, 2011
Revised: December 26, 2011
Accepted: March 6, 2012
Published: April 26, 2012

REFERENCES

Akkers, R.C., van Heeringen, S.J., Jacobi, U.G., Janssen-Megens, E.M., Françoijs, K.J., Stunnenberg, H.G., and Veenstra, G.J. (2009). A hierarchy of H3K4me3 and H3K27me3 acquisition in spatial gene regulation in Xenopus embryos. Dev. Cell *17*, 425–434.

Azuara, V., Perry, P., Sauer, S., Spivakov, M., Jørgensen, H.F., John, R.M., Gouti, M., Casanova, M., Warnes, G., Merkenschlager, M., and Fisher, A.G. (2006). Chromatin signatures of pluripotent cell lines. Nat. Cell Biol. *8*, 532–538.

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., et al. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell *125*, 315–326.

Boettiger, A.N., and Levine, M. (2009). Synchronous and stochastic patterns of gene activation in the Drosophila embryo. Science *325*, 471–473.

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

Burdon, T., Smith, A., and Savatier, P. (2002). Signalling, cell cycle and pluripotency in embryonic stem cells. Trends Cell Biol. *12*, 432–438.

Canham, M.A., Sharov, A.A., Ko, M.S., and Brickman, J.M. (2010). Functional heterogeneity of embryonic stem cells revealed through translational amplification of an early endodermal transcript. PLoS Biol. *8*, e1000379.

Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R.S., and Zhang, Y. (2002). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science *298*, 1039–1043.

Chambers, I., and Smith, A. (2004). Self-renewal of teratocarcinoma and embryonic stem cells. Oncogene *23*, 7150–7160.

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

Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V.B., Wong, E., Orlov, Y.L., Zhang, W., Jiang, J., et al. (2008). Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. Cell *133*, 1106–1117.

Efroni, S., Duttagupta, R., Cheng, J., Dehghani, H., Hoeppner, D.J., Dash, C., Bazett-Jones, D.P., Le Grice, S., McKay, R.D., Buetow, K.H., et al. (2008). Global transcription in pluripotent embryonic stem cells. Cell Stem Cell *2*, 437–447.

Fujita, T., Ryser, S., Piuz, I., and Schlegel, W. (2008). Up-regulation of P-TEFb by the MEK1-extracellular signal-regulated kinase signaling pathway contributes to stimulated transcription elongation of immediate early genes in neuroendocrine cells. Mol. Cell. Biol. *28*, 1630–1643.

Gilchrist, D.A., Dos Santos, G., Fargo, D.C., Xie, B., Gao, Y., Li, L., and Adelman, K. (2010). Pausing of RNA polymerase II disrupts DNA-specified nucleosome organization to enable precise gene regulation. Cell *143*, 540–551.

Graf, T., and Stadtfeld, M. (2008). Heterogeneity of embryonic and adult stem cells. Cell Stem Cell *3*, 480–483.

Guo, G., Huss, M., Tong, G.Q., Wang, C., Li Sun, L., Clarke, N.D., and Robson, P. (2010). Resolution of cell fate decisions revealed by single-cell gene expression analysis from zygote to blastocyst. Dev. Cell *18*, 675–685.

Guttman, M., Garber, M., Levin, J.Z., Donaghey, J., Robinson, J., Adiconis, X., Fan, L., Koziol, M.J., Gnirke, A., Nusbaum, C., et al. (2010). Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. Nat. Biotechnol. *28*, 503–510.

Hanna, J., Markoulaki, S., Mitalipova, M., Cheng, A.W., Cassady, J.P., Staerk, J., Carey, B.W., Lengner, C.J., Foreman, R., Love, J., et al. (2009). Metastable pluripotent states in NOD-mouse-derived ESCs. Cell Stem Cell *4*, 513–524.

Hanna, J., Cheng, A.W., Saha, K., Kim, J., Lengner, C.J., Soldner, F., Cassady, J.P., Muffat, J., Carey, B.W., and Jaenisch, R. (2010). Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. Proc. Natl. Acad. Sci. USA *107*, 9222–9227.

Hayashi, K., Lopes, S.M., Tang, F., and Surani, M.A. (2008). Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states. Cell Stem Cell *3*, 391–401.

Hishida, T., Nozaki, Y., Nakachi, Y., Mizuno, Y., Okazaki, Y., Ema, M., Takahashi, S., Nishimoto, M., and Okuda, A. (2011). Indefinite self-renewal of ESCs through Myc/Max transcriptional complex-independent mechanisms. Cell Stem Cell *9*, 37–49.

Hu, M., Krause, D., Greaves, M., Sharkis, S., Dexter, M., Heyworth, C., and Enver, T. (1997). Multilineage gene expression precedes commitment in the hemopoietic system. Genes Dev. *11*, 774–785.

Ivanova, N., Dobrin, R., Lu, R., Kotenko, I., Levorse, J., DeCoste, C., Schafer, X., Lun, Y., and Lemischka, I.R. (2006). Dissecting self-renewal in stem cells with RNA interference. Nature *442*, 533–538.

Kaji, K., Caballero, I.M., MacLeod, R., Nichols, J., Wilson, V.A., and Hendrich, B. (2006). The NuRD component Mbd3 is required for pluripotency of embryonic stem cells. Nat. Cell Biol. *8*, 285–292.

Kalmar, T., Lim, C., Hayward, P., Muñoz-Descalzo, S., Nichols, J., Garcia-Ojalvo, J., and Martinez Arias, A. (2009). Regulated fluctuations in nanog expression mediate cell fate decisions in embryonic stem cells. PLoS Biol. *7*, e1000149.

Kaneko, S., Li, G., Son, J., Xu, C.F., Margueron, R., Neubert, T.A., and Reinberg, D. (2010). Phosphorylation of the PRC2 component Ezh2 is cell cycle-regulated and up-regulates its binding to ncRNA. Genes Dev. *24*, 2615–2620.

Kiyonari, H., Kaneko, M., Abe, S., and Aizawa, S. (2010). Three inhibitors of FGF receptor, ERK, and GSK3 establishes germline-competent embryonic stem cells of C57BL/6N mouse strain with high efficiency and stability. Genesis *48*, 317–327.

Kunath, T., Saba-El-Leil, M.K., Almousailleakh, M., Wray, J., Meloche, S., and Smith, A. (2007). FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. Development *134*, 2895–2902.

Lee, J.S., Smith, E., and Shilatifard, A. (2010). The language of histone crosstalk. Cell *142*, 682–685.

Leeb, M., Pasini, D., Novatchkova, M., Jaritz, M., Helin, K., and Wutz, A. (2010). Polycomb complexes act redundantly to repress genomic repeats and genes. Genes Dev. *24*, 265–276.

Li, P., Tong, C., Mehrian-Shai, R., Jia, L., Wu, N., Yan, Y., Maxson, R.E., Schulze, E.N., Song, H., Hsieh, C.L., et al. (2008). Germline competent embryonic stem cells derived from rat blastocysts. Cell *135*, 1299–1310.

Loh, K.M., and Lim, B. (2011). A precarious balance: pluripotency factors as lineage specifiers. Cell Stem Cell *8*, 363–369.

Ma, Z., Swigut, T., Valouev, A., Rada-Iglesias, A., and Wysocka, J. (2011). Sequence-specific regulator Prdm14 safeguards mouse ESCs from entering extraembryonic endoderm fates. Nat. Struct. Mol. Biol. *18*, 120–127.

Marks, H., Chow, J.C., Denissov, S., Françoijs, K.J., Brockdorff, N., Heard, E., and Stunnenberg, H.G. (2009). High-resolution analysis of epigenetic changes associated with X inactivation. Genome Res. *19*, 1361–1373.

Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.K., Koche, R.P., et al. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature *448*, 553–560.

Min, I.M., Waterfall, J.J., Core, L.J., Munroe, R.J., Schimenti, J., and Lis, J.T. (2011). Regulating RNA polymerase pausing and transcription elongation in embryonic stem cells. Genes Dev. *25*, 742–754.

Nakamura, A., and Seydoux, G. (2008). Less is more: specification of the germline by transcriptional repression. Development *135*, 3817–3827.

Nechaev, S., and Adelman, K. (2011). Pol II waiting in the starting gates: Regulating the transition from transcription initiation into productive elongation. Biochim. Biophys. Acta *1809*, 34–45.

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

Niwa, H., Ogawa, K., Shimosato, D., and Adachi, K. (2009). A parallel circuit of LIF signalling pathways maintains pluripotency of mouse ES cells. Nature *460*, 118–122.

Papp, B., and Müller, J. (2006). Histone trimethylation and the maintenance of transcriptional ON and OFF states by trxG and PcG proteins. Genes Dev. *20*, 2041–2054.

Penny, G.D., Kay, G.F., Sheardown, S.A., Rastan, S., and Brockdorff, N. (1996). Requirement for Xist in X chromosome inactivation. Nature *379*, 131–137.

Rahl, P.B., Lin, C.Y., Seila, A.C., Flynn, R.A., McCuine, S., Burge, C.B., Sharp, P.A., and Young, R.A. (2010). c-Myc regulates transcriptional pause release. Cell *141*, 432–445.

Rinn, J.L., Kertesz, M., Wang, J.K., Squazzo, S.L., Xu, X., Brugmann, S.A., Goodnough, L.H., Helms, J.A., Farnham, P.J., Segal, E., and Chang, H.Y. (2007). Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell *129*, 1311–1323.

Rossant, J. (2008). Stem cells and early lineage development. Cell *132*, 527–531.

Schmitges, F.W., Prusty, A.B., Faty, M., Stützer, A., Lingaraju, G.M., Aiwazian, J., Sack, R., Hess, D., Li, L., Zhou, S., et al. (2011). Histone methylation by PRC2 is inhibited by active chromatin marks. Mol. Cell *42*, 330–341.

Smith, A.G. (2001). Embryo-derived stem cells: of mice and men. Annu. Rev. Cell Dev. Biol. *17*, 435–462.

Smith, A. (2010). Pluripotent stem cells: private obsession and public expectation. EMBO Mol. Med. *2*, 113–116.

Srinivasan, S., Dorighi, K.M., and Tamkun, J.W. (2008). Drosophila Kismet regulates histone H3 lysine 27 methylation and early elongation by RNA polymerase II. PLoS Genet. *4*, e1000217.

Stead, E., White, J., Faast, R., Conn, S., Goldstone, S., Rathjen, J., Dhingra, U., Rathjen, P., Walker, D., and Dalton, S. (2002). Pluripotent cell division cycles are driven by ectopic Cdk2, cyclin A/E and E2F activities. Oncogene *21*, 8320–8333.

Stock, J.K., Giadrossi, S., Casanova, M., Brookes, E., Vidal, M., Koseki, H., Brockdorff, N., Fisher, A.G., and Pombo, A. (2007). Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells. Nat. Cell Biol. *9*, 1428–1435.

Takahashi, K., Mitsui, K., and Yamanaka, S. (2003). Role of ERas in promoting tumour-like properties in mouse embryonic stem cells. Nature *423*, 541–545.

Tang, F., Barbacioru, C., Bao, S., Lee, C., Nordman, E., Wang, X., Lao, K., and Surani, M.A. (2010). Tracing the derivation of embryonic stem cells from the inner cell mass by single-cell RNA-Seq analysis. Cell Stem Cell *6*, 468–478.

