Endogenous fluctuations of OCT4 and SOX2 bias pluripotent cell fate decisions

Daniel Strebinger [1,2], Cédric Deluz [1,2], Elias T. Friman [1,2], Subashika Govindan [1], Andrea B. Alber [1], David M. Suter [1]

[1]Sponsored Stem Cells Research Chair (UPSUTER), The Institute of Bioengineering (IBI), School of Life Sciences, Swiss Federal Institute of Technology, 1015 Lausanne, Switzerland

[2]These authors contributed equally

Correspondence and requests for materials should be addressed to D.M.S.

(email: david.suter@epfl.ch)

Running title : Fluctuation of TFs impact cell fate

Abstract

SOX2 and OCT4 are pioneer transcription factors playing a key role in embryonic stem (ES) cell self-renewal and differentiation. However, how temporal fluctuations in their expression levels bias lineage commitment is unknown. Here we generated knock-in reporter fusion ES cell lines allowing to monitor endogenous SOX2 and OCT4 protein fluctuations in living cells and to determine their impact on mesendodermal and neuroectodermal commitment. We found that small differences in SOX2 and OCT4 levels impact cell fate commitment in G1 but not in S phase. Elevated SOX2 levels modestly increased neuroectodermal commitment and decreased mesendodermal commitment upon directed differentiation. In contrast, elevated OCT4 levels strongly biased ES cell towards both neuroectodermal and mesendodermal fates. Using ATAC-seq on ES cells gated for different endogenous SOX2 and OCT4 levels, we found that high OCT4 levels increased chromatin accessibility at differentiation-associated enhancers. This suggests that small endogenous fluctuations of pioneer transcription factors can bias cell fate decisions by concentration-dependent priming of differentiation-associated enhancers.

Keywords

Embryonic stem cells/Endogenous fluctuations/SOX2/OCT4/Differentiation

Introduction

Embryonic stem (ES) cells can be maintained in a self-renewing state in vitro or be driven to commit to specific fates when exposed to differentiation signals. However, ES cells often exhibit asynchrony and divergences in fate commitment when subjected to the same differentiation cues. This obscures the interpretation of how instructive signals impact cell fate decisions, and limits the generation of pure ES cell-derived cell populations for future regenerative medicine applications (Cohen & Melton, 2011). Heterogeneity in cell fate commitment points at the coexistence of different cellular states, but these remain largely uncharacterized at the molecular level. Intercellular variability in expression levels of cell fate regulators constitutes a potential source of variable cellular states. One well-studied example is the heterogeneity of NANOG expression in serum + LIF culture conditions, which reflects reversible transitions of ES cells between the naïve and primed states (Filipczyk *et al*, 2015). However, protein expression variability of cell fate regulators in ES cells maintained in a more homogeneous, naïve state is poorly explored.

The transcription factors SOX2 and OCT4 (also known as POU5F1) are expressed in ES cells and are strictly required to maintain their pluripotent state (Chew *et al*, 2005; Okumura-Nakanishi *et al*, 2005; Tapia *et al*, 2015; van den Berg *et al*, 2010; Zhao *et al*, 2004; Niwa *et al*, 2000; Masui *et al*, 2007). Recent studies reported that quantitative properties of SOX2 and OCT4 such as expression levels or DNA-binding properties are predictive of cell fate decisions at the four-cell stage (White *et al*, 2016; Goolam *et al*, 2016). Differences in SOX2 concentrations were also shown to change its enhancer binding profile (Mistri *et al*, 2018), however whether this translates into differences in cell fate commitment is unknown. SOX2 and OCT4 were also reported to play antagonistic roles in the differentiation of ES cells towards the neuroectodermal (NE) and mesendodermal (ME) fates (Zhao *et al*, 2004; Thomson *et al*, 2011). However, these conclusions were largely based on long-term overexpression/knockdown or indirect correlations from fixed cells. A recent study in

human ES cells has shown that OCT4 levels are predictive for the cell fate choice between pluripotent self-renewal and extra-embryonic mesoderm commitment (Wolff *et al*, 2018). Nevertheless, how endogenous expression levels of SOX2 and OCT4 fluctuate over time in naïve ES cells and whether these fluctuations bias germ layer cell fate commitment remains unknown.

Here we investigated how endogenous variability in SOX2 and OCT4 protein levels impact the ability of ES cells to differentiate towards the NE or ME fates. To do so, we generated knock-in mouse ES cell lines allowing to monitor endogenous fluctuations of SOX2 and OCT4 proteins in live cells. We found that both proteins fluctuate over a 2 to 3-fold range with timescales of approximately one cell cycle in naïve mouse ES cells. Endogenous expression levels of OCT4 and to a lesser extent of SOX2 at the onset of differentiation impact the ability of pluripotent cells to differentiate towards NE and ME. Using ATAC-seq on cells with different endogenous OCT4 levels, we show that OCT4 fluctuations are associated to changes of chromatin accessibility of enhancers involved in cell differentiation.

Results

Generation of a sox2-SNAP / oct4-halo knock-in ES cell line

We first aimed at generating a cell line allowing monitoring of SOX2 and OCT4 protein levels in single living cells. To do so, we used CRISPR-Cas9 genome editing (Yang *et al*, 2013; Cong *et al*, 2013) to knock-in fluorescent tags in fusion to the C-terminus of endogenous SOX2 and OCT4 proteins. We generated a cell line in which both alleles of SOX2 are fused to a C-terminal SNAP tag, and one allele of OCT4 is fused to a C-terminal HALO tag (Fig.1a, Supplementary Fig.1a-c). Since SNAP and HALO tags allow orthogonal labelling with fluorescent dyes, endogenous SOX2 and OCT4 levels can be measured independently in individual living cells (Fig.1b). These knock-ins were generated in a previously established reporter cell line for ME and NE

commitment (SBR cell line (Deluz *et al*, 2016)), thus allowing to monitor SOX2 and OCT4 protein levels in live cells and to track differentiation outcomes. The resulting cell line was named SBROS for **S**ox1/**B**rachyury/**R**eporter/**O**CT4-HALO/**S**OX2-SNAP. Because of the heterozygocity of the OCT4-HALO knock-in, we assessed whether OCT4-HALO levels are a good proxy for total OCT4 levels at the single cell level. To do so, we combined immunofluorescence against total OCT4 with direct labelling of OCT4-HALO using the Halo-TMR dye. We found that total OCT4 and OCT4-HALO levels were well correlated in single cells, suggesting that OCT4-HALO levels allow estimating total OCT4 levels (Fig.1c and Supplementary Fig.1d). We then verified whether this cell line expresses normal levels of pluripotency markers by quantitative immunofluorescence analysis. In contrast to Western blotting, immunofluorescence is not biased by differences in membrane transfer efficiency of proteins with different molecular weights (such as OCT4 and OCT4-HALO) and allows obtaining distributions of protein levels in the cell population. We found that the SBROS cell line expressed on average 89% of wild type mean OCT4 levels, 111% of wild type mean SOX2 levels, and 129% of wild type mean NANOG levels. The distributions and median expression levels of these proteins were similar to those of wild type E14 ES cells (Fig.1d). We also found the half-lives of both OCT4-HALO (7.8 ±1.3 h) and SOX2-SNAP (8.1 ±1.0 h) to be close to published half-life values for these proteins (Pan *et al*, 2016; Fang *et al*, 2014; Liu *et al*, 2017a) (Fig.1e), and the average cell cycle length of SBROS and wild type E14 ES cells were similar (Fig.1f). mRNA levels of pluripotency markers of SBROS cells were mostly unaltered (Supplementary Fig.1e). We also verified cloning efficiency and alkaline phosphatase activity after one week of clonal growth, and found it to be comparable between SBROS and its parental SBR cell line (Supplementary Fig.1f).

While the normal pluripotent phenotype of SBROS cells implies that the homozygous *Sox2-SNAP* alleles are functional, this is not necessarily the case for the heterozygous *Oct4-Halo* allele. To verify the functionality of the OCT4-HALO fusion

protein, we drove its expression from the constitutive EF1-α promoter in Zhbtc4 cells, which allow doxycycline (dox)-inducible OCT4 knockout. After 24 hours of dox induction, endogenous OCT4 expression is lost (Niwa *et al*, 2000) and thus cells rely only on the OCT4-HALO protein to maintain their pluripotent state. In these conditions, we found that the OCT4-HALO protein was able to fully rescue pluripotency, thus confirming its functionality (Supplementary Fig.1g). We also found the interaction of the OCT4-HALO protein with SOX2 to be preserved by co-immunoprecipitation experiments in the SBROS cell line (Supplementary Fig.1h), and the ChIP-seq profile of OCT4-HALO to be enriched at wt OCT4 peaks (Supplementary Fig.1i).

SOX2 fluctuations regulate OCT4 levels

It has been reported that SOX2 and OCT4 protein levels are positively correlated in single cells (Filipczyk *et al*, 2015), however the mechanism underlying this correlation is not understood. While ChIP-seq studies have shown that SOX2 and OCT4 bind to the regulatory sequences of the *Sox2* and *Oct4* genes (Chew *et al*, 2005; Loh *et al*, 2006), the functional impact of their binding on *Sox2* and *Oct4* expression is unclear. To determine how SOX2 and OCT4 impact their own and each other's expression levels, we established doxycycline (dox)-inducible ES cell lines allowing precisely timed expression of fusions of SOX2 or OCT4 to the YPet fluorescent protein. The functionality of these fusion proteins was validated by their ability to rescue pluripotency of inducible SOX2 or OCT4 knockout cells (2TS22C (Masui *et al*, 2007) or Zhbtc4 cells, respectively) (Supplementary Fig.1g and 1j), and by ChIP-seq analysis that showed genome-wide binding profiles similar to wt SOX2 and OCT4, respectively (Supplementary Fig.1i).

We then used these constructs to investigate the impact of YPet-OCT4 and YPet-SOX2 expression on endogenous SOX2 and OCT4 protein levels. To do so, YPet-OCT4 and YPet-SOX2 inducible ES cell lines were treated with dox for 0h, 4h or 7h, followed by immunofluorescence staining with anti-OCT4 and anti-SOX2 antibodies.

While upon YPet-OCT4 overexpression, SOX2 levels remained stable (Supplementary Fig.2a-c), SOX2 overexpression modestly increased OCT4 levels already four hours after SOX2 induction (Supplementary Fig.2d-f). Interestingly, SOX2 levels slightly decreased as a function of overexpressed YPet-SOX2 over time (Supplementary Fig.2g), raising the possibility that SOX2 negatively regulates its own expression, as suggested by previous studies (Kopp *et al*, 2008; Ormsbee Golden *et al*, 2013).

To confirm negative SOX2 autoregulation, we then monitored endogenous SOX2 levels in live cells upon exogenous SOX2 expression in another knock-in cell line in which both endogenous *Sox2* alleles are fused to *nanoluc*, allowing real-time monitoring of endogenous SOX2 levels by luminescence microscopy. Additionally, A P2A-Firefly Luciferase (FLUC) cassette was also knocked-in in fusion to one allele of *Sox1* to monitor NE commitment (Supplementary Fig.2h, Supplementary Fig.1a-c). We called this cell line **S**ox2-**N**anoluc-**S**ox1-**F**luc (SNSF). Quantitative immunofluorescence analysis confirmed that the NANOLUC tag did not strongly alter expression level distributions of pluripotency factors, as cells retained 86% of wildtype mean OCT4 levels, 79% of wildtype mean SOX2 levels, 93% of wildtype mean NANOG levels, and distributions and median expression levels of these proteins were similar to wild type E14 cells (Supplementary Fig.2i). Growth rates were also unaltered (Supplementary Fig.2j). Expression levels of pluripotency markers at the mRNA level (Supplementary Fig.1e), cloning efficiency and alkaline phosphatase activity after one week of clonal growth were also similar to wt ES cells (Supplementary Fig.1f).

We transduced the SNSF cell line with lentiviral constructs allowing dox-inducible overexpression of SOX2-SNAP, or YPet-SOX2-delDBD that lacks the SOX2 DNA binding domain. We then monitored SOX2 protein levels over time by luminescence microscopy after dox induction, as well as SOX2-SNAP or YPet-SOX2-delDBD levels by fluorescence microscopy. While YPet-SOX2-delDBD expression did not impact endogenous SOX2 levels, SOX2-SNAP overexpression reduced SOX2 protein levels

to 50% of their initial levels after approximately 9 hours (Fig.2a) ($p=1.88*10^{-14}$, Mann-Whitney-Wilcoxon test). This time scale is close to the half-life of the SOX2 protein (8.2 +/- 1h (SE) for the half-life of SOX2-SNAP, Fig.1e) and thus raises the possibility of a very rapid arrest of SOX2 protein accumulation. To determine whether SOX2 directly alters *Sox2* mRNA levels, we performed RT-QPCR shortly after overexpression of SOX2-SNAP or YPet-SOX2-delDBD. We found that *Sox2* mRNA levels were decreased already 2 hours after dox induction, and were down to 20% of their initial level after 6 hours (Fig.2b). This suggests that SOX2 rapidly represses the expression of its own mRNA, in line with earlier studies (Kopp *et al*, 2008; Ormsbee Golden *et al*, 2013). We next aimed to determine how SOX2 overexpression increased OCT4 levels. Surprisingly, *Oct4* mRNA levels were unaffected after 6 hours of SOX2 overexpression (Fig.2b). Since SOX2 and OCT4 form heterodimers, we reasoned that SOX2 overexpression could increase OCT4 levels by increasing the stability of the OCT4 protein. We thus measured OCT4 half-life by pulse-labelling OCT4-HALO with the Halo-Sir 647 dye in the SBROS cell line (as described previously in (Alber *et al*, 2018)) after overnight dox induction of SOX2-SNAP expression. We found the half-life of OCT4 to be increased by 50 % (Fig.2c), suggesting that SOX2 increases OCT4 levels by decreasing its degradation rate. In contrast, SOX2 overexpression does not alter its own half-life, as we have shown previously (Alber *et al*, 2018).

We next aimed to determine how endogenous variations in SOX2 and OCT4 levels affect each other's expression levels. To do so, we labelled SBROS cells with SNAP-SiR647 and Halo-TMR, and sorted cells for either high or low SOX2-SNAP levels, but with the same, intermediate OCT4-HALO expression levels. In order to minimize the effects of cell cycle progression on differences in SOX2 and OCT4 levels, we sorted cells that were in G1 phase based on DNA content (Fig.2d, Supplementary Fig.2k). The converse experiment was performed to determine how endogenous OCT4 levels impact SOX2 expression (Fig.2d, Supplementary Fig.2k). After sorting, cells were kept

either in pluripotency maintenance conditions (N2B27+2iLIF) or in differentiation conditions (N2B27). In cells sorted for SOX2-high or SOX2-low levels, OCT4 levels were increased and decreased 8 hours after sorting, respectively (Fig.2e, Supplementary Fig.2l). In contrast, in cells sorted for OCT4-high or OCT4-low levels, SOX2 levels remained unchanged 8 hours after sorting (Fig.2f, Supplementary Fig.2l). These results suggest that SOX2-high and low cells tend to increase and decrease OCT4 expression levels over time, respectively, and that low amplitude, endogenous variations in SOX2 levels regulate dynamic changes in OCT4 levels.

Characterisation of SOX2 and OCT4 fluctuations

The intercellular variability and the reversion of SOX2 and OCT4 levels towards their mean levels 8 hours after sorting for SOX2-high/low or OCT4-high/low levels (Fig.2e and Fig.2f, respectively) suggest that these proteins fluctuate over a time scale of hours in individual cells. We thus decided to use our knock-in ES cell lines to directly measure protein expression level fluctuations at the single cell level. We monitored absolute SOX2 levels in the SNSF cell line by luminescence microscopy, using a signal calibration approach we reported previously (Mandic *et al*, 2017). As expected, SOX2 levels doubled over one cell cycle (Fig.2g, Supplementary Fig.3a), and SOX2 concentrations calculated after normalization on the inferred nuclear volume (described in (Filipczyk *et al*, 2015)) were constant on average (Supplementary Fig.3b). In single cells, we found SOX2 concentrations to fluctuate over a 2 to 3-fold range (Fig.2g). To measure the time scale of SOX2 concentration fluctuations, individual cells were assigned a rank based on their initial SOX2 level (Fig.2h). We then used a rank-based autocorrelation function to determine the time scale of protein level fluctuations. This time scale is referred to as the mixing time (Sigal *et al*, 2006), and describes how long it takes for a cell to lose its expression rank and thereby to “mix” its expression level with other cells. Using data from either two or a single full cell cycle, we found SOX2 mixing times on the order of one cell cycle (Fig.2i and

Supplementary Fig.3c,d). Since < 2-fold changes in SOX2 expression were reported to compromise pluripotency maintenance (Kopp *et al*, 2008), our data suggest that rapid readjustment of SOX2 levels may be required to maintain pluripotency despite fluctuation amplitudes of 2-3 fold. We then performed analogous experiments using the SBROS cell line to monitor OCT4-HALO levels over the cell cycle by live fluorescence microscopy, revealing similar fluctuation amplitudes (Fig.2j, Supplementary Fig.3e,f), and mixing times (Fig.2k,l) as compared to SOX2. Thus, both SOX2 and OCT4 display 2 to 3-fold, rapid expression level fluctuations in the pluripotent state.

SOX2 and OCT4 fluctuations impact NE and ME commitment

We next aimed to determine how endogenous SOX2 levels tune the probability of NE differentiation by monitoring SOX2-NLUC levels and SOX1-P2A-FLUC expression after removal of 2i and LIF (Fig.3a,b). We tracked NLUC and FLUC signals in individual cells over time, and individual cell traces were aligned and normalized for cell cycle duration *in silico*, using linear resampling of the time variable (see Methods). We then grouped cells according to their expression of FLUC, and traced their SOX2 expression levels one cell cycle before, during the cell cycle of FLUC expression onset, and in the subsequent cell cycle. Traces of cells that did not express FLUC were temporally aligned to traces of cells that expressed FLUC, so that both groups of cells were compared at the same average time after which self-renewal signals were withdrawn (Fig.3c). Interestingly, during the cell cycle before Sox1+ cells appeared, higher SOX2 levels at the beginning and end, but not in the middle of the cell cycle correlated with a higher probability of NE differentiation (Fig.3c, marked by *). This suggests that SOX2 levels may play a role in NE commitment at the M-G1 transition. This finding is in line with an earlier study from our laboratory, in which we found that the absence of SOX2 at the M-G1 transition suppresses its ability to enhance neuroectodermal fate commitment (Deluz *et al*, 2016).

We next investigated whether different SOX2 levels at the very onset of differentiation impact NE and ME commitment. To do so, we sorted G1-gated SBROS cells stained with SNAP-SiR647 for low, medium and high SOX2 levels (Fig.3d and Supplementary Fig.4a). Cells were then released from self-renewal conditions by seeding in N2B27 medium devoid of 2i and LIF, and four days later NE and ME commitment were assessed by flow cytometry using the SOX1-P2A-eGFP and the BRA-P2A-mCherry reporters as readout. The fraction of eGFP+ cells scaled with initial SOX2 levels (Fig.3e), suggesting that high SOX2 levels at the time of release from self-renewal enhance NE fate commitment. In contrast, SOX2 levels had only a weak impact on ME commitment in these conditions (Fig.3e).

To determine the impact of OCT4 levels at the onset of pluripotency exit on NE and ME commitment, we sorted G1-gated SBROS cells stained with Halo-TMR in OCT4-low and OCT4-high subpopulations (Fig.3d and Supplementary Fig.4b) and cultured them for four days in the absence of LIF and 2i. Surprisingly, we found a large difference in NE and ME commitment between these populations (Fig.3f), even though average OCT4 levels differed by < 2-fold between them (Supplementary Fig.4b). We next aimed to explore the potential causal relationship between high OCT4 levels and increased NE/ME commitment. To do so, we generated a cell line allowing for inducible expression of SNAP-OCT4 in the SBR background, and treated it with dox for 12h to overexpress OCT4 for a brief period of time, thereby mimicking the timescales of endogenous OCT4 fluctuations (note that we previously demonstrated the functionality of the SNAP-OCT4 fusion protein (Deluz *et al*, 2016)). Subsequently, cells were sorted for G1 phase and differentiated for four days in N2B27 medium without 2i and LIF (Fig.3g and Supplementary Fig.4c). Strikingly, this led to a ~ 2-fold increase in the fraction of mCherry+ cells, suggesting that even a brief increase of OCT4 levels prior to self-renewal release enhances ME commitment. However, NE commitment was not enhanced as the fraction of eGFP+ cells remained unchanged (Fig.3g and Supplementary Fig.4c). We reasoned that this discrepancy

could be caused by opposing roles of high OCT4 levels before and after removal of pluripotency signals. To investigate the impact of OCT4 overexpression after release of pluripotency signals, we treated SNAP-OCT4 cells with dox throughout differentiation. While this also led to a similar increase in ME differentiation as obtained with a short dox treatment, NE commitment was strongly inhibited (Supplementary Fig.4d,e). In summary, our data suggests that after a short pulse of dox induction in the SNAP-OCT4 inducible SBR cell line, high OCT4 levels before and at the onset of differentiation both enhance ME differentiation but have an opposite impact on NE differentiation.

We next measured the combinatorial impact of SOX2 and OCT4 on differentiation by sorting G1-gated SBROS cells into four different subpopulations (Fig.3h and Supplementary Fig.4f) followed by four days of differentiation (Fig.3i and Supplementary Fig.4g). As expected, SOX2-high/OCT4-high (SHOH) cells were the most efficient to differentiate towards NE and ME, SOX2-low/OCT4-high (SLOH) cells were less capable to differentiate towards NE as compared to SHOH, and OCT4-low (SHOL and SLOL) populations were strongly impaired in differentiating towards both fates. To ask whether endogenous SOX2 and OCT4 level variability impacts NE and ME commitment during later cell cycle stages, we performed the same experiments in S phase-gated cells. We found the impact of SOX2 on NE commitment and the impact of OCT4 on both NE and ME commitment to be decreased (Fig.3j and Supplementary Fig.4h). Note that the differentiation efficiency in our conditions was not lower in S phase-sorted as compared to G1 phase sorted cells in contrast to previous studies (Pauklin & Vallier, 2013; Coronado *et al*, 2013). In fact, while ME differentiation of S phase-sorted cells was comparable to G1 phase-gated cells, NE differentiation was increased (Supplementary Fig.4i). These discrepancies likely result from differences in differentiation protocols and/or endpoint readouts. Regardless, our data suggest that differentiation to NE and ME depends on the levels of OCT4 and SOX2 in G1 but not

in S phase, which in the case of SOX2 is in line with its reported function at the Mitosis-G1 transition in NE induction (Deluz *et al*, 2016).

We next aimed at determining whether SOX2 and OCT4 levels also impact cell fate commitment upon directed differentiation. To do so, we cultured G1-sorted SHOH, SHOL, SLOH and SLOL subpopulations in N2B27 supplemented with 3μM CHIR99021 during three days to drive ME differentiation (Turner *et al*, 2014). Surprisingly, we found that OCT4-high and OCT4-low cells displayed a similar potential to differentiate towards ME (Fig.3k). In contrast, SOX2-low cells were almost two-fold more efficient than SOX2-high cells to differentiate towards ME (Fig.3k). This suggests that the sustained activation of the wnt pathway overcomes the influence of OCT4 in priming ME differentiation, and that SOX2 expression levels now become limiting for ME differentiation efficiency.