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

Wang, W., Yang, J., Liu, H., Lu, D., Chen, X., Zenonos, Z., Campos, L.S., Rad, R., Guo, G., Zhang, S., et al. (2011). Rapid and efficient reprogramming of somatic cells to induced pluripotent stem cells by retinoic acid receptor gamma and liver receptor homolog 1. Proc. Natl. Acad. Sci. USA *108*, 18283–18288.

Wray, J., Kalkan, T., and Smith, A.G. (2010). The ground state of pluripotency. Biochem. Soc. Trans. *38*, 1027–1032.

Wray, J., Kalkan, T., Gomez-Lopez, S., Eckardt, D., Cook, A., Kemler, R., and Smith, A. (2011). Inhibition of glycogen synthase kinase-3 alleviates Tcf3 repression of the pluripotency network and increases embryonic stem cell resistance to differentiation. Nat. Cell Biol. *13*, 838–845.

Ying, Q.L., Nichols, J., Chambers, I., and Smith, A. (2003a). BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. Cell *115*, 281–292.

Ying, Q.L., Stavridis, M., Griffiths, D., Li, M., and Smith, A. (2003b). Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture. Nat. Biotechnol. *21*, 183–186.

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.

Young, R.A. (2011). Control of the embryonic stem cell state. Cell *144*, 940–954.

npj | Systems Biology and Applications

www.nature.com/npjsba

ARTICLE OPEN

An extended model for culture-dependent heterogenous gene expression and proliferation dynamics in mouse embryonic stem cells

Simon Godwin[1], Daniel Ward[1], Elisa Pedone[1,2], Martin Homer[1], Alexander G. Fletcher[3,4] and Lucia Marucci[1,2,5]

During development, pluripotency is a transient state describing a cell's ability to give rise to all three germ layers and germline. Recent studies have shown that, in vitro, pluripotency is highly dynamic: exogenous stimuli provided to cultures of mouse embryonic stem cells, isolated from pre-implantation blastocysts, significantly affect the spectrum of pluripotency. 2i/LIF, a recently defined serum-free medium, forces mouse embryonic stem cells into a ground-state of pluripotency, while serum/LIF cultures promote the co-existence of ground-like and primed-like mouse embryonic stem cell subpopulations. The latter heterogeneity correlates with temporal fluctuations of pluripotency markers, including the master regulator Nanog, in single cells. We propose a mathematical model of Nanog dynamics in both media, accounting for recent experimental data showing the persistence of a small Nanog Low subpopulation in ground-state pluripotency mouse embryonic stem cell cultures. The model integrates into the core pluripotency Gene Regulatory Network both inhibitors present in 2i/LIF (PD and Chiron), and feedback interactions with genes found to be differentially expressed in the two media. Our simulations and bifurcation analysis show that, in ground-state cultures, Nanog dynamics result from the combination of reduced noise in gene expression and the shift of the system towards a monostable, but still excitable, regulation. Experimental data and agent-based modelling simulations indicate that mouse embryonic stem cell proliferation dynamics vary in the two media, and cannot be reproduced by accounting only for Nanog-dependent cell-cycle regulation. We further demonstrate that both PD and Chiron play a key role in regulating heterogeneity in transcription factor expression and, ultimately, mouse embryonic stem cell fate decision.

npj Systems Biology and Applications (2017)3:19 ; doi:10.1038/s41540-017-0020-5

INTRODUCTION

Mouse embryonic stem cells (mESCs) are pluripotent cells, isolated from the inner cell mass, which can be indefinitely expanded and retain pluripotency, or be pushed into specific differentiated states by proper stimuli in vitro, contributing to all germ layers when injected into host embryos.[1] In recent years, significant research effort has been put into defining optimal culture conditions for pluripotency maintenance, and identifying protocols for efficient differentiation of mESCs.

The standard mESC culture medium is serum/LIF; it contains serum factors and the cytokine leukaemia inhibitory factor (LIF). This medium has been reported to confer on mESCs a heterogeneous expression and temporal fluctuations of pluripotency factors and regulators, including Nanog, Rex1, Stella, Esrrb and β-catenin.[2–6] Notably, many of the metastable mESC genes are, directly or indirectly, regulated by Nanog,[7] a master regulator of pluripotency and development.[8] Importantly, mosaic expression patterns in serum/LIF result in an inhomogeneous propensity to differentiate in mESC subpopulations. This phenotype is reversible however: cells expressing Nanog Low levels, which are more prone to differentiate, can in time switch Nanog on, and vice versa for Nanog High cells.[3]

Such results prompted the identification of new culture conditions able to reduce heterogeneity and allow a standardised pluripotency phenotype. Recently, a new culture medium (2i/LIF) has been proposed[9]; it is serum-free and contains the two chemical inhibitors PD0325901 (MEK inhibitor,[10] hereafter named PD) and CHIR99021 (glycogen synthase kinase-3 (Gsk3) inhibitor,[11] hereafter named Chiron). When cultured in 2i/LIF, mESC reporter cell lines commonly used to monitor expression of Nanog and its direct target Rex1 (TNGA, which carries a stable green fluorescent protein (GFP) in one of the Nanog alleles,[3] and Rex1GFPd2, in which a destabilised GFP is inserted into the Rex1 locus,[12] respectively) show an almost complete elimination of the Nanog Low and Rex1 Low subpopulations observed under serum/LIF cultures, suggesting that 2i/LIF enables ground state pluripotency in vitro.[13]

The observed abolition of Nanog heterogeneity in 2i/LIF has recently been challenged,[14, 15] as well as the reliability of TNGA mESCs as a reporter cell line, due to differences between endogenous Nanog and GFP degradation rates.[16] A more recently-engineered Nanog reporter mESC cell line (Nd mESC) carries, under the Nanog regulatory regions, a destabilised Venus protein (Venus-Nuclear-PEST) with mRNA and protein half-lives comparable to endogenous Nanog.[17] Of note, while enabling a

[1]Department of Engineering Mathematics, University of Bristol, Bristol BS8 1UB, UK; [2]School of Cellular & Molecular Medicine, University of Bristol, Bristol BS8 1TD, UK; [3]School of Mathematics and Statistics, University of Sheffield, Sheffield S3 7RH, UK; [4]Bateson Centre, University of Sheffield, Sheffield S10 2TN, UK and [5]BrisSynBio, University of Bristol, Bristol BS8 1TQ, UK
Correspondence: Lucia Marucci (lucia.marucci@bristol.ac.uk)

Received: 25 September 2016 Revised: 31 May 2017 Accepted: 20 June 2017
Published online: 03 August 2017

highly dynamic Nanog readout, Nd mESCs have both Nanog alleles intact. As compared to TNGA mESCs, Nanog is still bistable in serum/LIF Nd cells, although with a smaller difference between the two Nanog Low and Nanog High (named NL and NH hereafter) subpopulations. Also, sorted cells re-establish the original bimodal distribution on a shorter time-scale.[15, 17] In ground-state cultures, a small NL subpopulation is still present in Nd mESCs; consistently, stochastic fluctuations between the two states can be observed within a single cell-cycle (CC), with the same amplitude (i.e., difference between maximum and minimum fluorescence levels measured in single-cell time-lapse experiments) observed in serum/LIF cultured cells.[15]

Diverse molecular mechanisms have been proposed to explain the divergent dynamics in the two media, including Nanog transcriptional bursting,[18] epigenetic mechanisms[19] and micro-RNA regulation.[20] Alongside these mechanisms, feedbacks in Gene Regulatory Networks (GRNs) of core pluripotency and differentiation factors play a key role in determining transcriptional dynamics and decision-making of mESCs.[21]

Mathematical modelling can be a powerful tool, not only to formalise dynamics observed experimentally, but also to suggest ad hoc perturbations to influence fate decision in mESCs.[22] Different mathematical formalisms have been proposed to describe mESC culture-dependent transcriptional dynamics (see ref. 22 for a review), focusing on Nanog regulations. Although based on different hypotheses about the origin of heterogeneity and fluctuations of pluripotency genes in serum/LIF cultures, all these models fail to predict the above-mentioned recent findings of the persistence of a NL subpopulation in 2i/LIF.

Furthermore, in the systems biology literature we are not aware of any mathematical attempt to explain and compare proliferation dynamics of mESCs in the two media. cell-cycle progression and stem cell commitment have been shown to be linked processes[23]: pluripotent mESCs display a unique and singular cell-cycle defined by a fast proliferation rate, long S/G2 phase and short G1 phase.[24] Conversely, during differentiation, the G1 phase duration increases in mESCs, hESCs, mouse and human iPSCs.[25, 26] It has thus been proposed that a long G1 phase plays a key role in differentiation, while a short G1 phase might be involved in actively sustaining the pluripotent state.[26] On the other hand, accumulation of mESCs in the G1 phase by inhibition of Cdk2 reduces their proliferation but does not affect pluripotency.[27] Agent-based models, that incorporate GRN dynamics into cellular agents, can aid in unravelling the coupling between gene expression and cellular phenotypes in both systems and synthetic biology,[28, 29] and help to explain the link between mESC proliferation and pluripotency. The only existing attempt in this direction recapitulates Nanog heterogeneity and sorting experiments of mESCs cultured in serum/LIF, but the model does not analyse the 2i/LIF scenario.[30]

Here, by mathematical modelling, we are able to recapitulate both the aforementioned persistence of a NL subpopulation in 2i/LIF, and mESC proliferation dynamics in both serum/LIF and 2i/LIF. We modelled the dynamics of an extended GRN, which includes genes differentially expressed in the two media, their transcriptional and posttranslational mutual interactions and the two inhibitors present in 2i/LIF. Numerical simulation and continuation results of our stochastic differential equation (SDE) model recapitulate existing data obtained with Nd mESCs, and indicate which of the network components are required to reproduce the experimental data. Agent-based simulations, calibrated to match mESC proliferation dynamics we measured experimentally, show that 2i/LIF cells possess a lower proliferation capacity than their serum/LIF counterpart, and that differences in Nanog levels alone cannot explain this phenotype, while Mycn qualifies as a regulator of cell-cycle progression. Overall, our results indicate that the complex interplay between feedback regulations, transcriptional noise, and the cell-cycle determines mESC dynamics and pluripotency, and that excitability might be an inherent feature of pluripotent cells.

RESULTS

Network derivation: identification of transcriptional core interactions

We started from an existing GRN-based model,[31] which formalised interactions among the core pluripotency transcription factors (Oct4, Sox2, Nanog and Rex1) to recapitulate mESC dynamics in serum/LIF and 2i/LIF media. The original network included Oct4-Sox2 heterodimer auto-activation, Oct4-Sox2 activation of Nanog, Nanog auto-activation, and the activation of Rex1 by both Nanog and the Oct4-Sox2 heterodimer. To account for medium-dependent differentiation signalling, an Fgf/Erk module was included, which inhibited Nanog's auto-activation and was activated by Oct4. An additional phenomenological external differentiation signal, inhibiting the core pluripotency network and inhibited by Nanog, was also present.

With the aim of refining this model, and encompassing genes crucial for pluripotency maintenance and transcriptional heterogeneity, we used the workflow summarised in Supplementary Fig. S1. First, we re-analysed a published RNA-sequencing data set of mESCs cultured in serum/LIF and 2i/LIF,[13] performing Gene Ontology (GO) analysis of the genes differentially expressed between the two media (3434 in total, fold change >2 and p-value <0.2) using the platform DAVID[32] (Supplementary Table S1). We categorised genes by their associated biological processes (BPs), and retained only those whose BPs related to development and differentiation, proliferation, cell cycle, morphology or cell death ('GO filter', Supplementary Fig. S1). Genes differentially expressed in the two media and involved in the relevant BPs (2288 in total) were then filtered to include only those that interact with the core pluripotency factors (i.e., Oct4, Sox2 and Nanog). This was done using an updated compendium for mESC transcription factors, CODEX,[33, 34] based on chromatin immunoprecipitation (ChIP) for transcription factors and histone modifications coupled to deep sequencing (ChIP-Seq), DNase-Seq and RNA-Seq data sets (Supplementary Fig. S1); genes regulating at least one of the core pluripotency factors were considered (34 genes, Supplementary Table S1). Rex1 present in the original network was removed, since it is only an output of the system (i.e., a marker of pluripotency) and does not directly regulate the pluripotent genes (thus, is not required to model the system dynamics).

From this analysis, we identified Mycn, Rest (RE1-silencing transcription factor) and Cdh7 (chromodomain helicase DNA binding protein 7) within the genes enriched in serum/LIF vs 2i/LIF, and Prdm14 (PR domain zinc finger 14) within the genes enriched in 2i/LIF. Of note, Rest, Mycn and Prdm14 had already been reported to significantly affect mESC pluripotency and the reprogramming of somatic cells.[35–37] Chd7 was not included in the final network; although it targets gene enhancer elements and co-localises with ESC master regulation, it is not essential for self-renewal and pluripotency.[38]

To add gene interaction directions (activation or repression), we used the NIA mESC cell bank,[39] generated by measuring the global expression patterns of mESCs upon perturbation of 54 transcription factors (Supplementary Fig. S1). Interactions between the core network genes Nanog, Oct4 and Sox2 were kept the same as in,[31] where the authors considered the Oct4-Sox2 heterodimer complex,[40] rather than the two single genes, regulating Nanog. Modelling the former regulation only has already been proven to enable good reproduction of mESC dynamics in serum/LIF[30, 31, 41–44] and reduces the number of unknown parameters in the model. In the NIA bank analysis we found that Oct4 represses Nanog; consistently, in our network, the complex both activates Nanog,[40] and also indirectly represses Nanog via the Fgf/Erk module (see below). The NIA bank analysis showed Nanog to be an auto-inhibitor, rather than an activator as in ref. 31. This interaction was kept as an auto-activator, again due to its demonstrated ability to provide a good reproduction of mESC dynamics; a comparison with the auto-inhibition case is provided below.

Since there were no data in the NIA bank for Rest and Prdm14, only interactions reported in the literature were included for the interaction of these genes with the rest of the network (see below). Of note, interactions found using CODEX but not confirmed in the NIA bank were removed (i.e., regulation of Mycn by Oct4 and Rest). In contrast, interactions found in the NIA bank but not present in CODEX were included (i.e., interaction between Mycn and Prdm14) as, although possibly being indirect interactions, they can affect the system dynamics. Regarding Mycn and Rest, in the NIA bank we found Oct4 and Sox2 to activate and inhibit both, respectively. Taken together as a heterodimer, with one activating (Oct4) and the other inhibiting (Sox2), there would be no resulting effect. Therefore no interaction from Oct4-Sox2 was included on Mycn and Rest.

Medium-dependent signalling components

To account for medium-dependent Nanog dynamics, the original model[31] considered the interaction of the MEK inhibitor (PD) with the core pluripotency network. However, the effects of the Gks3 inhibitor (Chiron) present in the 2i/LIF medium, and the resulting β-catenin stabilisation, were ignored. As in ref. 31, we model PD as a signal that weakens the inhibitory effect of Fgf/Erk signalling on Nanog. Given experimental evidence that PD alone is unable to sustain mESC clonal propagation,[1, 12] and the key role of the Wnt/β-catenin pathway in pluripotency and reprogramming of somatic cells,[5, 45] we additionally explored whether and how Chiron affects mESC expression patterns. We used a recently published ChiP-seq data set[46] to identify the interactions of β-catenin with the core pluripotency network and the additional genes identified (see above). Using the list of genes differentially expressed between Chiron (Wnt pathway activator through inhibition of Gsk3) and XAV (tankyrase inhibitor antagonising Wnt signalling through stabilisation of the Axin2 inhibitor) reported in[46] (Supplementary Table S1, Supplementary Fig. S1), β-catenin was found to inhibit directly both Mycn and Rest. We also included, using evidence from the literature, non-transcriptional Wnt/β-catenin pathway interactions known to be fundamental for pluripotency homoeostasis: Wnt/β-catenin inhibition of Tcf3 and Tcf3 inhibition of Oct4-Sox2 regulation of Nanog.[12] Additional protein interactions were added from the literature for the positive feedback loop between Nanog and Rest,[37] and Prdm14 inhibition of Fgf/Erk signalling.[47]

In summary, the complete network represented in Fig. 1 is composed as follows:

- the heterodimer Oct4-Sox2 represses Prdm14 and activates both itself and Nanog, with the latter activation inhibited by Tcf3;
- Nanog activates itself, Rest, and Prdm14, and inhibits Mycn;
- Mycn inhibits itself and activates Prdm14;
- Rest activates Oct4-Sox2 and Nanog;
- Ffg4/Erk is activated by Oct4-Sox2 and represses Nanog auto-regulation; the latter regulation is inhibited by both PD and Prdm14;
- β-catenin is activated by Chiron (for simplicity, we merged Chiron inhibition of Gsk3, and Gsk3 inhibition of β-catenin as a single interaction), and inhibits Rest, Mycn and Tcf3.

Dynamics in serum/LIF: Nanog bistability and noise-induced fluctuations

The SDE model represents activation/inhibition regulations in the network using Hill kinetics to encompass saturation with zero-mean Gaussian noise, as in ref. 31 to represent molecular background noise. For further details about model derivation, parameter settings and simulations, and full equations of the system, please refer to the Material and Methods section and Supplementary Information.

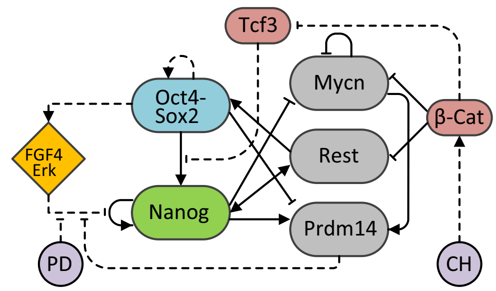


Fig. 1 Scheme of the regulatory network. The core network is composed of the Oct4-Sox2 heterodimer and Nanog; Mycn, Rest and Prdm14 were added as genes differentially expressed in serum/LIF vs. 2i/LIF, and involved in feedback loops with the core network (detailed derivation in Supplementary Fig. S1). Fgf4/Erk, β-catenin and Tcf3 are added to account for the inputs (i.e., PD and Chiron -CH in the Figure-, the two inhibitors present in 2i/LIF). *Solid* and *dashed lines* indicate transcriptional and non-transcriptional interactions, respectively

We performed 50 simulations, each for 1000 individual cells (as independent realisations of the SDE model), to estimate the model parameters that consistently describe flow cytometry [fluorescence-activated cell sorting (FACS)] data of Nd mESCs cultured in serum/LIF.[15, 17] The concentration of Nanog was recorded for each cell and distributions were captured at steady-state. Figure 2a shows a representative simulated distribution with a roughly equal number of cells in the NH and NL steady-states (the NH cells being 49.6 ± 1.5% of the overall population), correctly reproducing the bimodal distribution of Nanog observed experimentally in Nd cells (56.2 ± 8.0% of NH cells[15, 17]). Of note, the model can reproduce the fact that the Nd cells FACS distribution is narrower than the one measured in TNGA mESC experimental data,[48] due to the higher stability of the reporter protein in the latter cells. In single-cell simulations, stochastic fluctuations of Nanog are observed between the Low and High state, in contrast to monostable dynamics of Oct4-Sox2 (Fig. 2b). The time-scale of our simulations matches experimental single-cell measurements of live Nd mESCs obtained in time-lapse experiments, which show fluctuations between the NH and NL states within a single cell-cycle (circa 12 h) (Fig. S2 in ref. 15), thus on a much shorter time-scale than the TNGA cells.[48]

We performed numerical bifurcation analysis (i.e., a numerical study of the changes in the dynamics and stability of a system upon variations in its parameters[49, 50]) continuing the Nanog steady-state as a function of the maximal auto-regulative transcription rate s_4 (bifurcation parameter). A typical multi-stationary bifurcation plot is observed (Fig. 2c): within a certain s_4 region, two stable steady-states (*blue lines*, corresponding to NL and NH states) co-exist, separated from the unstable steady-state (red line) by two saddle-node bifurcations. This result is in line with the hypothesis that the Nanog multimodal distribution in serum/LIF arises from its auto-regulatory feedback loops.[41, 43, 31]

Finally, we sought to reproduce sorting experiments, in which NL and NH Nd mESCs were FACS sorted, separately plated in serum/LIF medium and analysed for Nanog reporter expression over a 4-day time-course. Our simulations correctly match the experimental observation that both subpopulations can restore the original distribution within 4 days of culture (compare Fig. 2d here and experimental results in Fig. 3b in ref. 17).

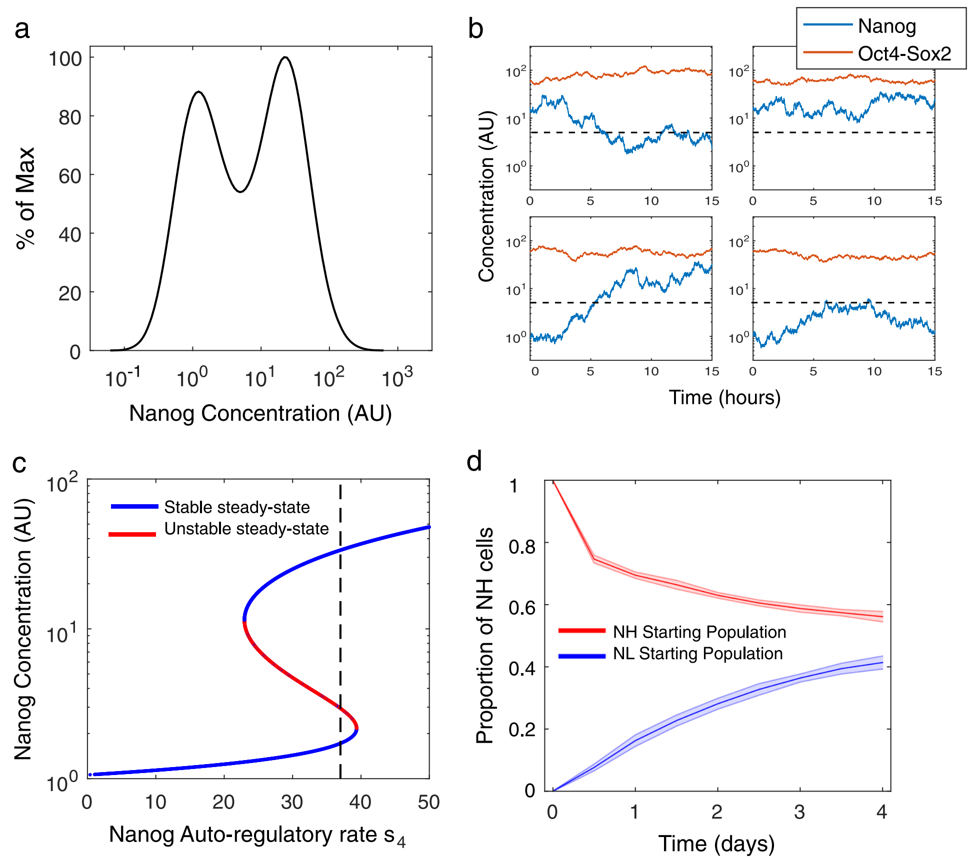


Fig. 2 System dynamics and stability in serum/LIF (PD = Chiron = 0 AU) **a** Steady-state distribution of Nanog simulated in 1000 cells showing that approximately 50% of the cells are in Nanog High (NH) and 50% Nanog Low (NL) states. **b** Four representative time-course simulations of Nanog (*blue line*) and Oct4-Sox2 (*red line*) in single cells over 15 h. Oct4-Sox2 dimer concentration remains fairly constant compared to Nanog. The *horizontal black dotted line* represents the threshold between NH and NL. **c** Continuation of Nanog steady-state for the deterministic system (no noise). *Blue lines* are stable steady-states, while the *red line* is unstable. The maximal transcription rate, s_4 (*dotted black line*), intersects both NH and NL steady-states, indicating bistability. **d** Sorting simulation: the dynamics of NH or NL cells (1000 for each state) are simulated over 4 days, and the proportion of cells in the NH state is recorded every 12 h ($n = 10$, *shaded areas* indicate standard deviation). The simulations show the cells tending towards the steady-state distribution of approximately 50% NH and 50% NL