OCT4-high cells open differentiation enhancers

We then aimed to identify the molecular mechanisms by which small and transient endogenous fluctuations of SOX2 and OCT4 result in major biases in differentiation potential. As SOX2 and OCT4 were shown to regulate chromatin accessibility (King & Klose, 2017; Raccaud *et al*, 2019), we reasoned that small changes in their expression level could prime cells for different fates by altering the chromatin accessibility landscape. We thus performed ATAC-seq in G1-sorted SHOH, SHOL, SLOH and SLOL cells. We quantified the fraction of reads in 81’132 open regions identified by peak calling and found no overall differences between conditions (Supplementary Fig.5a). We compared the accessibility of each open region in high vs low conditions for SOX2 and OCT4 as well as for SHOH vs SLOL. 3’914 loci (4.8%) were significantly up- or downregulated (FDR < 10%) upon changes in OCT4 alone (538 loci), SOX2 alone (1’259 loci), or SOX2 and OCT4 together, i.e. SHOH vs SLOL (2’117 loci). We grouped these loci into OCT4-regulated, SOX2-regulated and co-regulated loci that were either upregulated (more open in high cells) or downregulated

(more open in low cells) (Fig.4a,b and Supplementary Fig.5b,c). Loci in all groups were close to differentiation-associated genes (Fig.4b) but those upregulated by OCT4 alone were the most enriched for genes involved in differentiation processes (Fig.4c, Supplementary Fig.5d-h). Using a subset of OCT4-upregulated loci that overlap OCT4 peaks and with a more stringent false discovery rate (< 5%) also revealed enrichment for differentiation terms (Supplementary Fig.5i), showing that OCT4 binds to regulatory elements near differentiation genes that are more accessible in cells with higher OCT4 levels. In contrast, loci overlapping pluripotency-associated super-enhancers were unaffected by SOX2 and OCT4 levels (Fig.4d). Loci in which accessibility was positively correlated to SOX2 and OCT4 levels were enriched for SOX2 and OCT4 binding, while those that were negatively correlated showed less overlap with SOX2 and OCT4 ChIP-seq peaks and were enriched for H3K4me3, marking active promoters and transcription start sites (Fig.4e). Note that while we used SOX2 and OCT4 peaks from two ChIP-seq datasets generated in the same conditions for meaningful comparison, which yielded ~30% of overlap of upregulated loci, including peaks from two additional OCT4 ChIP-seq datasets (with no comparable SOX2 data) increased this number to ~70% (Supplementary Fig. 5j), indicating that most of these loci are bound by OCT4 and possibly also by SOX2. Loci with increased accessibility in OCT4-high cells lose accessibility upon rapid OCT4 depletion (data from (King & Klose, 2017)), in line with OCT4 directly regulating accessibility at these sites (Supplementary Fig.5k) even though we cannot exclude that some loci may be regulated by indirect mechanisms. This suggests that fluctuations of endogenous SOX2 and OCT4 lead to temporal changes in chromatin accessibility, and that high OCT4 levels result in the opening of differentiation-associated enhancers.

Discussion

While gene expression fluctuations are increasingly being recognized as an important source of protein level variability in single cells, how these impact cellular functions remains largely unclear. Here we show that endogenous fluctuations in levels of pluripotency regulators have a major impact on ES cell differentiation potential. While Nanog displays prolonged, large amplitude fluctuations that alter ES cell differentiation potential in serum + LIF (Kalmar *et al*, 2009), these are caused by transitions between naïve and primed ES cells (Filipczyk *et al*, 2015) and thus reflect fluctuations between different phenotypic states. In contrast, ES cells maintained in a naïve state display small amplitude, transient fluctuations of SOX2 and OCT4, which nonetheless bias differentiation potential. Some of our findings such as enhancement of NE commitment by elevated SOX2 levels (Thomson *et al*, 2011; Wang *et al*, 2012; Zhao *et al*, 2004), or the impact of low OCT4 levels on deficient differentiation (Karwacki-Neisius *et al*, 2013; Radzisheuskaya *et al*, 2013), are in line with earlier studies. However, the enhancement of NE commitment by high endogenous OCT4 levels and the absence of SOX2 impact on ME commitment contrasts with previous work (Thomson *et al*, 2011). Thus, overexpression or indirect correlations from fixed samples have to be interpreted with caution to assess the function of cell fate regulators.

The large impact of small amplitude, transient OCT4 fluctuations on differentiation is surprising, suggesting a sensitive and rapid downstream mechanism modulating cell responsiveness to differentiation. While changes in the chromatin accessibility landscape as a function of OCT4 levels offers an explanation for these observations, the reason for differential responses of pluripotency regulatory elements as compared with differentiation-related enhancers is unclear. The potential role for cooperativity with other pluripotency TFs or differential affinity of OCT4 binding sites will require further investigation. We also observed widespread SOX2-dependent accessibility changes, including at many differentiation enhancer-associated loci. However, loci with increased accessibility in SOX2-high cells were less enriched for differentiation-

associated enhancers than in OCT4-high cells (Supplementary Fig.5d), which is in line with the weaker effects of SOX2 levels on differentiation. Finally, the fact that cells are more sensitive to SOX2 and OCT4 level variability in G1 phase than S phase raises the possibility that these transcription factors mainly act shortly after mitosis to re-open closed enhancer regions, in line with their reported pioneer transcription factor activity (Soufi *et al*, 2015) and their essential function in cell fate decisions at the mitosis-G1 transition (Deluz *et al*, 2016; Liu *et al*, 2017b).

Materials and Methods

Cell Culture

The different ES cell lines used in this study are listed in Table 1. The SBR (Deluz *et al*, 2016) and SBROS cell lines were generated from CGR8 ES cells (Sigma, Cat#07032901-1VL), and the E14 cell line (kindly provided by Didier Trono, EPFL) was used for all ES cell experiments. Cells were routinely cultured on dishes coated with 0.1% gelatin type B (Sigma, Cat#G9391-100G), in GMEM (Sigma, Cat#G5154-500ML) supplemented with 10% ES-cell qualified fetal bovine serum (Gibco, Cat#16141-079), nonessential amino acids (Gibco, Cat#11140-050), 2mM L-glutamine (Gibco, Cat#25030-024), sodium pyruvate (Sigma, Cat#S8636-100ML), 100µM 2-mercaptoethanol (Sigma, Cat#63689-25ML-F), penicillin and streptomycin (BioConcept, Cat#4-01F00-H), homemade leukemia inhibitory factor (LIF), CHIR99021 (Merck, Cat#361559-5MG) at 3µM and PD184352 (Sigma PZ0181-25MG) at 0.8µM. Cells were passaged by trypsinisation (Sigma, Cat#T4049-100ML) every two to three days at a ratio of 1:10.

For imaging experiments, ES cells were cultured on dishes coated with 5µg/ml E-Cadherin, in N2B27 medium supplemented with LIF, CHIR99021 at 3µM and PD184352 at 0.8µM (N2B27+2iLIF). E-Cadherin coating was performed as previously described(Nagaoka *et al*, 2006). Briefly, 5µg/ml E-Cadherin (R&D, Cat#8875-EC or

Cat#748-EC) in PBS (with Ca2+ and Mg2+; Amimed, Cat#3-05K00-I) were added to the culture vessel and incubated for 90 minutes at 37°C. Just before seeding, the E-Cadherin solution was removed, the surface of the vessel rinsed once with PBS and filled with the appropriate cell culture medium.

N2B27 medium was prepared by 1:1 mixing of DMEM/F12 (Gibco, Cat#11320-033) + N2 supplement (Gibco, Cat#17502-001) medium with Neurobasal (Gibco, Cat#21103-049) + B27 supplement (Gibco, Cat#17504-001) medium, supplemented with penicillin (1000IU/ml) and streptomycin (1000mg/ml), 2mM L-Glutamine and 0.1mM 2-mercaptoethanol.

HEK 293T cells (ATCC) were cultured in DMEM (Gibco, Cat#41966-029) supplemented with 10% fetal bovine serum (Gibco, Cat#10270-106), penicillin and streptomycin and passaged every 2 days at a ratio of 1:8.

For the selection of transduced and transfected cells, the following antibiotic concentrations were used: 8µg/ml of Blasticidin (Gibco A11139-03), 2µg/ml of Puromycin (Gibco A11138-03) and 200µg/ml of Hygromycin B (Invitrogen 10687010). ES cells were transfected using X-tremeGENE 9 transfection reagent (Sigma, Cat#06 365 809 001).

Cell Line	**Reference/Origin**
SBR	(Deluz *et al*, 2016)
SNSF	This study
Calibration cells	(Mandic *et al*, 2017)
2TS22C	(Masui *et al*, 2007)
Zhbtc4	(Niwa *et al*, 2000)
E14	Kindly provided by the laboratory of Didier trono, EPFL
CGR8	Sigma
HEK293T	ATCC

Table 1: Cell Lines used in this study

Generation of knock-in cell lines

The SBROS and SNSF cell lines were generated using CRISPR/Cas9-mediated homology-directed repair (HDR). The repair templates were designed to contain a knock-in cassette flanked by homology arms (HAs) with the target sequence missing the endogenous STOP codon. Guide RNA sequences were designed to overlap with the endogenous STOP codon, and the repair templates contain mutations in the PAM sequence, thus ensuring that the repair plasmids are not cut (Supplementary Figure 1a).

The knock-in cassette (between the HAs) contains the coding sequence for the tag in frame with the protein of interest and a selection marker. For the SOX2 knock-ins, the cassette consists of SNAP-IRES-Hygro (SBROS), or NLuc-loxP-P2A-Puro-sfGFP-loxP and NLuc-loxP-P2A-Bsd-sfGFP-loxP (SNSF). The knock-in cassette for OCT4 consists of a linker (WRAASRLTS)-Halo-IRES-Blasticidin (SBROS). The Sox1 knock-in cassette consists of P2A-FLuc-Stop-loxP-pGK-Hygro-Stop-loxP.

Guide RNAs targeting the *Oct4*, *Sox2* and *Sox1* loci were designed using the Zhang Lab toolbox (www.genome-engineering.org/crispr) and cloned into the pX330 vector (*Sox2* and *Oct4*), expressing Cas9 and the guide RNA, or pX335 (*Sox1*), expressing Cas9n and the guide RNA (Cong *et al*, 2013). The guide RNA sequences are listed in Table 2.

gRNA targets	**gRNA sequence (5’ to 3’)**
Oct4	GACTGAGGCACCAGCCCTCCC
Sox2	GCAGCCCTCACATGTGCGACA
Sox1	GACGCACATCTAGCGCCGCG

Table 2: sequences of guide RNAs used to generate knock-in cell lines

To generate the SBROS cell line, SBR cells (Deluz *et al*, 2016) were transfected with pX330-Sox2 and pKI-SOX2-SNAP-IRES-Hygro at a 1:3 ratio. After two days, selection was started with Hygromycin B. After 11 days of selection, cells were stained with 12nM SNAP-SiR 647, and single cells were sorted for high SNAP expression into 96-well plates and grown out. One clone identified as homozygously targeted by PCR on genomic DNA (Supplementary Fig.1b) was further validated by Western Blot (Supplementary Fig.1c) and used to knock-in a Halo-Tag at the C-terminus of OCT4. To do so, this clone was transfected with pX330-Pou5f1 and pKI-OCT4-HALO-IRES-Bsd at a 1:3 ratio followed by blasticidin selection two days after transfection. Single colonies were then picked manually and grown out, and one clone in which one allele was targeted as indicated by PCR on genomic DNA (Supplementary Fig.1b) was further analysed by Western Blot (Supplementary Fig.1c).

To generate the SNSF cell line, E14 cells were co-transfected with pX330-Sox2 and pKI-Sox2-NLuc-loxP-P2A-Puro-sfGFP-loxP at a 1:3 ratio. After two days, selection with Puromycin was started. After six days of selection, cells were co-transfected with pX330-Sox2 and pKI-Sox2-NLuc-loxP-P2A-Bsd-eGFP-loxP, and selection with Blasticidin was initiated two days later. A homozygous Blasticidin-resistant clone was identified by PCR on genomic DNA and subsequently recombined by transient transfection of a plasmid expressing Cre recombinase. Successful excision of the selection cassette was confirmed by PCR on genomic DNA (Supplementary Fig.1b). This resulting intermediate cell line (Sox2-NLuc) was co-transfected with pX335-Sox1 and pKI-Sox2-P2A-FLuc-loxP-pGK-Hygro-loxP, and selection with Hygromycin B was started two days later. Single clones were picked manually ten days later and the knock-in was confirmed using PCR on genomic DNA (Supplementary Fig.1b). The fusion of NLUC to SOX2 was confirmed by Western blotting (Supplementary Fig.1c).

All knock-in and corresponding wild-type alleles were verified by Sanger sequencing of the PCR products. All sequences were preserved except for the presence of a

single nucleotide insertion in the 3'UTR of the wild-type *Oct4* allele of the SBROS cell line.

Lentiviral Vector Production and generation of stable cell lines

Lentiviral vectors were produced by Calcium Phosphate co-transfection of HEK 293T cells with the envelope (PAX2), packaging (MD2G), and lentiviral construct of interest. The viral vectors were concentrated 120-fold by ultracentrifugation as described previously (Suter *et al*, 2006). Stable cell lines were generated by transducing 50,000 cells per well of a 24-well plate with 50µl of concentrated lentiviral vector particles. Antibiotic selection was started 48-72 hours later and maintained throughout passaging.