Dynamics in 2i/LIF: persistence of noise-induced fluctuations

As described in the Introduction, Nd mESCs contain a persistent small subpopulation of NL cells even in ground-state cultures (2i/LIF, experimental FACS distributions in Fig. 4b in ref. 17 and Fig. 5a in ref. 15), in contrast to previous evidence of a complete elimination of the NL subpopulation using a stable GFP reporter for Nanog.[9, 13, 51] Existing mathematical formalisms, at least for parameters previously considered in the literature, predict a complete elimination of Nanog heterogeneity and temporal fluctuations in 2i/LIF cultures due to PD reducing Erk's inhibition of Nanog.[31, 41]

To reproduce the system dynamics in 2i/LIF, we reduced the standard deviation of Nanog transcriptional noise by 20% as compared to its value in serum/LIF, consistent with experimental measurements of the number of mRNA molecules per cell obtained by single-molecule fluorescent in situ hybridisation in Nd mESCs.[15] The model shows that 88.46 ± 1% of the cells are in the NH steady-state (Fig. 3a), matching experimental data (91.1 ± 3.1% in NH state[15, 17]).

Consistently, single-cell simulations over 15 h show fluctuations (Fig. 3b), in agreement with Nd single-cell time-lapses in ground-state cultures reported in ref. 15. Of note, for the specific levels of Chiron and PD used (Supplementary Information), the system is very close to the saddle-node bifurcation point (Fig. 3c). Figure 3d shows simulation results of the sorting experiment, with isolated NL and NH subpopulations approaching the original steady-state distribution after 4 days of culture.

To understand how the reduction of noise affects the system dynamics and stability, we sought to reproduce the same data keeping the Nanog noise term the same as in serum/LIF simulations (Supplementary Information). Obtaining a Nanog steady-state distribution and temporal dynamics equivalent to the previous reduced-noise scenario (giving now 87.27 ± 0.99% of NH cells, compare Figs. 3a, b with Supplementary Figure S2a and S2b, respectively) required an increase of the inputs PD and Chiron (Supplementary Information). With these parameters, the continuation of the Nanog steady-state against its auto-transcription rate s_4 shows that the deterministic system is monostable (see the *black dotted line* only intersecting the NH steady-state in Supplementary Fig. S2c). Although the deterministic system has only NH as a stable attractor, due to the Erk signalling block, the transcriptional noise of Nanog causes about 10% of cells to enter NL. Supplementary Figure S2d shows a sorting simulation, where populations starting from purely NL and NH populations were simulated over 4 days. The cells tend towards the steady-state distribution seen in Supplementary Fig. 2a. However, when compared to simulations in Fig. 3d, the rate at which NL cells tend towards the steady-state is much faster. Sorting experiments in 2i/LIF cultures have not been performed, but would inform about the suitability of one approach vs. the other.

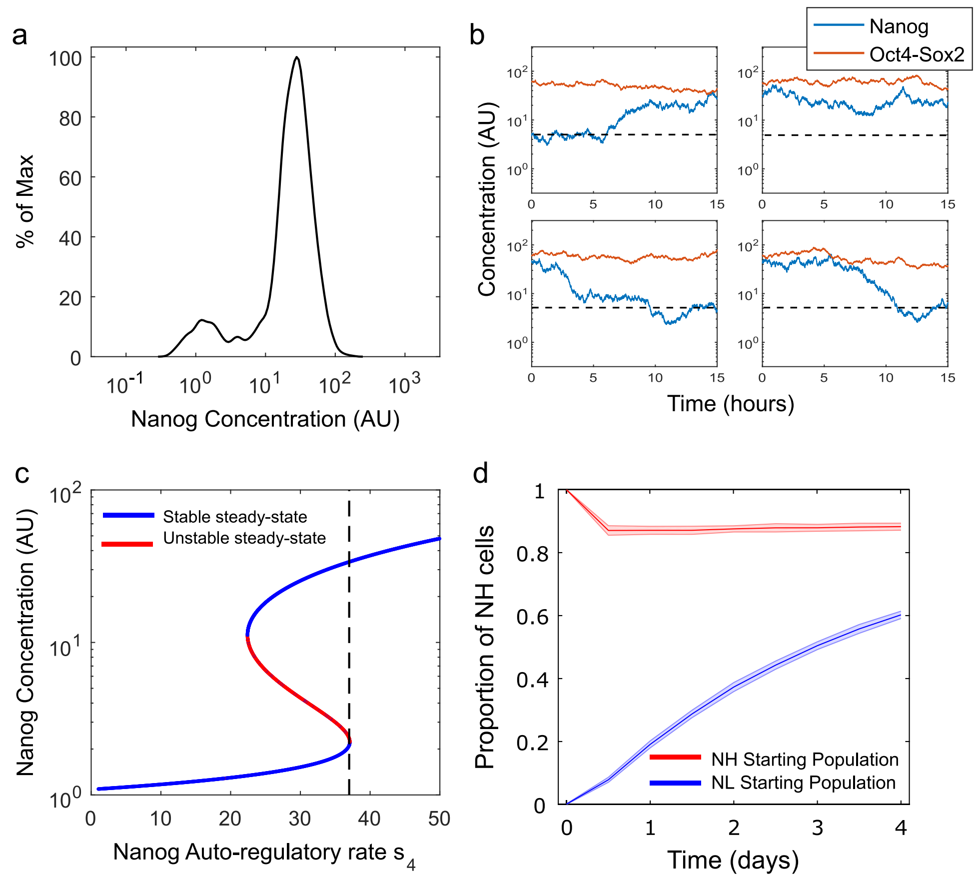


Fig. 3 System dynamics and stability in 2i/LIF (PD = Chiron = 2 AU, Nanog transcriptional noise reduced by 20%) **a** Steady-state distribution of Nanog simulated in 1000 cells, showing that approximately 90% and 10% of the cells are in the NH and NL states, respectively. **b** Four representative time-course simulations of Nanog (*blue line*) and Oct4-Sox2 (*red line*) in single cells over 15 h. Similarly to the serum/LIF case, Oct4-Sox2 dimer concentration remains fairly constant compared to Nanog. The horizontal *black dotted line* represents the threshold between NH and NL. **c** Continuation of Nanog steady-state for the deterministic system (no noise). *Blue lines* are stable steady-states, while the *red line* is unstable. The maximal transcription rate, s_4 (*dotted black line*), intersects both NH and NL steady-states, though is very close to the saddle-node bifurcation. **d** Sorting simulation: 1000 cells in the NH or NL state are simulated over 4 days, and the proportion of cells in the NH state is recorded every 12 h ($n = 10$, shaded areas indicate standard deviation). Cells restore the steady-state distribution of approximately 90% NH and 10% NL

Taken together, our results show that, accounting for reduced noise in Nanog expression in ground-state cultures, 2i/LIF shifts the system towards the monostable regime while keeping it excitable.

Dependence of the system stability on the GRN interactions

To understand how the shift of the system towards a monostable Nanog regime depends on the two inhibitors present in 2i/LIF and on the GRN topology, two-parameter continuations were performed on the variables PD and Chiron for the original network, and under single gene deletion (Fig. 4a). The area above each curve denotes the set of values of PD and Chiron for which the system is predicted to be monostable (NH state only).

The two-parameter continuation of the Nanog steady-state in the original network (no genes deleted) shows that the presence of both inhibitors is needed to leave the bistability region (Fig. 4a, *red line*). In particular, PD inhibits the auto-regulatory feedback of Nanog on itself, responsible for bistability, and Chiron strengthens Oct4-Sox2 activation of Nanog, facilitating the shift to the NH state.

The model predicts that Rest deletion (Fig. 4a, *black dotted line*) does not significantly impact system stability, while in the case of Mycn removal, monostability can be obtained only if PD and Chiron are increased (Fig. 4a, *blue line*); this is a consequence of the indirect activation of Nanog by Mycn through Prdm14 (Fig. 1). Indeed, the same trend is observed upon simulated Prdm14 deletion, but in this case a significant increase of inhibitors would be required (Fig. 4a, *green line*). Experimentally, Prdm14−/− mESCs cultured in 2i/LIF show heterogeneous Nanog expression, and defective differentiation potential.[35] This confirms our results about the key role of Prdm14 for the system dynamics and consequent pluripotency phenotype. Simulating our SDE model for the Prdm14 deletion scenario, we found, at steady-state, a much greater proportion of cells expressing NL in serum/LIF (Supplementary Fig. S3a) and persistence of Nanog heterogeneity in 2i/LIF (Supplementary Fig. S3b), matching the mentioned experimental data in Prdm14−/− mESCs.

To understand the effect of β-catenin and Tcf3 on the stability of the deterministic system, continuations of the Nanog steady-state against its auto-transcription rate were performed upon β-catenin deletion, and Tcf3 deletion. Continuation upon β-catenin deletion for serum/LIF parameters (Fig. 4b) shows that the system shifts slightly further into the bistable regime, as indicated by the *dashed black line*, though the change is subtle (compare with Fig. 2c). A similar result is seen for 2i/LIF parameters as seen in Fig. 4c; here the system is shifted further into the bistable area (compare with Fig. 3c). More pronounced results are seen for

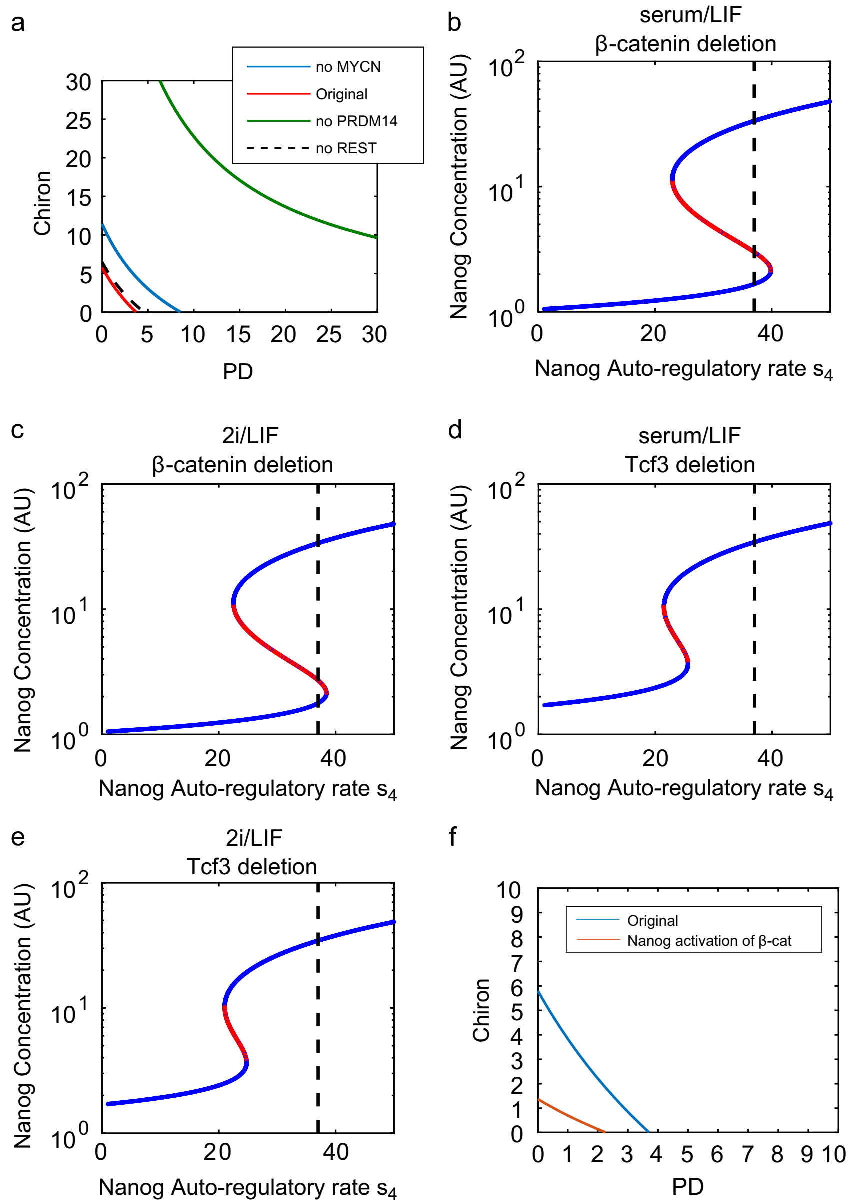


Fig. 4 Bifurcation analysis of the full system, and upon single network factor deletion **a** Two-parameter continuation (PD and Chiron). The area above each curve denotes the parameter (i.e., PD and Chiron) region in which the deterministic system exhibits monostability (NH only) for the full network (*red line*), and upon single deletion of Rest (*dashed line*), Mycn (*blue line*), and PRDM14 (*green line*). The graph shows that both PD and Chiron, with varying levels depending on gene deletion, are required to enter the monostable region. **b**–**e** Continuation of Nanog steady-state for the deterministic system, varying maximal transcription rate s_4, upon β-catenin deletion in serum/LIF (**b**) and 2i/LIF (**c**), and upon Tcf3 deletion in serum/LIF (**d**) and 2i/LIF (**e**). *Blue lines* are stable steady-states, while the *red lines* are unstable. **f** Two-parameter continuation (PD and Chiron). The area above the curve denotes the values of PD and Chiron in which the deterministic system exhibits monostability (NH only) for the full network, including an activation of Nanog on β-catenin (*red line*) comparing to the original full network with no additional activation of β-catenin (*blue line*)

Tcf3 deletion: using serum/LIF parameters (Fig. 4d), the system has moved entirely into the NH, monostable regime (compare with Fig. 2c). The same shift is seen, though even further, for 2i/LIF parameters as seen in Fig. 4e (compare again with Fig. 3c). These results are in line with experimental evidence that Tcf3 inhibition is associated with a significant delay in mESC differentiation.[52] Note that bifurcation analysis was performed on the system without noise. As shown in the previous section (and also in ref. 53), noise can be sufficient to cause bimodality and temporal fluctuations in a positive feedback loop, even in a monostable regime.

Nanog activation of β-catenin induces its stochastic fluctuations, but does not affect the overall system dynamics

Recently, we reported an indirect activation of Nanog on β-catenin (i.e., Nanog repression of β-catenin inhibitor Dkk1), which has important consequences for the successful reprogramming of mouse embryonic fibroblast to induced pluripotent stem cells.[5]

We analysed the dynamics of the system adding this extra interaction (for simplicity, we assumed that Nanog directly activates β-catenin; as the extra interaction with Dkk1 would eventually add some delay to the system). Simulating the system

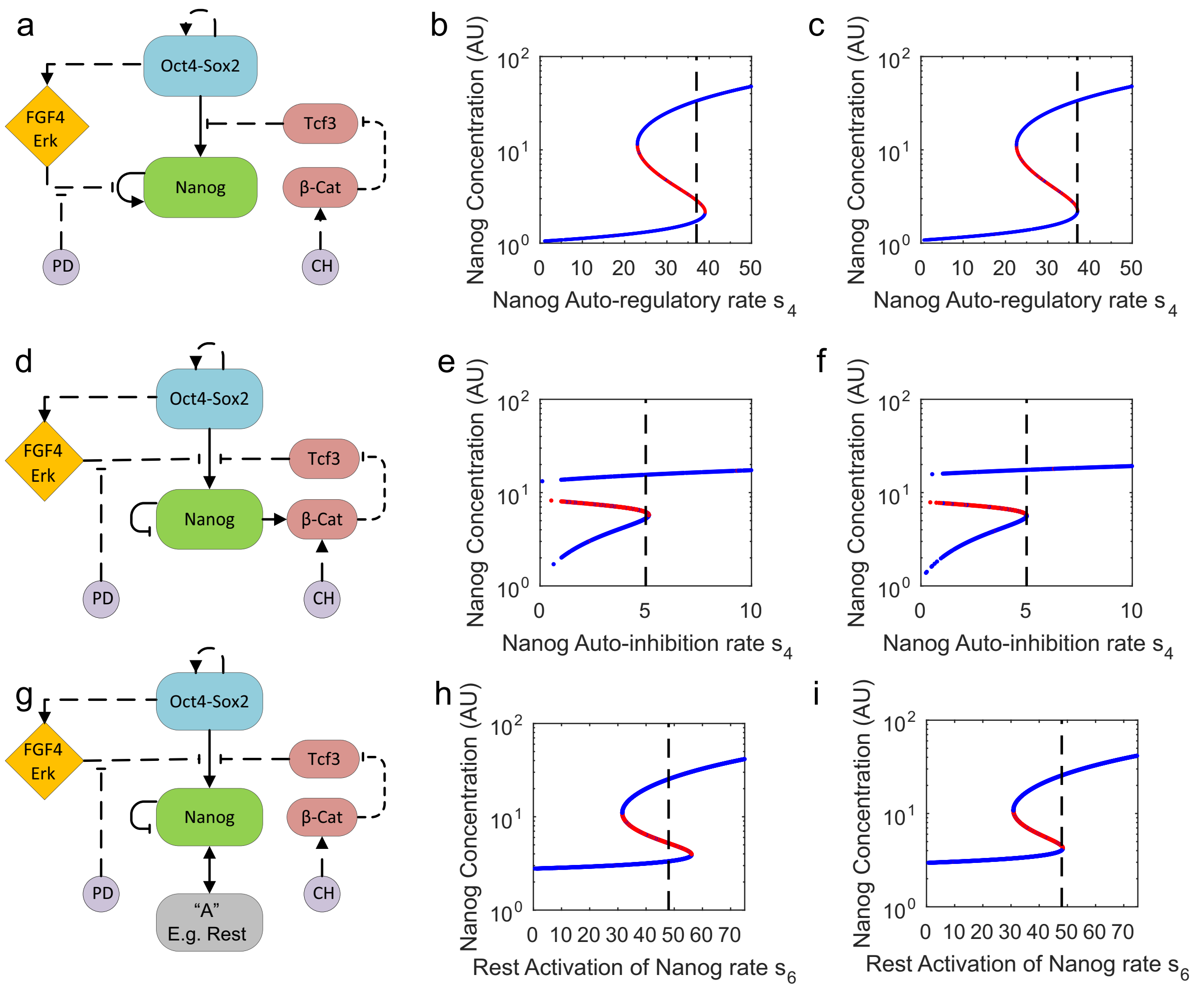


Fig. 5 Simpler network models demonstrating bistability of Nanog. **a** Scheme of simplified network, composed of core factors Nanog, Oct4, Sox2; Fgf4/Erk, β-catenin and Tcf3 are added to account for the inputs (i.e., PD and Chiron (CH), the two inhibitors present in 2i/LIF). *Solid* and *dashed lines* indicate transcriptional and non-transcriptional interactions, respectively. **b**, **c** Continuation of Nanog steady-state for the deterministic system (no noise) in serum/LIF (**b**) and 2i/LIF (**c**). *Blue lines* are stable steady states, while *red lines* are unstable. The maximal transcription rate, s_4 (*dotted black line*), intersects both NH and NL steady-states, indicating bistability in serum/LIF (**b**), but is close to the saddle-node bifurcation in 2i/LIF (**c**). **d** Scheme of Nanog auto-inhibitory simplified network, with additional activation of β-catenin by Nanog. **e**, **f** Continuation of Nanog steady-state for the deterministic system (no noise) in serum/LIF (**e**) and 2i/LIF (**f**). *Blue lines* are stable steady-states, while *red lines* are unstable. The maximal transcription rate of Nanog auto-inhibition, s_4 (*dotted black line*), intersects both NH and NL steady-states, indicating bistability in serum/LIF (**e**), while being closer to the saddle-node bifurcation in 2i/LIF (**f**). **g** Scheme of Nanog auto-inhibitory simplified network, with additional positive feedback supplied via an extra factor, i.e. Rest. **h**, **i** Continuation of Nanog steady-state for the deterministic system (no noise) in serum/LIF (**h**) and 2i/LIF (**i**). *Blue lines* are stable steady-states, while *red lines* are unstable. The maximal transcription rate of Rest activation on Nanog, s_6 (*dotted black line*), intersects both NH and NL steady-states, indicating bistability, in serum/LIF (**h**), and is very close to the saddle-node bifurcation in 2i/LIF (**i**)

to reproduce serum/LIF cultures (i.e., using the same parameter set as in Fig. 2), the new interaction introduces a bimodal distribution at steady-state also for β-catenin (compare Supplementary Fig. S4b with Supplementary Fig. S4d). Thus, we successfully fitted the distribution observed in the β-catenin tagged EL55 mESC cell line and the difference between the β-catenin Low and High states reported in ref. 5. Nanog distribution in serum/LIF is almost unchanged after adding the new interaction (compare Supplementary Fig. S4a, S4c). Simulating the system in 2i/LIF culture conditions (noise and inputs as in Fig. 3), the Nanog distribution is again unaffected (compare Supplementary Fig. S4e, S4g), but cells are pushed toward a β-catenin High state (compare Supplementary Fig. S4f, S4h), with a small population of β-catenin-Low cells, qualitatively reproducing data in ref. 5. Interestingly, Nanog activation of β-catenin changes the dependency of the system on PD and Chiron, reducing the amount of inducers required to push the system towards a monostable regime (Fig. 4f). These results suggest that Nanog regulation of β-catenin does not affect the overall system stability dynamics, though it can sustain the shift from a bistable to a monostable regime.