Alkaline phosphatase assays

ES cells were plated at 400 cells per well of a gelatinated 6-well plate in medium with 10% serum, 2i and LIF. The medium was exchanged every two to three days, and after 7 days alkaline phosphatase staining was performed following the manufacturer's instruction (Sigma, Cat#86R-1KT).

Pluripotency rescue colony scoring

Pluripotency rescue colony scoring was performed on the cell colonies right before alkaline phosphatase assays. Dox treatment of Zhbtc4 cells (inducible Oct4 knockout) results in an almost complete loss of colonies, while dox treatment of 2TS22C cells (inducible Sox2 knockout) results in a high number of flat colonies. Therefore, in Zhbtc4-derived cell lines, pluripotency rescue was scored by measuring the ratio of colonies obtained with and without dox treatment. In 2TS22C-derived cell lines, pluripotency rescue was scored by quantifying the ratio of dome-shaped versus flat colonies.

DNA constructs and cloning

To generate the pKI Sox2-Nluc-loxP-P2A-Puro-sfGFP-loxP, two multiple cloning sites, the downstream one including a loxP site, were inserted into a pCMV backbone by Oligo annealing. Next, a P2A construct was inserted by oligo annealing between a ClaI and a BamHI site, and the Nluc or Fluc coding sequence fused to a loxP site was cloned between SpeI and ClaI of the resulting plasmid. The selection cassette consisting of sfGFP-Puro or eGFP-Bsd was created using fusion PCR (Hobert, 2002) and inserted between a BamHI and a XhoI site. Subsequently, the Sox2 homology arms (HAs) were inserted. For the 5' HA, the PCR-amplified 5' HA and the backbone were digested with BsmBI and the 3' HA was inserted by In-Fusion cloning into an XbaI site. The pKI Sox1-P2A-Fluc-loxP-pGK-Hygro-loxP construct was based on a previously published plasmid pKI Sox1-P2A-loxP-eGFP-pGK-Hygro-loxP (Deluz *et al*, 2016), in which eGFP was replaced by Fluc using restriction cloning with AclI and SacI. For all pX330 (Addgene Cat#42230) and pX335 (Addgene Cat#42335) constructs, the vector was opened using BbsI and the guide RNAs were inserted using oligo annealing. The pKI Oct4-HALO-IRES-Bsd was based on pKI Oct4-Fluc-P2A-eGFP-Bsd, in which the Fluc-P2A-eGFP-Bsd cassette was replaced by Halo-IRESbsd using SpeI and NotI. To generate the pKI Sox2-SNAP-IRES-Hygro construct, pKI Sox2-Nluc-loxP-P2A-Puro-sfGFP-loxP was digested with NdeI and re-ligated, thus removing the Nluc-loxP-P2A-Puro-sfGFP cassette and resulting in pKI Sox2-NdeI. pLV pGK-rtTA3G-IRES-Hygro was digested with EcoRI and AgeI to remove the rtTA3G sequence, which was replaced with a P2A sequence by oligo annealing. The resulting vector pLV pGK-P2A-IRES-Hygro was digested with XbaI and the coding sequence for the SNAP-tag was inserted, resulting in the pLV pGK-P2A-SNAP-IRES-Hygro construct. Next, the P2A-SNAP-IRES-Hygro cassette was amplified by PCR and used for In-Fusion cloning with pKI-Sox2-NdeI linearized by

PCR from the start of the 3' HA to the end of the 5' HA of SOX2 (excluding the stop codon). The resulting pKI-Sox2-P2A-SNAP-IRES-Hygro vector was amplified by inverse PCR to remove the P2A-SNAP cassette, which was then replaced by a SNAP-tag using In-Fusion cloning, resulting in the pKI SOX2-SNAP-IRES-Hygro construct. The different constructs used in this study are summarized in Table 3.

Name	Reference
pKI Sox2-Nluc-loxP-P2A-Puro-sfGFP-loxP	This study
pKI Sox2-Nluc-loxP-P2A-Bsd-eGFP-loxP	This study
pKI Sox2-SNAP-IRES-Hygro	This study
pKI Oct4-HALO-IRES-Bsd	This study, Addgene #115689
pKI Sox1-P2A-FLuc-loxP-pGK-Hygro-loxP	This study
pX330 Sox2	This study
pX330 Oct4	This study
pX335 Sox1	This study
pLV TRE3G-Ypet-Sox2	(Deluz *et al*, 2016)
pLV TRE3G-Ypet-Oct4	(Deluz *et al*, 2016)
pLV TRE3G-SOX2-SNAP	(Deluz *et al*, 2016)
pLV TRE3G-YPet-Sox2-delDBD	(Deluz *et al*, 2016)
pLV pGK-rtTA3G-IRES-Hygro	(Mandic *et al*, 2017)
pLV pGK-rtTA3G-IRES-Bsd	(Deluz *et al*, 2016)
pLV-pGK-Cre	(Deluz *et al*, 2016)
pLV pGK-Nluc-Fluc-NLS-P2A-H2B-mCherry-IRES-Hygro	(Mandic *et al*, 2017), Addgene #115684

Table 3: DNA constructs used in this study

Confirmation PCRs

For all knock-in verification PCRs, genomic DNA was purified (Sigma, Cat#G1N350-1KT) and subsequently used to identify clones with correctly targeted alleles. PCR was done using Phusion High Fidelity DNA Polymerase (ThermoScientific, Cat#F530L). Primers are listed in Table 4.

Name	Sequence (5’ to 3’)	Target
Sox1-KI_F2	GTGCCCCTGACGCACAT	In 5’ HA
Sox1-KI_R2	CGCTGTGTGCCTCCTCTG	In 3’ HA
Seq_Sox2_KI_fw	AGGTGCCGGAGCCCG	In 5’HA
Sox2_verif_rv_3’	GCATGCTAGCCACAAAGAAA	Downstream of 3’ HA
Oct4_KI_verif_fw	GCTCCTCCACCCACCC	Within *Oct4*
(sf)GFP_verif_fw	CTCGGCATGGACGAGC	In sfGFP
Seqout_IRES_rv	AGACAGGGCCAGGTTTCC	In IRES
COct4_5'_500bp_f	GATCGTCGACTAGCACAATCCCTTAGCGGT	Upstream of 5’ HA

Table 4: Primers for knock-in verification

Western Blotting

Cells were trypsinised, collected by centrifugation and washed once in ice-cold PBS. 10 million cells were then resuspended in 500µl Hypotonic Buffer (20mM Tris-HCl pH7.4, 10mM NaCl, 3mM MgCl2), supplemented with 1mM PMSF (AppliChem, Cat#A0999.0005) and Protease Inhibitors (Sigma, Cat#P8340-5ML). After 15 minutes of incubation on ice, cells were lysed by the addition of 25µl of 10% NP-40 (Roche, Cat#11.754599001) and subsequent vortexing for 15 seconds. Nuclei were collected

by centrifugation (10 min, 4°C, 14.000 rpm) and lysed by resuspension in 30µl RIPA Buffer per million of cells (50mM Tris-HCl pH7.4, 1% NP-40, 0.5% Sodium Deoxycholate, 0.1% SDS, 2mM EDTA, 150mM NaCl), supplemented with PMSF (AppliChem, Cat#A0999.0005) and Protease Inhibitors (Sigma Cat#P8340-5ML). Lysis was allowed to proceed for 30 minutes during which cells were incubated on ice. Every 10 minutes, the sample was vortexed and further incubated on ice. To separate soluble nuclear proteins from debris, lysed nuclei were centrifuged for 30 minutes (4°C, 14.000rpm). The protein concentration of the supernatant was determined by performing a Bicinchoninic acid assay (BCA) (ThermoFisher, Cat#23235) and 15µg of protein were mixed with Laemmli sample buffer (Invitrogen, Cat#NP0007) and loaded on an SDS gel (BioRad, Cat#456-1094) for separation (SDS Running Buffer 25mM Tris, 190mM Glycine, 0.1%SDS). Proteins were subsequently transferred (Transfer Buffer 25mM Tris, 190mM Glycine, 20% Methanol, 0.1% SDS) from the gel onto a PVDF membrane (Merck, Cat#IPVH07850) using a wet transfer system. The membrane was blocked with 5% milk (Roth, Cat#T145.3) in PBS-T to reduce unspecific binding and incubated with the appropriate concentration of primary antibody overnight. The next day, the membrane was rinsed once with PBS-T, incubated with a secondary antibody in 5% milk in PBS-T, and washed extensively with PBS-T. Finally, chemiluminescence was revealed using Clarity Western ECL Substrate (BioRad, Cat#170-5060) and the signal was detected on a Fusion FX 7 apparatus (Vilber). The antibodies and concentrations used are summarized in Table 5.

Target	**Species**	**Dilution**	**Supplier**
Primary Antibodies			
αSOX2	Rabbit	1:250	Invitrogen 48-1400
αOCT4	Mouse	1:200	Santa Cruz sc5279
Secondary Antibodies			

αMouse-IgG-HRP	Goat	1:10.000	Promega W402B
αRabbit-IgG-HRP	Goat	1:10.000	Promega W401B

Table 5: Antibodies used for Western Blotting

Quantitative PCR (RT-QPCR)

Total RNA was extracted using the GenElute™ Mammalian Total RNA Miniprep Kit Q-PCR (Sigma, Cat#RTN350), and reverse transcription was performed using an oligoDT primer using superscript II (Thermofisher, Cat#18064014). QPCR was performed on a 7900HT Fast Real-Time PCR System (Thermofischer) with SYBR green reagent (Roche, Cat#04707516001). The *Rps9* cDNA was used for data normalization. Primers used for RT-QPCR are listed in Table 6.

Primer name	Primer sequence (5' to 3')
Nanog_qPCR_f	AACCAAAGGATGAAGTGCAAGCGG
Nanog_qPCR_r	TCCAAGTTGGGTTGGTCCAAGTCT
Esrrb_qPCR_f	GCACCTGGGCTCTAGTTGC
Esrrb_qPCR_r	TACAGTCCTCGTAGCTCTTGC
Sox2_qPCR_f	GCACATGAACGGCTGGAGCAACG
Sox2_qPCR_r	TGCTGCGAGTAGGACATGCTGTAGG
Oct4_qPCR_f	GAGGAGTCCCAGGACATGAA
Oct4_qPCR_r	AGATGGTGGTCTGGCTGAAC
Klf4_qPCR_f	CGAACTCACACAGGCGAGAA
Klf4_qPCR_r	CGGAGCGGGCGAATTT
Rps9_qPCR_f	TTGTCGCAAAACCTATGTGACC
Rps9_qPCR_r	GCCGCCTTACGGATCTTGG
Zfp42_qPCR_f	CCCTCGACAGACTGACCCTAA
Zfp42_qPCR_r	TCGGGGCTAATCTCACTTTCAT

Table 6: Primers used for QPCR

ChIP-seq

The 2TS22C EF1α-YPet-Sox2, Zhbtc4 EF1α-YPet-Oct4 and Zhbtc4 EF1α-Oct4-HALO cells were treated with 1 μg/ml doxycycline for 40-48 hours before fixation to ensure complete depletion of endogenous SOX2 (Masui *et al*, 2007) and OCT4 (Niwa *et al*, 2000). At least 10^7 cells were fixed in 1% formaldehyde for 10 minutes at room temperature, subsequently quenched with 200 mM Tris-HCl pH 8.0, washed with PBS, spun down, and stored at -80°C. The cell pellet was then resuspended in 1.5 ml LB1 (50 mM HEPES-KOH pH 7.4, 140 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 10% Glycerol, 0.5% NP40, 0.25% TritonX-100), incubated 10 min at 4°C, spun down, resuspended in 1.5 ml LB2 (10 mM Tris-HCl pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA), and incubated 10 min at 4°C. The pellet was spun down and rinsed twice with SDS shearing buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA, 0.15% SDS), and resuspended in 0.9 ml SDS shearing buffer. All buffers contained Protease Inhibitor Cocktail in DMSO (Sigma, Cat#P8340) diluted at 1:100. The suspension was transferred to a milliTUBE 1 ml AFA fiber and sonicated on a E220 focused ultrasonicator (Covaris) using the following settings: 20 min, 200 cycles, 5% duty, and 140W. Sonicated chromatin was incubated with 5□g/10^7 cells of the αSOX2 Y-17 (Santa Cruz, Cat#sc-17320) or 5μl of antibody/4x10^6 cells of the αOCT4A C30A3C1 (Cell Signaling Technology, Cat#5677S) at 4°C overnight. 0.5% BSA-blocked Protein G Dynabeads (Thermo Fischer, Cat#10003D) were added to the chromatin and incubated for 3 hours at 4°C. The chromatin was washed several times at 4°C with 5 min incubation between each wash and 2 min magnetization to collect beads; twice with Low Salt Wash Buffer (10 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.15% SDS, 1 mM PMSF), once with High Salt Wash Buffer (10 mM Tris-HCl pH 8.0, 500 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.15% SDS, 1 mM PMSF), once with LiCl Wash Buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA, 0.5 mM EGTA, 250 mM LiCl, 1% NP40, 1% sodium deoxycholate, 1 mM PMSF), and with TE

buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA). Beads were subsequently resuspended in Elution buffer (TE buffer with 1% SDS and 150 mM NaCl), treated with 400 ng/ml Proteinase K and reverse crosslinked at 65°C overnight. Samples were purified using MinElute PCR purification kit (Qiagen, Cat#28006). For SBR, Zhbtc4 EF1α-YPet-Oct4 and Zhbtc4 EF1α-Oct4-HALO, cells were first fixed with 2 mM Disuccinimidyl glutarate (DSG) (Thermo Fisher, Cat#20593) for 50 minutes in PBS at room temperature before proceeding with 1% formaldehyde fixation and chromatin immunoprecipitation as described above. One replicate of each ChIP-seq library was prepared with NEBNext ChIP-seq Library Prep Master Mix Set (NEB, Cat#E6240S) using insert size selection of 250 bp. All libraries were sequenced with 75-nucleotide read length paired-end sequencing on a Illumina NextSeq 500 with 30-50 million reads being sequenced for each sample.