Reduced and modified network topologies can reproduce Nanog dynamics in the presence of at least one positive feedback

As described above, Nanog noise-induced temporal fluctuations arise in our model from the presence of a positive feedback loop in the network topology (i.e., Nanog auto-regulation, Fig. 1). We investigated whether reduced networks, which still keep such a loop, could recapitulate Nanog dynamics in both media.

First, we analysed the dynamics of a reduced network comprising the core factors Nanog, Oct4, Sox2 with an Fgf/Erk module only, the Wnt/β-catenin genes β-catenin and Tcf3 and the two inhibitors PD and Chiron (Fig. 5a). Interactions kept from the full network have the same parameter values; only interactions that were not directly comparable (i.e., those involving Fgf/Erk, PD and Prdm14) were

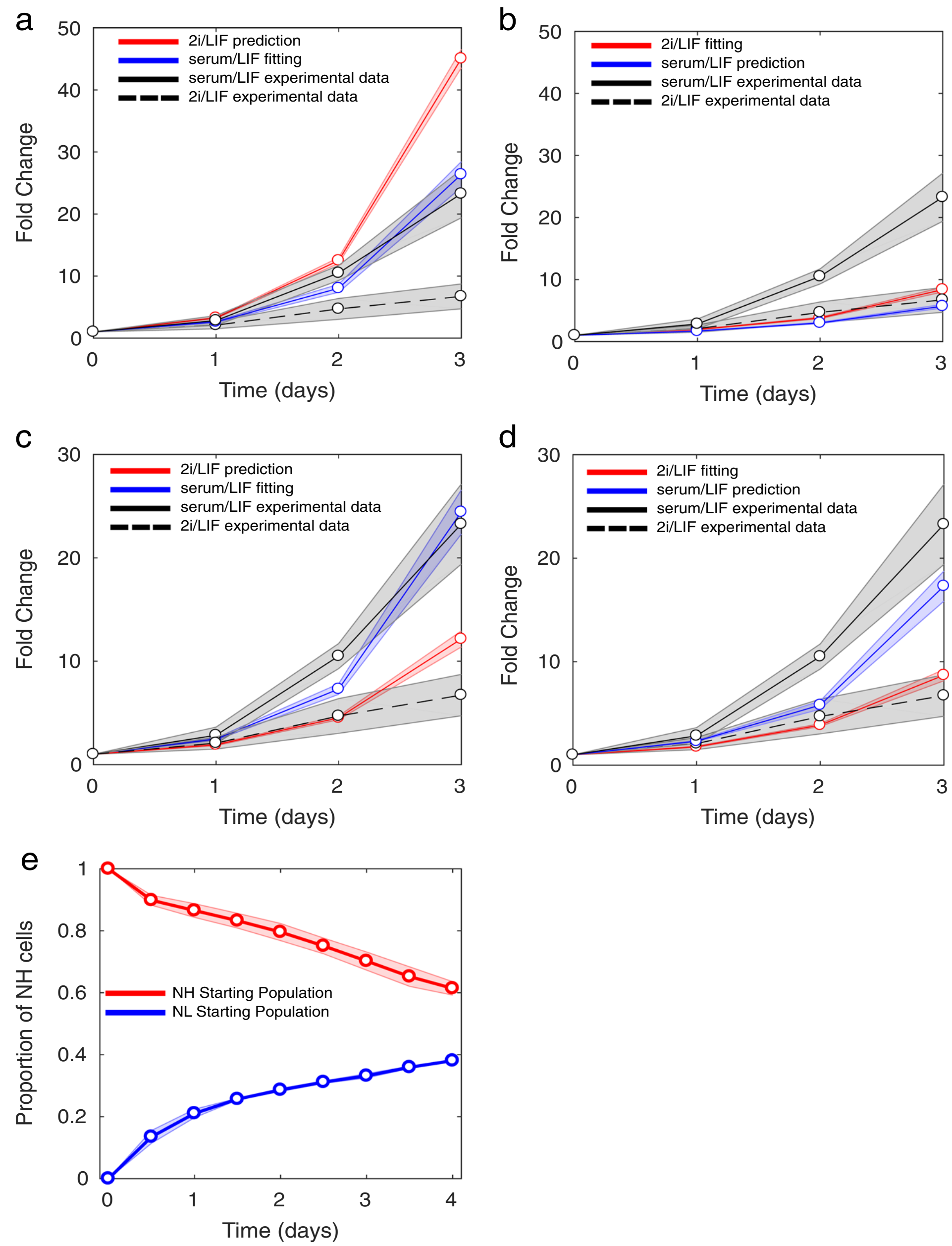


Fig. 6 Agent-based modelling results for Nanog-dependent or Mycn-dependent cell-cycle (**a**, **b**) Nanog-dependent cell-cycle model simulations. **a** The *blue line* depicts the Nanog-dependent cell-cycle model fitted to serum/LIF proliferation data (data shown as *solid black line*, RMSE 2.35), while the *red line* is the predicted cell-growth in 2i/LIF (RMSE 22.59 as compared to 2i/LIF data, shown as *dashed black line*). Each cell cycle was independently sampled from a normal distribution with mean 8.5 h for NH and 13.5 h for NL, and standard deviation 1 h for both NH and NL. **b** The *red line* shows fitting of the Nanog-dependent cell-cycle model to 2i/LIF data (RMSE 1.07), with prediction of growth in serum/LIF plotted in *blue* (RMSE 11.01); cell cycles were allocated with mean 13.175 h for NH and 20.925 h for NL. **c**, **d** Mycn-dependent cell-cycle model simulations. **c** The *blue line* depicts the Mycn-dependent cell-cycle model fitted to serum/LIF data (*solid black line*, RMSE 2.02), while the *red line* is the predicted cell growth in 2i/LIF (data shown as *dashed black line*, RMSE 3.08). In this case the cell cycles were allocated from a normal distribution with mean 8.5 h for Mycn High (MH) and 15 h for Mycn Low (ML). **d** The *red line* shows the Mycn-dependent cell-cycle model fitted to 2i/LIF data (RMSE 1.33) with serum/LIF predicted growth dynamics plotted in blue (RMSE 4.41). The cell cycles were allocated with means 9.5 h for MH and 16 h for ML. (**e**) Agent-based simulations of Nanog sorting experiment (Mycn-dependent cell-cycle model) in serum/LIF. The *red* and *blue lines* depict the changing ratio of NH cells in the population starting from NH or NL state, respectively. In **a**–**d**, data are the average of four independent experiments. Cell count, performed every 24 h for 3 days, is plotted as fold change vs. count at day 0; Standard Error (SE) is shown as a *shaded grey region*. In **a**–**e**, all model simulations are the average of 15 independent simulations, with the *shaded regions* showing SE

altered to match experimental data (Supplementary Information). Results of continuation of the Nanog steady-state against its auto-regulatory rate are comparable to those obtained with the full network under the same parameters for inputs and noise (compare Fig. 5b, c with Figs. 2c and 3c for serum/LIF and 2i/LIF, respectively). These results confirm that Nanog auto-activation is sufficient to explain bistability of the system.

As in our NIA bank analysis, two independent studies have recently reported that Nanog might function as an auto-inhibitor of itself rather than an activator.[54, 55] We investigated if we could still predict bistability considering auto-inhibition instead of auto-activation. Starting from the reduced model in Fig. 5a, we further added Nanog activation of β-catenin, which results in an indirect positive feedback loop acting on Nanog through β-catenin and Tcf3 (Fig. 5d). Changing the original parameters in such a way that this indirect positive feedback loop is strong enough (i.e., by increasing the system's non-linearity via Hill coefficient values and sensitivity to β-catenin via Michaelis–Menten constants) we obtained a bistable regime in serum/LIF and a shift of the system towards the monostable regime in 2i/LIF (Fig. 5e, f). Similarly, we reproduced the system dynamics neglecting Nanog activation of β-catenin, but instead including another positive feedback loop present in the full network (i.e., formed by the reciprocal interaction between Nanog and Rest, Fig. 5g; see continuations of Nanog steady-state against Rest activation rate in serum/LIF and 2i/LIF in Fig. 5h, i, respectively). The interaction with Rest was slightly altered in the reduced network, compared to the full network, to increase its nonlinearity (i.e., increasing Hill coefficient values, Supplementary Information). Taken together, these results indicate that if Nanog auto-activation is absent other positive feedback loops in the full network can still recapitulate bistability.

Another interaction present in the full network, which is debated in the literature, is β-catenin's inhibition of Mycn (predicted in the data set we used,[46] although other reports suggest the existence of a positive feedback between Mycn and β-catenin[56]). We performed simulations for two different scenarios: no interaction between Mycn and β-catenin, and β-catenin activation of Mycn (Supplementary Figs. S5 and S6, respectively). Comparing the dynamics using these topologies (Supplementary Figs. S5b–S5e and S6b–S6e) with those of the original network (Figs. 2c, 2a, 3c, 3a and 4a) we found that β-catenin's inhibition of Mycn is not necessary to reproduce Nanog dynamics in the two media.

Agent-based analysis of the system dynamics

We developed an agent-based version of the model, accounting for cell division (shown schematically in Supplementary Fig. S7a) in order to reproduce not only Nanog dynamics, but also the progression of mESCs through the cell cycle in the two media. The few reports that compared experimentally the progression of the G1 phase in metastable vs. ground-state mESC cultures disagree in their conclusions. One recent study showed that the G1-phase of mESCs grown in 2i/LIF is shorter than in serum/LIF.[57] This is in contrast to measurements of a slower cell cycle in 2i/LIF due to a longer G1 phase reported in ref. 58, and comparable proliferation rates in the two media reported in an independent study.[59] We therefore performed 3-day proliferation assays of wild-type mESCs cultured in serum/LIF or 2i/LIF (Materials and Methods); as shown in Fig. 6a–d, ground-state cultured mESCs show a significantly slower proliferation rate, as compared to serum/LIF cultured cells.

In order to reproduce our experimental data in agent-based simulations, we used the serum/LIF data for fitting and the 2i/LIF data for validation, and vice versa (Supplementary Information). We first coupled cell-cycle (CC) progression to the expression of Nanog, given recent experimental and modelling results in ref. 30, which suggest that at least in serum/LIF, Nanog expression levels correlate positively with mESC proliferation rate. By allocating different CC times to NH and NL cells (mean $CC_{NH} = 8.5$ h and $CC_{NL} = 13.5$ h, and standard deviations of 1 h), we could match the serum/LIF experimental proliferation data, with a root-mean squared error (RMSE) between the data and model of 2.35 (Fig. 6a). Furthermore, simulations showed a distribution of Nanog of approximately 50% NH and 50% NL cells (Supplementary Fig. S7b), consistent with the SDE model (Fig. 2) and experimental data.[15,17] Using the parameters fitted to the serum/LIF data to predict cell growth in 2i/LIF resulted in an increase in the NH subpopulation, as observed in the SDE model (Supplementary Fig. S7c); however, there was a large discrepancy between the proliferation profile and the recorded experimental values, with a RMSE of 22.59 (approx. 10× the error in the serum fitting, Fig. 6a).