Co-immunoprecipitation

The SBROS and SBR cell lines were plated at a density of about 2 million cells per 10cm dish, and 48 hours later these were treated with 1% PFA for 10 min. Cells were subsequently washed with PBS, and then collected in 1ml lysis buffer (50 mM Tris [pH 7.4], 150 mM NaCl, 1% Triton X-100, inhibitors for proteinase and phosphatase). The lysates were then incubated at 4°C with or without αSOX2 antibody (Santa Cruz, Cat#sc-17320) using 5μg/10^7 cells. The next day, 5% BSA-blocked Protein G Dynabeads (Thermo Fischer, Cat#10003D) were added to the samples, which were then washed four times in 1ml wash buffer (50 mM Tris pH 7.4, 500 mM NaCl, 1% Triton X-100, inhibitors for proteinase and phosphatase). Samples were subsequently incubated at 98°C for 10 min in 40μl of 2× Laemmli buffer, centrifuged at 2,000 × g for 3 min, and frozen at -80°C. Samples were then used for Western Blotting using αOct4 antibody at a dilution of 1:200 (Santa Cruz, Cat#sc5279).

Immunofluorescence and Image Acquisition

ES cells were fixed for 15 to 30 min with ice-cold 2% PFA (AppliChem, Cat#A0877,0500) in PBS, permeabilised and blocked with chilled PBS-Triton (0.5%, AppliChem, Cat#A1388,0500) and 1% FBS for 30 - 60 min. Samples were incubated with the primary antibody in PBS and 1% FBS overnight at 4°C, washed twice in PBS, and incubated with the secondary antibody in PBS and 1% FBS for 45 - 60 min. Samples were then washed three times with 0.1% PBS-Tween (Fisher Scientific, Cat#BP337-500), incubated with 1 µg/ml DAPI for 15 minutes, washed twice with 0.1% PBS-Tween and once with PBS. The antibodies used are listed in Table 7.

Immunofluorescence stainings were imaged using a 20x magnification objective (Olympus UPlanSApo 20x, NA 0.75) on an Olympus Cell XCellence or using a 20x magnification objective (Nikon PlanApo 20x, NA 0.75, CFI/60) on a GE InCell Analyzer 2200 apparatus. (GE, Cat#29027886, Biomolecular Screening Facility at EPFL)

Target	Species	Dilution	Supplier
Primary Antibodies			
αSOX2	Rabbit	1:200	Invitrogen Cat#48-1400
αOCT4	Mouse	1:50	Cell Signaling Cat#75463
αOCT4	Mouse	1:300	Santa Cruz Cat#sc5279
αNANOG	Rabbit	1:500	Abcam Cat#ab80892
Secondary Antibodies			
αMouse-IgG-555	Donkey	1:1000	ThermoFisher Cat#A31570
αMouse-IgG-488	Goat	1:1000	ThermoFisher Cat#A11001
αMouse-IgG-647	Goat	1:1000	ThermoFisher Cat#A21235
αRabbit-IgG-647	Chicken	1:1000	ThermoFisher Cat#A21443

Table 7: Antibodies used for Immunofluorescence staining

Live cell luminescence imaging

Time-lapse luminescence recordings were performed on an Olympus LuminoView LV200 microscope equipped with an EM-CCD cooled camera (Hamamatsu photonics, EM-CCD C9100-13), a 60x magnification objective (Olympus UPlanSApo 60x, NA 1.35, oil immersion) in controlled environment conditions (37°C, 5% CO2). One day before the experiment, cells were seeded on fluorodishes (WPI, Cat#FD35-100) coated with E-Cadherin in 2ml of N2B27+2iLIF as described above. For quantitative NANOLUC imaging, knock-in cells were mixed with Calibration cells at a ratio of 1:10 as described previously (Mandic *et al*, 2017). The medium was supplemented with 1mM Luciferin (NanoLight Technology, Cat#306A) and 0.5µl of RealTime Glo Cell Viability Assay Substrate (Promega, Cat#G9711). For imaging in pluripotency maintenance conditions, images were acquired every 299 s in the NANOLUC channel and between 59 and 178 s in Firefly Luciferase channel with a cycle time of 8 to 17 minutes for up to 75 hours. For overexpression experiments of SOX2-SNAP and YPet-SOX2-delDBD, the exposure time in the NANOLUC channel was 58 s and the cycle time 22 minutes. For imaging in differentiation conditions, images were acquired every 59 s in the NANOLUC channel and every 599 s in the FLUC channel with a cycle time of 15 to 16 minutes for up to 70 hours.

For time-lapse imaging of SOX2-NLUC / SOX1-P2A-Fluc, we used an arbitrary threshold of 500 FLUC molecules per cell for at least 4 hours to classify cells as Sox1-positive.

Fluorescence Time-lapse microscopy

For time-lapse imaging of the induction kinetics of SOX2-SNAP and YPet-SOX2delDBD as well as for OCT4-HALO single cell imaging, cells were seeded in E-Cadherin-coated wells of an imaging-grade 96-well plate in N2B27+2iLIF as described above. The next day, the medium was supplemented with the SNAP-SiR647 dye

(NEB, Cat#S9102S) for SOX2-SNAP imaging or HALO-SiR647 (gift from Suliana Manley, EPFL) for OCT4-HALO imaging at a concentration of 12nM for SOX2-SNAP imaging or 50nM for OCT4-HALO imaging. Cells were imaged using a 20x magnification objective (Olympus UPlanSApo 20x, NA 0.75) on an Olympus Cell XCellence in controlled conditions (37°C and 5% CO2) for 24 hours. For the induction experiments, doxycycline was added to a final concentration of 500ng/ml to induce transgene expression after one hour. For the OCT4-HALO imaging, since we observed a mild global fluorescence decrease in the cell population, we corrected for the loss of intensity on the population average to obtain single cell traces (Supplementary Fig.3e,f).

Flow Cytometry and Fluorescence-Activated Cell Sorting (FACS)

SBROS cells were stained with the HaloTag TMR Ligand (Promega, Cat#G8251) and the SNAP-SiR647 dye (NEB, Cat#S9102S) at a concentration of 100nM or 12nM, respectively, for one hour. Subsequently, cells were incubated in with Hoechst 33342 (Invitrogen, Cat#H3570) at a concentration of 1.62µM for 15 minutes. Cells were then trypsinised, washed in PBS and resuspended in PBS/1% FBS for sorting on a BD FACSAria II.

To determine how endogenous SOX2 levels regulate OCT4 expression, we used the following sorting strategy: Cells were gated for G1 based on Hoechst staining and on a narrow window of intermediate OCT4-HALO expression levels (~25% of cells). This window was further subdivided into two windows defined by the highest or lowest ~30% of SOX2-SNAP expression (Supplementary Fig.2k). The converse strategy was used to determine how endogenous OCT4 levels regulate SOX2 expression. After FACS, cells were spun down, resuspended in N2B27 medium and seeded in a gelatinated 24-well plate in N2B27 or N2B27+2iLIF. After 7 hours, cells were incubated with both TMR and SNAP-SiR647 dyes at a concentration of 100nM or 12nM, respectively, for one hour. Thereafter, cells were again stained with Hoechst

33342 for 15 minutes and subsequently trypsinized and collected by centrifugation for Flow Cytometry Analysis on a BD Fortessa or BD LSR II, followed by analysis using the FlowJo software (Supplementary Fig.2l).

To evaluate the impact of SOX2 or OCT4 levels on differentiation, we gated cells in G1 based on their Hoechst profile and defined three sub-bins of ~25% in SOX2-SNAP of all G1 cells (Supplementary Fig.4a) or into ~25% of OCT4-HALO high and low cells (Supplementary Fig.4b), respectively.

For the quadruple sorts based on a combination of SOX2/OCT4 high and low cells, we gated in all G1 or S phase cells on four windows corresponding to ~20% of the total cell population each (Supplementary Fig.4f).

Cells were washed once in PBS, trypsinised, collected by centrifugation and resuspended in PBS/1% FBS before flow cytometry analysis. All data acquisition was performed on a BD Fortessa and analysis was performed using the FlowJo software. E14 cells were used as a negative control to gate for Sox1-eGFP and Bra-mCherry.

In vitro Differentiation

For live-cell luminescence microscopy, cells were cultured in N2B27+2iLIF for at least two passages before 30,000 cells were seeded on E-Cadherin in fluorodishes and incubated in N2B27+2iLIF overnight. The next day, the medium was changed to N2B27 supplemented with 1mM luciferin and 0.5µl of RealTime Glo Cell Viability Assay Substrate, and image acquisition was started.

For differentiation assays after cell sorting, cells were seeded at a density of 60,000 cells/well of a 6-well plate coated with gelatin. Two days later, the medium was exchanged for fresh N2B27 and after four days differentiation outcomes were assessed by flow cytometry on a BD Fortessa.

For directed differentiation experiments, cells were seeded at a density of 60,000 cells/well of a gelatin-coated 12-well plate in N2B27 medium supplemented with 3µM

of CHIR99021. Three days later, differentiation outcomes were assessed by flow cytometry on a BD Fortessa.

ATAC-Seq

ATAC seq was performed on 50,000 cells for each condition as previously described (Buenrostro *et al*, 2013). Briefly, 50,000 cells were sorted by FACS, pelleted and washed with 1X ice cold PBS at 800g for 5 min. Cells were gently resuspended in 50 µl of ice-cold ATAC lysis buffer (10 mM Tris-HCl pH 7.4, 10 mM NaCl, 3 mM MgCl2, 0.1% NP40), and immediately pelleted at 800g for 10 min at 4°C. To transpose open chromatin regions, cells were resuspended in 50 µl of transposition reaction mix containing 0.5 µM of Tn5 transposase (Chen *et al*, 2017) (gift from Prof. Bart Deplancke lab, EPFL) in TAPS-DMF buffer (10 mM TAPS-NaOH, 5 mM Mgcl2, 10% DMF) and incubated at 37°C for 30 min. The transposed DNA was purified using a DNA purification kit (Zymo Research, Cat#D4003) and eluted in 12 µl of water. A 65 µl PCR reaction was setup with 10 µl of transposed DNA, 0.5 µM of forward primer Ad1_noMX, 0.5 µM of multiplexing reverse primer Ad2.x (Buenrostro *et al*, 2013), 0.6x SYBR® Green I, and 1x PCR Master Mix (NEB, Cat#M0544). The samples were thermocycled at 72°C for 5 minutes, 98°C for 30 s, followed by 5 cycles of 98°C for 10 s, 63°C for 30 s and 72°C for 1 min. A 15 µl aliquot was analyzed by qPCR to determine the number of additional cycles needed to avoid amplification saturation as described in (Buenrostro *et al*, 2013). The amplified ATAC libraries were purified using a DNA purification kit (Zymo Research, Cat#D4003) and size selected using Agencourt AMPure beads (Beckman Coulter, Cat#63881) (0.55X unbound fraction followed by 1.2X bound fraction). All libraries were sequenced with 75-nucleotide read length paired-end sequencing on an Illumina NextSeq 500 with 30-60 million reads being sequenced for each sample.

Immunofluorescence Image Analysis

Immunofluorescence images were first background-corrected using the built-in function in the Fiji software. Semi-automated image analysis was then performed using a custom CellProfiler (Kamentsky *et al*, 2011) pipeline. Images were segmented based on their DAPI signal and manually corrected for misidentified objects. Subsequently, fluorescence intensity was measured in the identified nuclei in all channels. The intensities were used to generate histograms of protein expression (NANOG, OCT4, and SOX2), to evaluate the effects of overexpression of OCT4 or SOX2 and to estimate the correlation between OCT4-HALO and total OCT4.

Cell Tracking and Single Cell Analysis

Cells were tracked manually using Fiji (ImageJ) by defining regions of interest (ROIs) throughout the movie. Next, all ROIs for a single cell were measured and the background (part of the image in the vicinity of the cell but devoid of cells) was subtracted. We used a previously reported method to convert the observed light intensity to absolute molecule numbers (Mandic *et al*, 2017).

To determine SOX2 levels in pluripotency conditions, cells were *in silico* synchronised for cell cycle progression using linear interpolation of the time variable, and absolute molecule numbers were converted to nuclear concentration, using a model for the nuclear size increase during the cell cycle (Filipczyk *et al*, 2015) and a reported estimate of the nuclear volume of ES cells (Chalut *et al*, 2012). To evaluate how cells readjust their SOX2 levels over time, we used a rank-based autocorrelation (Sigal *et al*, 2006) using data from cells tracked over one or two consecutive cell cycles. To compare the autocorrelation function between data tracked for one and two cell cycles, we selected 100 random single cell traces from the SOX2 data and calculated the protein memory based on a conservative mixing time estimation (Sigal *et al*, 2006). As the results using data from one and two cell cycles were similar, we used a

single cell cycle from the OCT4-HALO imaging to calculate the rank-based autocorrelation of OCT4-HALO.

To determine how SOX2 levels predict neuroectodermal differentiation, we classified tracked cell cycles based on their FLUC signal in four groups, using an arbitrary threshold of 500 AU in FLUC maintained for at least four hours: “negative cells” were defined as cells below the threshold throughout the movie; “before SOX1+” were defined as negative cells that pass the threshold in the next cell cycle; “turning SOX1+” were defined as cells passing the threshold in the current cell cycle; “SOX1+” were defined as cells with FLUC levels above the threshold. The “before SOX1+” cell population also contains traces that did not cover a full cell cycle before becoming Sox1-positive. All single cell traces were *in silico* synchronised using a linear interpolation of the time variable. A two-sample t-test with unequal variance (MatLab) was performed for the mean SOX2 levels in the cell cycle before cells turn SOX1 positive to evaluate statistical significance.

To determine the induction kinetics in the YPet-SOX2-delDBD and SOX2-SNAP overexpressing cell lines, single cells were tracked over divisions in one daughter cell.

ATAC-seq and ChIP-seq analysis

ATAC-seq and ChIP-seq reads were aligned to the mouse reference genome mm10 using STAR (Dobin *et al*, 2013) with settings ‘--alignMatesGapMax 2000 --alignIntronMax 1 --alignEndsType EndtoEnd --outFilterMultimapNmax 1’. Duplicate reads were removed with Picard and reads not mapping to chromosomes 1-19, X, or Y were removed. For each sample, peaks were called with MACS2 (Zhang *et al*, 2008) with settings ‘-f BAMPE -q 0.01 -g mm’. For ATAC-seq data, peaks from all samples were merged with BEDOPS (Neph *et al*, 2012). Peaks overlapping peaks called for ChIP-seq Input data from asynchronous mouse ES cells (GSE89599) were discarded. The HOMER2 (Heinz *et al*, 2010) function annotatePeaks.pl was used with settings ‘-noadj -len 0 -size given’ to count the number of reads for each sample in

peaks. Analysis of differentially abundant regions was done with edgeR (Robinson *et al*, 2010) and limma (Ritchie *et al*, 2015) using TMM normalization. The analysis was done using three different contrasts; (i) SHOH vs SLOL, design ~0+Condition+Replicate, (ii) SOX2 high vs SOX2 low, design ~0+SOX2+OCT4+Replicate, (iii) OCT4 high vs OCT4 low, design ~0+OCT4+SOX2+Replicate. Regions with an adjusted p-value < 0.1 for at least one test were used in the analysis. Groupings were made according to fold-change direction and if loci were significantly different for OCT4 high vs OCT4 low only (OCT4 regulated), SOX2-high vs SOX2-low only (SOX2 regulated), or regulated by both SOX2 and OCT4 or OCT4-high/SOX2-high vs OCT4-low/SOX2-low (co-regulated). SOX2 and OCT4 peaks used to determine overlap were taken from GSE87820 and GSE92846 (King & Klose, 2017; Liu *et al*, 2017b). Additional OCT4 ChIP-seq data was taken from GSE56138 (Buecker *et al*, 2014). H3K4me3 peaks in ES-Bruce4 cells from ENCODE (An integrated encyclopedia of DNA elements in the human genome, 2012) and ES-cell super-enhancers from (Whyte *et al*, 2013) were converted to mm10 using liftOver. Gene ontology analysis was done using the closest UCSC-annotated gene to each peak with Fisher's exact test in topGO using genes closest to all peaks as background. bigWig files for both ATAC-seq and ChIP-seq were generated by merging replicate bam files with SAMTools (Li *et al*, 2009) followed by the deepTools (Ramírez *et al*, 2014) functions bamCoverage (with setting '--normalizeUsingRPKM'). Average lineplots and ChIP-seq heatmaps were generated using deepTools computeMatrix (with setting 'reference-point') and plotHeatmap. The wt SOX2 ChIP-seq profile in Supplementary Fig. 1i is based on data from GSE89599 (Deluz *et al*, 2016). Genome tracks were generated in the UCSC genome browser.

Determination of SOX2 and OCT4 protein half-lives

SBROS cells were seeded at 30,000 cells/cm^2 on E-cadherin in N2B27 medium. After 24 h, cells were pulse-labeled with 12 nM of SNAP-SiR 647 or 20 nM Halo-SiR647

ligand (gift from Suliana Manley, EPFL) for 30 min at 37 °C. Cells were then washed 3x with PBS and incubated in N2B27 medium for 15 min at 37 °C. This washing step was repeated once more, then cells were washed 2x with PBS and phenol-free N2B27 medium was added. The decay of the fluorescence signal was imaged using an InCell Analyzer 2200 microscope (GE Healthcare Life Sciences) with a 20x Objective, 10% laser power, 300 ms exposure (Cy5: excitation filter 632/22 nm, emission filter 679/34 nm), and 2x2 binning for 12.5 h at intervals of 15 min. Images were analysed in FiJi, where the background was subtracted from all images in the stack (rolling ball radius = 50 pixels). The integrated fluorescence intensity was then quantified by manual tracking. ROIs were drawn around each cell of interest at each time point and the integrated fluorescence intensity was calculated by multiplying the area of each ROI with its mean intensity. The local background was calculated by drawing a ROI close to the cell of interest in each time frame, followed by multiplying its mean intensity with the area of the cellular ROI. The background intensity was subtracted from the cellular intensity for each time frame. In case of cell divisions, both daughter cells were tracked separately and their integrated intensities were summed. Each fluorescence intensity trace was normalized to the value of the first frame and the single cell decay rates (b) were determined by exponential curve fitting, using the curve fitting tool in Matlab (fitted equation: $f(t) = e^{-bt}$). Half-lives were then calculated as follows: $T_{1/2} = b/\ln(2)$. Single cell half-lives in 20 cells were quantified for SOX2-SNAP, OCT4-HALO and OCT4-HALO after 6-8 hours of SOX2 overexpression.

Statistical analysis

The cell cycle times (Fig.1f and Supplementary Fig.2j) were compared using two-sided t-tests, showing no statistical significance. The delta Ct values in Fig.2b were analysed using two-sided t-tests comparing the *Sox2* or *Oct4* mRNA levels between 0 and 2 hours as well as 0 and 6 hours. The OCT4-HALO half-lives (Fig.2c) were assessed using a two-sided t-test with unequal variance. Differences between SOX2-

low and SOX2-high or OCT4-low and OCT4-high were analysed using a one-sided t-test (Fig.2e,f). For the autocorrelation functions in Figures 2i,l and Supplementary Fig.3d the error bars denote the SE estimated by bootstrapping. In Fig.3c, a two-sided t-test with unequal variance was performed. For Fig.3e-g,i-k , and Supplementary Fig. 4e,i,j, all depicted tests are two-sided t-tests with unequal variance.

Code availability

The semi-automated image analysis pipeline is available from the corresponding author upon reasonable request.

Data Availability

Data supporting the findings of this manuscript are available from the corresponding author upon reasonable request. ATAC-seq and ChIP-seq data from this study have been deposited in GEO (Gene Expression Omnibus) with the accession code GSE126554 (reviewer access token afsngogsplmbvcp).

Acknowledgments

This work was supported by the Swiss National Science Foundation (PP00P3_1144828 and PP00P3_172905) to D.M.S, the Pierre Mercier Foundation, and the generous support of the Fondazione Teofil Rossi di Montelera e di Premuda and an anonymous donor advised by CARIGEST SA. We thank the Swiss Federal Institute of Technology (EPFL), the Biomolecular Screening Facility (EPFL-BSF) and the EPFL Bioimaging and Optics Core Facility (EPFL-BIOP) for assistance in imaging and the EPFL Flow Cytometry Core Facility (EPFL-FCCF) for the fluorescence activated cell sorting. We thank Antonio Meireles-Filho for help with ATAC-seq experiments.

Author contributions:

Conceptualization, D.S and D.M.S.; Methodology, D.S, D.M.S., E.T.F, S.G., A.B.A.; Formal Analysis, D.S, D.M.S., E.T.F., A.B.A.; Investigation, D.S, D.M.S., C.D., E.T.F, S.G., A.B.A.; Resources, D.M.S.; Writing – Original Draft, D.S and D.M.S.; Writing – Review & Editing, D.S, D.M.S., C.D., E.T.F, S.G., A.B.A., Funding Acquisition, D.M.S.; Supervision, D.M.S.

Conflicts of interests

The authors declare no competing interests.

References

Alber AB, Paquet ER, Biserni M, Naef F & Suter DM (2018) Single Live Cell Monitoring of Protein Turnover Reveals Intercellular Variability and Cell-Cycle Dependence of Degradation Rates. *Mol. Cell* **71:** 1079-1091.e9

An integrated encyclopedia of DNA elements in the human genome (2012) *Nature* **489:** 57–74

van den Berg DLC, Snoek T, Mullin NP, Yates A, Bezstarosti K, Demmers J, Chambers I & Poot RA (2010) An Oct4-centered protein interaction network in embryonic stem cells. *Cell Stem Cell* **6:** 369–381

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

Buenrostro JD, Giresi PG, Zaba LC, Chang HY & Greenleaf WJ (2013) Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10:** 1213–1218

Chalut KJ, Höpfler M, Lautenschläger F, Boyde L, Chan CJ, Ekpenyong A, Martinez-Arias A & Guck J (2012) Chromatin Decondensation and Nuclear Softening Accompany Nanog Downregulation in Embryonic Stem Cells. *Biophys. J.* **103:** 2060–2070

Chen W, Gardeux V, Meireles-Filho A & Deplancke B (2017) Profiling of Single-Cell Transcriptomes. *Curr. Protoc. Mouse Biol.* **7:** 145–175

Chew J-L, Loh Y-H, Zhang W, Chen X, Tam W-L, Yeap L-S, Li P, Ang Y-S, Lim B, Robson P & Ng H-H (2005) Reciprocal Transcriptional Regulation of Pou5f1 and Sox2 via the Oct4/Sox2 Complex in Embryonic Stem Cells. *Mol. Cell. Biol.* **25:** 6031–6046

Cohen DE & Melton D (2011) Turning straw into gold: directing cell fate for regenerative medicine. *Nat Rev Genet* **12:** 243–52

Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA & Zhang F (2013) Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* **339:** 819–823