We next used the 2i/LIF data for fitting, and the serum/LIF data for validation: by increasing the Nanog-dependent CC duration in both NH and NL populations by 55% ($CC_{NH} = 13.175$ h and $CC_{NL} = 20.925$ h) we were able to reproduce 2i/LIF proliferation dynamics, with a RMSE of 1.07 (Fig. 6b). However, fitted parameters performed poorly at predicting serum/LIF dynamics (RMSE 11.01, approx. 4× the error in our serum fitting, Fig. 6b). These results indicate that the previously suggested Nanog-dependent only model for CC progression[30] is unable to reproduce both the extensively reported increase in NH cells[1,5,9,12,60] and a reduction in proliferation in ground state cultures we measured here.

We therefore sought another possible candidate gene in the underlying GRN (Fig. 1) whose expression might be responsible for the different proliferation dynamics observed in the two media. The Myc family of transcription factors has been reported to positively regulate cell growth and proliferation in various cell types,[61] and is a master regulator of mESC proliferation.[62]

Therefore, we adapted the model by allocating the CC duration based on whether the mother of a newly divided cell is in the Mycn High (MH) or Mycn Low (ML) state. Of note, in our GRN, Mycn is inhibited by Nanog, and, therefore, also exhibits a bimodal distribution in serum/LIF, with reduced levels in 2i/LIF, consistent with experimental observations.[13] In a similar way to the Nanog-dependent CC model, we first fitted the serum/LIF data set, obtaining a RMSE of 2.02 (Fig. 6c, 14% improvement on the Nanog model) with $CC_{MH} = 8.5$ h and $CC_{ML} = 15$ h and a distribution of Nanog of approximately 50% NH and 50% NL cells (Supplementary Fig. S7d). Importantly, using these values to predict the 2i/LIF data set gave an 86% improvement on the Nanog model, with a RMSE of 3.08 (Fig. 6c) and Nanog distribution comparable to the SDE model (Supplementary Fig. S7e). When fitting 2i/LIF proliferation data, the Mycn model not only gave good fitting results ($CC_{MH} = 9.5$ h and $CC_{ML} = 16$ h, RMSE 1.33, Fig. 6d), but also performed acceptably in predicting dynamics in serum/LIF (RMSE 4.41, Fig. 6d). When fitting the two data sets independently, we found a difference in CC duration of 8%, as opposed to the 55% needed to adapt the Nanog model to the 2i/LIF data set. Finally, we further validated the agent-based model (using parameters as in Fig. 6c) performing NL and NH sorting simulations in serum/LIF (Fig. 6e): coupling Mycn to the cell cycle also reproduces experimental data,[17] with restoration of the original distribution within 4 days, as seen in our SDE model (Fig. 2d).

While the fitting results using a Mycn-dependent CC model were a significant improvement on the Nanog dependent model, the predictions were still not perfect (Figs. 6c and d); these discrepancies might be due to the simplified CC model we are using (in which daughter cells inherit the protein concentrations from the mother cell, and maintain a fixed cell cycle time until next division, Supplementary Information and ref. 30). Also, it may be noted that we are focusing on a limited number of genes that are differentially expressed between the two cultures conditions; other cell cycle regulators, not included here as not relevant to determine Nanog dynamics, might also affect the different proliferation kinetics that we measured.

DISCUSSION

Understanding the causes and consequences of heterogeneity in pluripotent stem cells is crucial to define optimal protocols for pluripotency maintenance and robust differentiation.

In this paper we focused on pluripotency-related factor dynamics of an isogenic mESC population, including feedback regulations, external stimuli provided in culture media, and coupling to the CC.

Our computational results show that heterogeneous expression and fluctuations of pluripotency genes in serum/LIF cultures arise as noise-induced bistable dynamics from the combination of positive feedback loops and transcriptional noise, in line with previous computational studies.[31, 41, 43, 60] Recent experimental results[63] based on mESC single-cell resolution data (single-molecule RNA-FISH and quantitative time-lapse imaging) are incompatible with alternative models, which suggest oscillatory[43] or excitable[48] dynamics in serum/LIF. In our model, bistability is tightly regulated not only by Fgf, known to destabilise the pluripotent state,[64] but also by the canonical Wnt pathway, demonstrated experimentally to avoid differentiation of mESCs into epiblast stem cells.[65] Indeed, our numerical continuation shows that the contribution of both PD and Chiron regulates the stability and dynamics of the system, and that deletion of Tcf3 causes a shift into the Nanog High state.

Our model suggests that the ground-state of pluripotency is excitable, i.e., a subpopulation of cells can still leave the Nanog High state. Our results re-conciliate experimental evidence of the overall homogeneity of pluripotency factors[9, 13, 51, 63] and the co-existence of subpopulations with different fate potential in 2i/LIF,[14, 66] and correctly predict the effect of gene knock-outs on the overall stability of the pluripotent state. It can be argued that other parameter settings might have been able to reproduce the experimental data considered here, given that we did not measure them directly. Hill kinetics are used in all the models we developed, with consequent assumptions on the system (e.g., the cell is well stirred, and concentrations and rate constants are not spatially dependent[67]). However, our qualitative analysis still provides useful insights into the system stability and dynamic behaviour. We showed that reduced or modified networks can reproduce Nanog dynamics, but each of the new genes introduced in our GRN has a key role in determining the system behaviour: Prdm14 interaction with Fgf/Erk maintains the right proportion of NH to NL cells; Rest, through its interaction with Nanog, ensures the presence of a positive feedback loop (even if Nanog auto-repression is considered), and Mycn is needed for CC regulation. Indeed, it was not possible to reproduce the measured mESC growth kinetics in both media by accounting only for the coupling between Nanog and cell proliferation previously proposed.[30] Our model and experimental data indicate, in agreement with recent evidence,[58] that 2i/LIF-cultured mESCs show reduced proliferation as compared to serum/LIF conditions. In our view, the causal relationship between CC progression and ground-state pluripotency needs to be better characterised, as well as the extent to which G1 phase lengthening affects pluripotency marker homogeneity. Extensions of the agent-based models proposed here, informed by single-cell data and including additional genes differentially expressed between serum/LIF and 2i/LIF and involved in CC regulation, might be highly informative.

Our results indicate that the perturbation of pluripotency and CC regulators must be properly balanced for pluripotency maintenance, and could inspire further research for the definition of chemical media affecting both signalling pathways and cell proliferation.

Lineage-associated gene expression is still present in 2i/LIF conditions,[14] reinforcing the biological hypothesis that there might still be functional heterogeneity related to extra-embryonic potential[1]; we also showed that multiple feedback loops can alter the dynamic equilibrium of ground-state pluripotency. Further investigation is needed to definitively assess if heterogeneous and fluctuating expression profiles might be detrimental for in vitro state pluripotency maintenance,[68] or might contribute to unbiased cell fate determination in the exit from pluripotency.[7, 63]

MATERIALS AND METHODS

Model derivation and parameterisation

We used in the SDE model Hill kinetics for transcriptional interactions to account for saturation; when a gene was regulated transcriptionally by more than one factor, Hill terms were summed. Regarding non-transcriptional interactions, Chiron was assumed to decrease the degradation of β-catenin, and β-catenin assumed to increase the degradation of Tcf3. Such regulations were modelled using Hill terms within the degradation terms of the respective equations. Tcf3 was assumed to decrease the binding rate of the Oct4-Sox2 heterodimer activation of Nanog. We used an mRNA steady-state assumption, modelling mRNA and protein dynamics in a single term, as in ref. 31. Noise was modelled with an additive zero-mean Gaussian term, where noise term parameters refer to the standard deviation of the normal distribution being randomly sampled from. Time was multiplied by a normalisation factor to match the time-scale of time-lapses in ref. 15. Parameters of the interactions kept from the reference GRN[31] were maintained as close as possible to the values in the model proposed therein, and adapted to match, at steady-state, the fold change direction of genes in serum/LIF vs. 2i/LIF reported in ref. 13, and Nanog steady-state distributions in ref. 15. Degradation rates were fixed, and Hill coefficients for interactions not present in the original model[31] were fixed at 1.

SDE simulations settings (including equations, parameter description and values), continuation analysis settings and details of the agent-based simulations are given in Supplementary Information.

Cell proliferation experiments

Mouse ESCs (E4Tg2a and GS) were grown in Dulbecco's modified Eagle's medium supplemented with 20% foetal bovine serum (Sigma), 1× nonessential amino acids, 1× GlutaMax, 1 × 2-mercaptoethanol and 1000U/ml murine recombinant LIF (Peprotech) or in NDiff227 Neural differentiation medium (Stem Cells, SCS-SF-NB-02) supplemented with 1000U/ml murine recombinant LIF (Peprotech), 3 µM CHIRON-99021 (Selleck) and 1 µM PD0325901(Selleck).

Mouse ESCs, grown in Serum/LIF or 2i/LIF for four passages, were seeded on gelatin-coated 24 wells/plates at the concentration of $10–20 \times 1\,0^3$ per well 24h before starting the experiment.

Every day, for 3 consecutive days, cells were tripsinized and 10µl of cell suspension were manually counted using a Neubauer chamber. Data are means ± Standard Error ($n = 4$, 2 replicates for each cell line).

Data availability

Files needed to reproduce simulations in the paper are available in the Figshare repository (10.6084/m9.figshare.5053330).

ACKNOWLEDGEMENTS

We thank N. Masuda (University of Bristol) for suggestions on the manuscript. This work was funded by MRC (MR/N021444/1) (L.M.), BrisSynBio, a BBSRC/EPSRC Synthetic Biology Research Centre (BB/L01386X/1) (L.M.) and EPSRC PhD scholarships (S.G., D.W.).

AUTHOR CONTRIBUTIONS

S.G., D.W., E.P., A.F., M.H., and L.M. conceived and designed the analysis; S.G. performed network derivation, simulations and continuation analysis; D.W. performed agent-based model simulations; E.P. performed experiments; A.F. supervised agent-based model development; L.M, S.G., D.W. and E.P. wrote the manuscript; S.G., D.W., E.P., A.F., M.H., L.M. commented on the manuscript; L.M. and M. H. supervised the project.

ADDITIONAL INFORMATION

Supplementary Information accompanies the paper on the *npj Systems Biology and Applications* website (doi:10.1038/s41540-017-0020-5).