Coronado D, Godet M, Bourillot P-Y, Tapponnier Y, Bernat A, Petit M, Afanassieff M, Markossian S, Malashicheva A, Iacone R, Anastassiadis K & Savatier P (2013) A short G1 phase is an intrinsic determinant of naïve embryonic stem cell pluripotency. *Stem Cell Res.* **10:** 118–131

Deluz C, Friman ET, Strebinger D, Benke A, Raccaud M, Callegari A, Leleu M, Manley S & Suter DM (2016) A role for mitotic bookmarking of SOX2 in pluripotency and differentiation. *Genes Dev.* **30:** 2538–2550

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M & Gingeras TR (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinforma. Oxf. Engl.* **29:** 15–21

Fang L, Zhang L, Wei W, Jin X, Wang P, Tong Y, Li J, Du JX & Wong J (2014) A methylation-phosphorylation switch determines Sox2 stability and function in ESC maintenance or differentiation. *Mol. Cell* **55:** 537–551

Filipczyk A, Marr C, Hastreiter S, Feigelman J, Schwarzfischer M, Hoppe PS, Loeffler D, Kokkaliaris KD, Endele M, Schauberger B, Hilsenbeck O, Skylaki S, Hasenauer J, Anastassiadis K, Theis FJ & Schroeder T (2015) Network plasticity of pluripotency transcription factors in embryonic stem cells. *Nat. Cell Biol.* **17:** 1235–1246

Goolam M, Scialdone A, Graham SJL, Macaulay IC, Jedrusik A, Hupalowska A, Voet T, Marioni JC & Zernicka-Goetz M (2016) Heterogeneity in Oct4 and Sox2 Targets Biases Cell Fate in 4-Cell Mouse Embryos. *Cell* **165:** 61–74

Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H & Glass CK (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38:** 576–589

Hobert O (2002) PCR fusion-based approach to create reporter gene constructs for expression analysis in transgenic C. elegans. *BioTechniques* **32:** 728–730

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

Kamentsky L, Jones TR, Fraser A, Bray M-A, Logan DJ, Madden KL, Ljosa V, Rueden C, Eliceiri KW & Carpenter AE (2011) Improved structure, function and compatibility for CellProfiler: modular high-throughput image analysis software. *Bioinforma. Oxf. Engl.* **27:** 1179–1180

Karwacki-Neisius V, Göke J, Osorno R, Halbritter F, Ng JH, Weiße AY, Wong FCK, Gagliardi A, Mullin NP, Festuccia N, Colby D, Tomlinson SR, Ng H-H & Chambers I (2013) Reduced Oct4 Expression Directs a Robust Pluripotent State with Distinct Signaling Activity and Increased Enhancer Occupancy by Oct4 and Nanog. *Cell Stem Cell* **12:** 531–545

King HW & Klose RJ (2017) The pioneer factor OCT4 requires the chromatin remodeller BRG1 to support gene regulatory element function in mouse embryonic stem cells. *eLife* **6:**

Kopp JL, Ormsbee BD, Desler M & Rizzino A (2008) Small increases in the level of Sox2 trigger the differentiation of mouse embryonic stem cells. *Stem Cells Dayt. Ohio* **26:** 903–911

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R & 1000 Genome Project Data Processing Subgroup (2009) The Sequence Alignment/Map format and SAMtools. *Bioinforma. Oxf. Engl.* **25:** 2078–2079

Liu L, Michowski W, Inuzuka H, Shimizu K, Nihira NT, Chick JM, Li N, Geng Y, Meng AY, Ordureau A, Kołodziejczyk A, Ligon KL, Bronson RT, Polyak K, Harper JW, Gygi SP, Wei W & Sicinski P (2017a) G1 cyclins link proliferation, pluripotency and differentiation of embryonic stem cells. *Nat. Cell Biol.* **19:** 177–188

Liu Y, Pelham-Webb B, Di Giammartino DC, Li J, Kim D, Kita K, Saiz N, Garg V, Doane A, Giannakakou P, Hadjantonakis A-K, Elemento O & Apostolou E (2017b) Widespread Mitotic Bookmarking by Histone Marks and Transcription Factors in Pluripotent Stem Cells. *Cell Rep.* **19:** 1283–1293

Loh Y-H, Wu Q, Chew J-L, Vega VB, Zhang W, Chen X, Bourque G, George J, Leong B, Liu J, Wong K-Y, Sung KW, Lee CWH, Zhao X-D, Chiu K-P, Lipovich L, Kuznetsov VA, Robson P, Stanton LW, Wei C-L, et al (2006) The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat. Genet.* **38:** 431–440

Mandic A, Strebinger D, Regali C, Phillips NE & Suter DM (2017) A novel method for quantitative measurements of gene expression in single living cells. *Methods San Diego Calif* **120:** 65–75

Masui S, Nakatake Y, Toyooka Y, Shimosato D, Yagi R, Takahashi K, Okochi H, Okuda A, Matoba R, Sharov AA, Ko MS & Niwa H (2007) Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. *Nat Cell Biol* **9:** 625–35

Mistri TK, Arindrarto W, Ng WP, Wang C, Lim LH, Sun L, Chambers I, Wohland T & Robson P (2018) Dynamic changes in Sox2 spatio-temporal expression promote the second cell fate decision through Fgf4/Fgfr2 signaling in preimplantation mouse embryos. *Biochem. J.* **475:** 1075–1089

Nagaoka M, Koshimizu U, Yuasa S, Hattori F, Chen H, Tanaka T, Okabe M, Fukuda K & Akaike T (2006) E-cadherin-coated plates maintain pluripotent ES cells without colony formation. *PloS One* **1:** e15

Neph S, Kuehn MS, Reynolds AP, Haugen E, Thurman RE, Johnson AK, Rynes E, Maurano MT, Vierstra J, Thomas S, Sandstrom R, Humbert R & Stamatoyannopoulos JA (2012) BEDOPS: high-performance genomic feature operations. *Bioinforma. Oxf. Engl.* **28:** 1919–1920

Niwa H, Miyazaki J & Smith AG (2000) Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat. Genet.* **24:** 372–376

Okumura-Nakanishi S, Saito M, Niwa H & Ishikawa F (2005) Oct-3/4 and Sox2 regulate Oct-3/4 gene in embryonic stem cells. *J. Biol. Chem.* **280:** 5307–5317

Ormsbee Golden BD, Wuebben EL & Rizzino A (2013) Sox2 expression is regulated by a negative feedback loop in embryonic stem cells that involves AKT signaling and FoxO1. *PloS One* **8:** e76345

Pan X, Cang X, Dan S, Li J, Cheng J, Kang B, Duan X, Shen B & Wang Y-J (2016) Site-specific Disruption of the Oct4/Sox2 Protein Interaction Reveals Coordinated Mesendodermal Differentiation and the Epithelial-Mesenchymal Transition. *J. Biol. Chem.* **291:** 18353–18369

Pauklin S & Vallier L (2013) The cell-cycle state of stem cells determines cell fate propensity. *Cell* **155:** 135–47

Raccaud M, Friman ET, Alber AB, Agarwal H, Deluz C, Kuhn T, Gebhardt JCM & Suter DM (2019) Mitotic chromosome binding predicts transcription factor properties in interphase. *Nat. Commun.* **10:** 487

Radzisheuskaya A, Chia GLB, dos Santos RL, Theunissen TW, Castro LFC, Nichols J & Silva JCR (2013) A defined Oct4 level governs cell state transitions of pluripotency entry and differentiation into all embryonic lineages. *Nat. Cell Biol.* **15:** 579–590

Ramírez F, Dündar F, Diehl S, Grüning BA & Manke T (2014) deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* **42:** W187-191

Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W & Smyth GK (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43:** e47

Robinson MD, McCarthy DJ & Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinforma. Oxf. Engl.* **26:** 139–140

Shahbazi MN, Scialdone A, Skorupska N, Weberling A, Recher G, Zhu M, Jedrusik A, Devito LG, Noli L, Macaulay IC, Buecker C, Khalaf Y, Ilic D, Voet T, Marioni JC & Zernicka-Goetz M (2017) Pluripotent state transitions coordinate morphogenesis in mouse and human embryos. *Nature* **552:** 239–243

Sigal A, Milo R, Cohen A, Geva-Zatorsky N, Klein Y, Liron Y, Rosenfeld N, Danon T, Perzov N & Alon U (2006) Variability and memory of protein levels in human cells. *Nature* **444:** 643–6

Soufi A, Garcia MF, Jaroszewicz A, Osman N, Pellegrini M & Zaret KS (2015) Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming. *Cell* **161:** 555–568

Suter DM, Cartier L, Bettiol E, Tirefort D, Jaconi ME, Dubois-Dauphin M & Krause K-H (2006) Rapid generation of stable transgenic embryonic stem cell lines using modular lentivectors. *Stem Cells Dayt. Ohio* **24:** 615–623

Tapia N, MacCarthy C, Esch D, Gabriele Marthaler A, Tiemann U, Araúzo-Bravo MJ, Jauch R, Cojocaru V & Schöler HR (2015) Dissecting the role of distinct OCT4-SOX2 heterodimer configurations in pluripotency. *Sci. Rep.* **5:** Available at: http://www.nature.com/articles/srep13533 [Accessed November 16, 2017]

Thomson M, Liu SJ, Zou L-N, Smith Z, Meissner A & Ramanathan S (2011) Pluripotency Factors in Embryonic Stem Cells Regulate Differentiation into Germ Layers. *Cell* **145:** 875–889

Turner DA, Rué P, Mackenzie JP, Davies E & Martinez Arias A (2014) Brachyury cooperates with Wnt/β-catenin signalling to elicit primitive-streak-like behaviour in differentiating mouse embryonic stem cells. *BMC Biol.* **12:** 63

Wang Z, Oron E, Nelson B, Razis S & Ivanova N (2012) Distinct Lineage Specification Roles for NANOG, OCT4, and SOX2 in Human Embryonic Stem Cells. *Cell Stem Cell* **10:** 440–454

White MD, Angiolini JF, Alvarez YD, Kaur G, Zhao ZW, Mocskos E, Bruno L, Bissiere S, Levi V & Plachta N (2016) Long-Lived Binding of Sox2 to DNA Predicts Cell Fate in the Four-Cell Mouse Embryo. *Cell* **165:** 75–87

Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, Rahl PB, Lee TI & Young RA (2013) Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153:** 307–319

Wolff SC, Kedziora KM, Dumitru R, Dungee CD, Zikry TM, Beltran AS, Haggerty RA, Cheng J, Redick MA & Purvis JE (2018) Inheritance of OCT4 predetermines fate choice in human embryonic stem cells. *Mol. Syst. Biol.* **14:** e8140

Yang H, Wang H, Shivalila CS, Cheng AW, Shi L & Jaenisch R (2013) One-Step Generation of Mice Carrying Reporter and Conditional Alleles by CRISPR/Cas-Mediated Genome Engineering. *Cell* **154:** 1370–1379

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W & Liu XS (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9:** R137

Zhao S, Nichols J, Smith AG & Li M (2004) SoxB transcription factors specify neuroectodermal lineage choice in ES cells. *Mol. Cell. Neurosci.* **27:** 332–342

Figure Legends

Figure 1: Knock-in cell lines

a Scheme of the knock-in alleles of the SOX1-BRA-Reporter-OCT4-HALO-SOX2-SNAP (SBROS) cell line. Red hexagons: STOP codons. Yellow triangles: loxP sites. **b** Example images showing the localization of SOX2-SNAP and OCT4-HALO by fluorescence microscopy of SBROS cells. Scale bar: 50µm. **c** Correlation between OCT4-HALO and total OCT4 levels determined by immunofluorescence and HALO labelling. **d** Distributions of OCT4 (n=2682), SOX2 (n=1236) and NANOG (n=2416) levels in wt E14 and SBROS cell lines as determined by quantitative immunofluorescence. Dashed lines: mean protein levels; Dotted lines: median protein levels. **e** Protein half-lives of OCT4-HALO and SOX2-SNAP in mouse ES cells. Whiskers: minimum and maximum values; Box: lower and upper quartiles; Solid line: median. Black dots: outliers **f** Cell cycle duration of E14 wt (n=6) and SBROS (n=8) cells. Error bars: SE. p value: two-sided t-test.