Competing Interests: The authors declare that they have no competing financial interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

1. Hackett, J. A. & Surani, M. A. Regulatory principles of pluripotency: from the ground state up. *Cell Stem Cell* **15**, 416–430 (2014).
2. Cahan, P. & Daley, G. Q. Origins and implications of pluripotent stem cell variability and heterogeneity. *Nat. Rev. Mol. Cell Biol.* **14**, 357–368 (2013).
3. Chambers, I. *et al.* Nanog safeguards pluripotency and mediates germline development. *Nature* **450**, 1230–1234 (2007).
4. Hayashi, K., Lopes, S. M., Tang, F. & Surani, M. A. Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states. *Cell Stem Cell* **3**, 391–401 (2008).
5. Marucci, L. *et al*. beta-catenin fluctuates in mouse ESCs and is essential for Nanog-mediated reprogramming of somatic cells to pluripotency. *Cell Rep.* **8**, 1686–1696 (2014).
6. van den Berg, D. L. *et al.* Estrogen-related receptor beta interacts with Oct4 to positively regulate Nanog gene expression. *Mol. Cell Biol.* **28**, 5986–5995 (2008).
7. Torres-Padilla, M. E. & Chambers, I. Transcription factor heterogeneity in pluripotent stem cells: a stochastic advantage. *Development* **141**, 2173–2181 (2014).
8. Chambers, I. *et al*. Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* **113**, 643–655 (2003).
9. Ying, Q. L. *et al*. The ground state of embryonic stem cell self-renewal. *Nature* **453**, 519–523 (2008).
10. Davies, S. P., Reddy, H., Caivano, M. & Cohen, P. Specificity and mechanism of action of some commonly used protein kinase inhibitors. *Biochem. J.* **351**, 95–105 (2000).
11. Bain, J. *et al*. The selectivity of protein kinase inhibitors: a further update. *Biochem. J.* **408**, 297–315 (2007).
12. Wray, J. *et al*. Inhibition of glycogen synthase kinase-3 alleviates Tcf3 repression of the pluripotency network and increases embryonic stem cell resistance to differentiation. *Nat. Cell Biol.* **13**, 838–845 (2011).
13. Marks, H. *et al*. The transcriptional and epigenomic foundations of ground state pluripotency. *Cell* **149**, 590–604 (2012).
14. Morgani, S. M. *et al*. Totipotent embryonic stem cells arise in ground-state culture conditions. *Cell Rep.* **3**, 1945–1957 (2013).
15. Abranches, E. *et al*. Stochastic NANOG fluctuations allow mouse embryonic stem cells to explore pluripotency. *Development* **141**, 2770–2779 (2014).
16. Faddah, D. A. *et al*. Single-cell analysis reveals that expression of nanog is biallelic and equally variable as that of other pluripotency factors in mouse ESCs. *Cell Stem Cell* **13**, 23–29 (2013).
17. Abranches, E., Bekman, E. & Henrique, D. Generation and characterization of a novel mouse embryonic stem cell line with a dynamic reporter of Nanog expression. *PLoS One* **8**, e5992 (2013).
18. Miyanari, Y. & Torres-Padilla, M. E. Control of ground-state pluripotency by allelic regulation of Nanog. *Nature* **483**, 470–473 (2012).
19. Guo, G. *et al*. Serum-based culture conditions provoke gene expression variability in mouse embryonic stem cells as revealed by single-cell analysis. *Cell Rep.* **14**, 956–965 (2016).
20. Kumar, R. M. *et al*. Deconstructing transcriptional heterogeneity in pluripotent stem cells. *Nature* **516**, 56–61 (2014).
21. Semrau, S. & van Oudenaarden, A. Studying lineage decision-making in vitro: emerging concepts and novel tools. *Annu. Rev. Cell Dev. Biol.* **31**, 317–345 (2015).
22. Herberg, M. & Roeder, I. Computational modelling of embryonic stem-cell fate control. *Development* **142**, 2250–2260 (2015).
23. Singh, A. M. Cell cycle-driven heterogeneity: on the road to demystifying the transitions between "poised" and "restricted" pluripotent cell states. *Stem Cells Int.* **2015**, 219514 (2015).
24. Boward, B., Wu, T. & Dalton, S. Control of cell fate through cell cycle and pluripotency networks. *Stem cells* **34**, 1427–1436 (2016).
25. Burdon, T., Smith, A. & Savatier, P. Signalling, cell cycle and pluripotency in embryonic stem cells. *Trends Cell Biol.* **12**, 432–438 (2002).
26. Orford, K. W. & Scadden, D. T. Deconstructing stem cell self-renewal: genetic insights into cell-cycle regulation. *Nat. Rev. Genet.* **9**, 115–128 (2008).
27. Wang, R. & Guo, Y. L. Transient inhibition of cell proliferation does not compromise self-renewal of mouse embryonic stem cells. *Exp. Cell Res.* **318**, 2094–2104 (2012).
28. Fiore, G. *et al*. In-silico analysis and implementation of a multicellular feedback control strategy in a synthetic bacterial consortium. *ACS Synth. Biol.* **6**, 507–517 (2016).
29. Gorochowski, T. E. Agent-based modelling in synthetic biology. *Essays Biochem.* **60**, 325–336 (2016).
30. Herberg, M. *et al*. Dissecting mechanisms of mouse embryonic stem cells heterogeneity through a model-based analysis of transcription factor dynamics. *J. R. Soc. Interface* **13**, 20160167 (2016).
31. Herberg, M., Kalkan, T., Glauche, I., Smith, A. & Roeder, I. A model-based analysis of culture-dependent phenotypes of mESCs. *PLoS One* **9**, e92496 (2014).
32. Huang, D. W. *et al*. DAVID bioinformatics resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. *Nucleic Acids Res.* **35**, 169–175 (2007).
33. Sanchez-Castillo, M. *et al*. CODEX: a next-generation sequencing experiment database for the haematopoietic and embryonic stem cell communities. *Nucleic Acids Res.* **43**, 1117–1123 (2015).
34. Martello, G. *et al*. Esrrb is a pivotal target of the Gsk3/Tcf3 axis regulating embryonic stem cell self-renewal. *Cell Stem Cell* **11**, 491–504 (2012).
35. Yamaji, M. *et al*. PRDM14 ensures naive pluripotency through dual regulation of signaling and epigenetic pathways in mouse embryonic stem cells. *Cell Stem Cell* **12**, 368–382 (2013).
36. Varlakhanova, N. V. *et al*. MYC maintains embryonic stem cell pluripotency and self-renewal. *Differentiation* **80**, 9–19 (2010).
37. Singh, S. K. *et al*. REST-miR-21-SOX2 axis maintains pluripotency in E14Tg2a.4 embryonic stem cells. *Stem Cell Res.* **15**, 305–311 (2015).
38. Schnetz, M. P. *et al*. CHD7 targets active gene enhancer elements to modulate ES cell-specific gene expression. *PLoS Genet.* **6**, e1001023 (2010).
39. Nishiyama, A. *et al*. Uncovering early response of gene regulatory networks in ESCs by systematic induction of transcription factors. *Cell Stem Cell* **5**, 420–433 (2009).
40. Rodda, D. J. *et al*. Transcriptional regulation of nanog by OCT4 and SOX2. *J. Biol. Chem.* **280**, 24731–24737 (2005).
41. Chickarmane, V., Olariu, V. & Peterson, C. Probing the role of stochasticity in a model of the embryonic stem cell: heterogeneous gene expression and reprogramming efficiency. *BMC Syst. Biol.* **6**, 98 (2012).
42. Chickarmane, V., Troein, C., Nuber, U. A., Sauro, H. M. & Peterson, C. Transcriptional dynamics of the embryonic stem cell switch. *PLoS Comput. Biol.* **2**, e123 (2006).
43. Glauche, I., Herberg, M. & Roeder, I. Nanog variability and pluripotency regulation of embryonic stem cells--insights from a mathematical model analysis. *PLoS One* **5**, e11238 (2010).
44. Lakatos, D., Travis, E. D., Pierson, K. E., Vivian, J. L. & Czirok, A. Autocrine FGF feedback can establish distinct states of Nanog expression in pluripotent stem cells: a computational analysis. *BMC Syst. Biol.* **8**, 112 (2014).
45. Kuhl, S. J. & Kuhl, M. On the role of Wnt/beta-catenin signaling in stem cells. *Biochim. Biophys. Acta* **1830**, 2297–2306 (2013).
46. Zhang, X., Peterson, K. A., Liu, X. S., McMahon, A. P. & Ohba, S. Gene regulatory networks mediating canonical Wnt signal-directed control of pluripotency and differentiation in embryo stem cells. *Stem Cells* **31**, 2667–2679 (2013).
47. Nakaki, F. & Saitou, M. PRDM14: a unique regulator for pluripotency and epigenetic reprogramming. *Trends Biochem. Sci.* **39**, 289–298 (2014).
48. Kalmar, T. *et al*. Regulated fluctuations in nanog expression mediate cell fate decisions in embryonic stem cells. *PLoS Biol.* **7**, e1000149 (2009).
49. Marucci, L. *et al*. How to turn a genetic circuit into a synthetic tunable oscillator, or a bistable switch. *PLoS One* **4**, e8083 (2009).
50. Ghosh, S., Matsuoka, Y., Asai, Y., Hsin, K. Y. & Kitano, H. Software for systems biology: from tools to integrated platforms. *Nat. Rev. Genet.* **12**, 821–832 (2011).
51. Nichols, J. & Smith, A. Naive and primed pluripotent states. *Cell Stem Cell* **4**, 487–492 (2009).
52. Yi, F. *et al*. Opposing effects of Tcf3 and Tcf1 control Wnt stimulation of embryonic stem cell self-renewal. *Nat. Cell Biol.* **13**, 762–770 (2011).
53. To, T. L. & Maheshri, N. Noise can induce bimodality in positive transcriptional feedback loops without bistability. *Science* **327**, 1142–1145 (2010).
54. Navarro, P. *et al*. OCT4/SOX2-independent Nanog autorepression modulates heterogeneous Nanog gene expression in mouse ES cells. *EMBO J.* **31**, 4547–4562 (2012).
55. Fidalgo, M. *et al*. Zfp281 mediates Nanog autorepression through recruitment of the NuRD complex and inhibits somatic cell reprogramming. *Proc. Natl. Acad. Sci U.S.A.* **109**, 16202–16207 (2012).
56. Fagnocchi, L. *et al*. A Myc-driven self-reinforcing regulatory network maintains mouse embryonic stem cell identity. *Nat. Commun.* **7**, 11903 (2016).
57. Coronado, D. *et al*. A short G1 phase is an intrinsic determinant of naive embryonic stem cell pluripotency. *Stem Cell Res.* **10**, 118–131 (2013).
58. Kolodziejczyk, A. A. *et al*. Single Cell RNA-sequencing of pluripotent states unlocks modular transcriptional variation. *Cell Stem Cell* **17**, 471–485 (2015).
59. Tamm, C., Pijuan Galito, S. & Anneren, C. A comparative study of protocols for mouse embryonic stem cell culturing. *PLoS One* **8**, e81156 (2013).
60. Luo, Y., Lim, C. L., Nichols, J., Martinez-Arias, A. & Wernisch, L. Cell signalling regulates dynamics of Nanog distribution in embryonic stem cell populations. *J. R. Soc. Interface* **10**, 20120525 (2012).
61. Dang, C. V. MYC, metabolism, cell growth, and tumorigenesis. *Cold Spring Harbor Perspect. Med.* **3**, a014217 (2013).

62. Scognamiglio, R. *et al*. Myc depletion induces a pluripotent dormant state mimicking diapause. *Cell* **164**, 668–680 (2016).
63. Singer, Z. S. *et al*. Dynamic heterogeneity and DNA methylation in embryonic stem cells. *Mol. Cell* **55**, 319–331 (2014).
64. Kunath, T. *et al*. FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. *Development* **134**, 2895–2902 (2007).
65. ten Berge, D. *et al*. Embryonic stem cells require Wnt proteins to prevent differentiation to epiblast stem cells. *Nat. Cell Biol.* **13**, 1070–1075 (2011).
66. Macfarlan, T. S. *et al*. Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature* **487**, 57–63 (2012).
67. Karlebach, G. & Shamir, R. Modelling and analysis of gene regulatory networks. *Nat. Rev. Mol. Cell Biol.* **9**, 770–780 (2008).
68. Smith, A. Nanog heterogeneity: tilting at windmills? *Cell Stem Cell* **13**, 6–7 (2013).