Figure 2: Cross-regulation and single-cell fluctuations of SOX2 and OCT4

a Endogenous SOX2 levels in single cells (turquoise) normalized to the value at t=0 in SNSF cells upon induction of SOX2-SNAP (n=37) or a truncated SOX2 version missing the DNA binding domain (YPET-SOX2-delDBD, n=47). Black line: average SOX2-SNAP levels after induction normalized on the maximum level (n=22); Yellow line: average YPET-SOX2-delDBD levels after induction, normalized on the maximum level (n=20). **b** *Sox2* and *Oct4* mRNA levels upon overexpression of YPet-SOX2 or YPET-SOX2-delDBD (n=3, p-values determined by two-sided t-test). **c** Changes of OCT4 half-life upon SOX2 overexpression (n=20 cells each, p-values determined by two-sided t-test). OE: overexpression. **d** Sorting strategy to evaluate the impact of endogenous fluctuations of SOX2 and OCT4 on OCT4 and SOX2 protein levels, respectively. **e** Changes of SOX2 and OCT4 levels after sorting for high and low SOX2 levels in cells expressing intermediate OCT4 levels (n=4; error bars: SE; p-values determined by two-sided t-test). **f** Changes of SOX2 and OCT4 levels after sorting for high and low OCT4 levels in cells expressing intermediate SOX2 levels (n=4, p-values determined by two-sided t-test). **g** Representative traces of the absolute number of SOX2-NLUC molecules (top) and inferred concentration in nM (bottom) in *in-silico* synchronized cells (n=10). **h** Single cells were ranked according to their SOX2 expression at t=0 and assigned a color code for their initial rank from low (blue) to high (red) levels, and changes in ranks over time are shown (n=59). **i** Rank-based autocorrelation function of the SOX2 ranks (n=59). Error bars: SE estimated by bootstrapping. **j** Representative single cell traces of the integrated intensity of OCT4-HALO in single cells (top) and the corresponding inferred concentration (bottom) in *in-silico* synchronized cells (n=10). **k** Representation of the ranks as in g. **l** Rank-based autocorrelation function as in i (n=47). Error bars: SE estimated by bootstrapping.

Figure 3: Endogenous SOX2 and OCT4 level fluctuations bias differentiation

a Scheme of single-cell tracking for SNSF cells turning SOX1-positive (turquoise cytoplasm) or remaining SOX1-negative throughout the experiment. Turquoise dashed contour: measured region for cells turning SOX1 positive; Red dashed contour: measured region for cells staying SOX1 negative. **b** Snapshot from luminescence movie of SNSF cells in differentiation. White arrowheads: cells starting to express SOX1-P2A-FLUC; Orange arrowheads: cells staying SOX1-negative throughout the experiment. Red: SOX2-NLUC, turquoise: Sox1-P2A-FLUC. **c** Single cell traces of SOX2 levels for cells in differentiation. Turquoise traces: cells becoming SOX1 positive in the middle cell cycle, average is shown in bold black; Red traces: cells remaining SOX1-negative throughout the experiment, average is shown in bold; Green line: average FLUC signal in cells that become SOX1-positive. * $p<0.05$ determined by two-sided t.test with unequal variance. **d** Strategies to sort SBROS cell populations with different SOX2 and OCT4 levels. **e** Differentiation outcome for cells sorted for different SOX2 levels (n=5 for low and high, n=4 for medium). **f** Differentiation outcome for cells sorted for different OCT4 levels (n=4). **g** Differentiation outcome of uninduced cells or cells induced 12 hours for OCT4-SNAP overexpression before sorting in G1 phase (n=6). **h** Sorting strategy for SBROS cells in G1 or S phase and into the four following subpopulations: SOX2 high & OCT4 high (SHOH), SOX2 high & OCT4 low (SHOL), SOX2 low & OCT4 high (SLOH) or SOX2 low & OCT4 low (SLOL). **i** Differentiation outcome for cells sorted in G1 phase (n=5). **j** Differentiation outcome for cells sorted in S phase (n=5). **k** Differentiation outcome for cells sorted in G1 phase in the presence of 3μM CHIR (n=7).

Figure 4: SOX2 and OCT4-dependent changes in chromatin accessibility

a Example tracks of the ATAC-seq signal in SHOH, SHOL, SLOH and SLOL populations for loci of which chromatin accessibility is unaffected, affected by OCT4 levels, SOX2 levels, or affected by both. **b** Heatmap of relative chromatin accessibility in SHOH/SHOL/SLOH/SLOL samples for all loci with significantly altered accessibility, with examples of genes close to affected regulatory regions. Groups are defined by their significance (OHvsOL/SHvsSL/SHOHvsSLOL) and fold-change direction. **c** Gene Ontology (GO) term enrichment analysis for genes close to regions with higher accessibility in OCT4-high cells. **d** Chromatin accessibility ratio of SHOH vs SLOL cells in loci overlapping ES cell super-enhancer regions. **e** Percentage of regions in each group showing an overlap with OCT4, SOX2 and H3K4me3 ChIP-seq peaks.

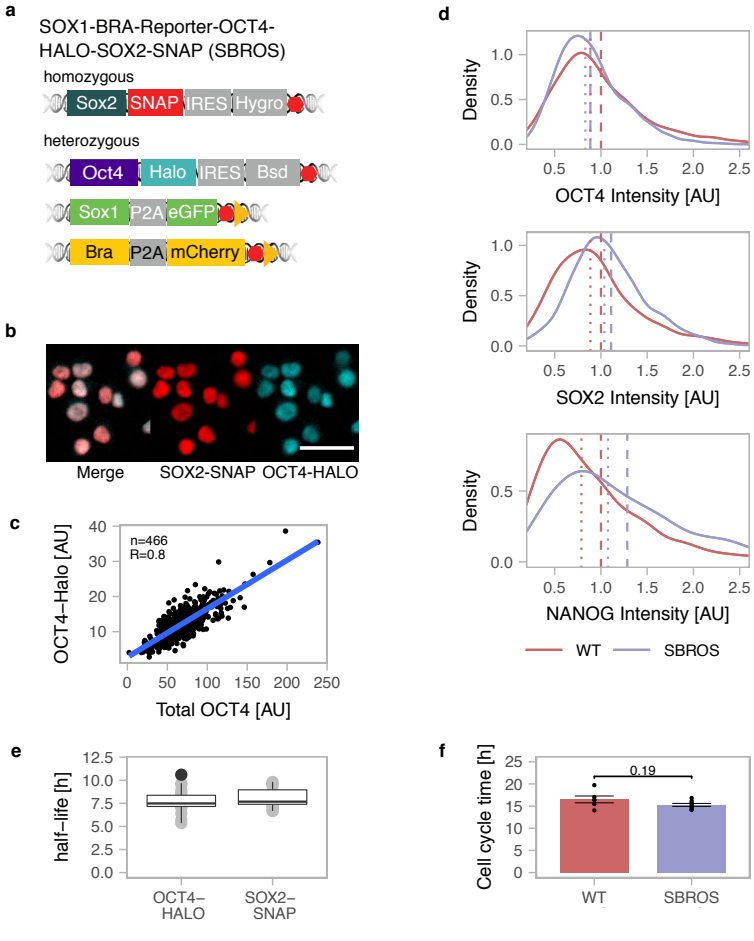
a
SOX1-BRA-Reporter-OCT4-HALO-SOX2-SNAP (SBROS)
homozygous
Sox2
SNAP
IRES
Hygro
heterozygous
Oct4
Halo
IRES
Bsd
Sox1
P2A
eGFP
Bra
P2A
mCherry
b
Merge
SOX2-SNAP
OCT4-HALO
c
n=466
R=0.8
OCT4−Halo [AU]
Total OCT4 [AU]
d
Density
OCT4 Intensity [AU]
SOX2 Intensity [AU]
NANOG Intensity [AU]
WT
SBROS
e
half−life [h]
OCT4−HALO
SOX2−SNAP
f
Cell cycle time [h]
0.19
WT
SBROS

Figure 1

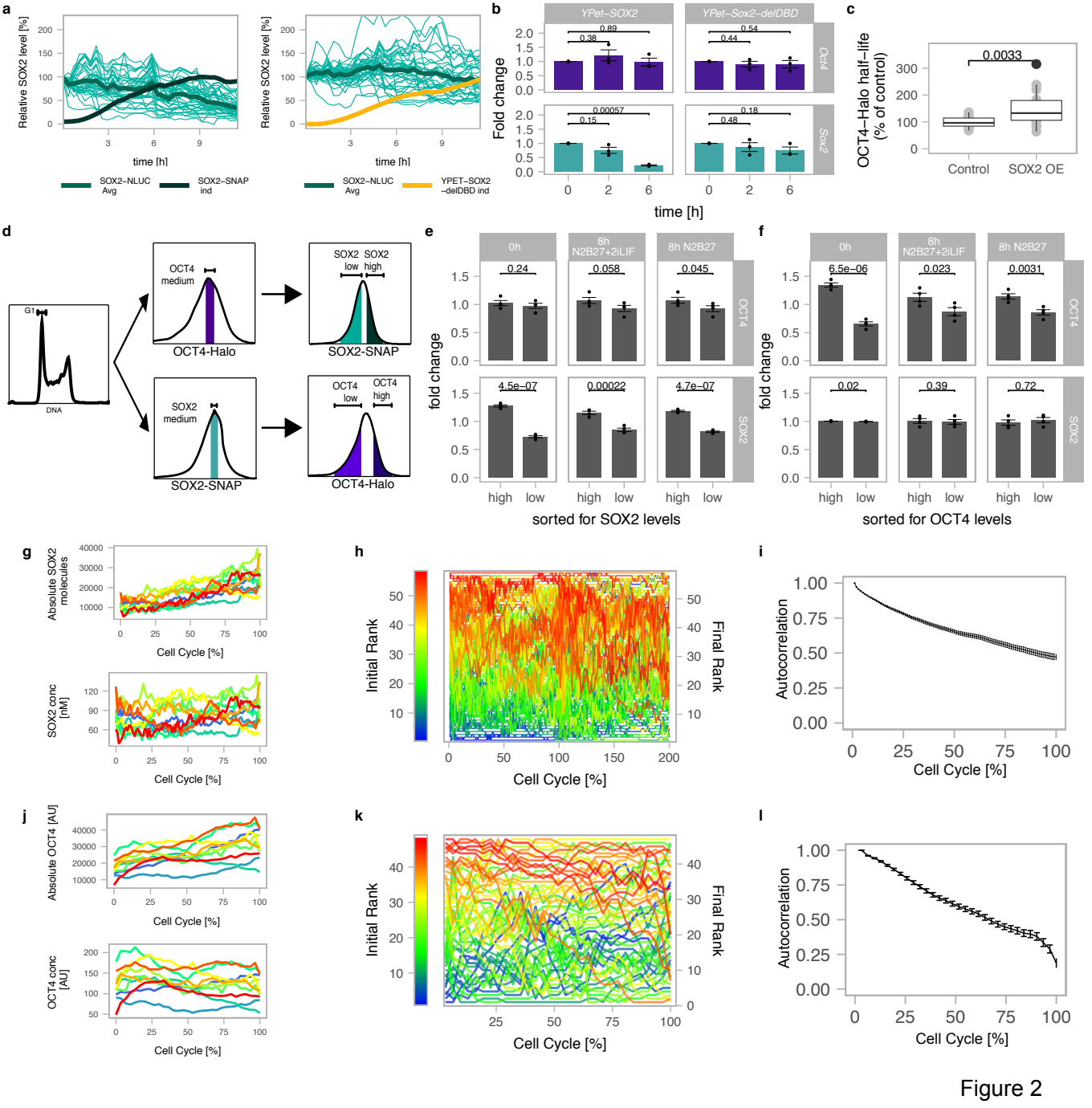
a
Relative SOX2 level [%]
time [h]
SOX2–NLUC Avg
SOX2–SNAP ind
YPET–SOX2 –delDBD ind
b
YPet–SOX2
YPet–Sox2–delDBD
Oct4
Sox2
Fold change
time [h]
c
OCT4–Halo half–life (% of control)
0.0033
Control
SOX2 OE
d
G1
DNA
OCT4 medium
OCT4-Halo
SOX2 medium
SOX2-SNAP
SOX2 low
SOX2 high
OCT4 low
OCT4 high
e
0h
8h N2B27+2iLIF
8h N2B27
fold change
sorted for SOX2 levels
f
sorted for OCT4 levels
g
Absolute SOX2 molecules
SOX2 conc [nM]
Cell Cycle [%]
h
Initial Rank
Final Rank
i
Autocorrelation
j
Absolute OCT4 [AU]
OCT4 conc [AU]
k
l

Figure 2

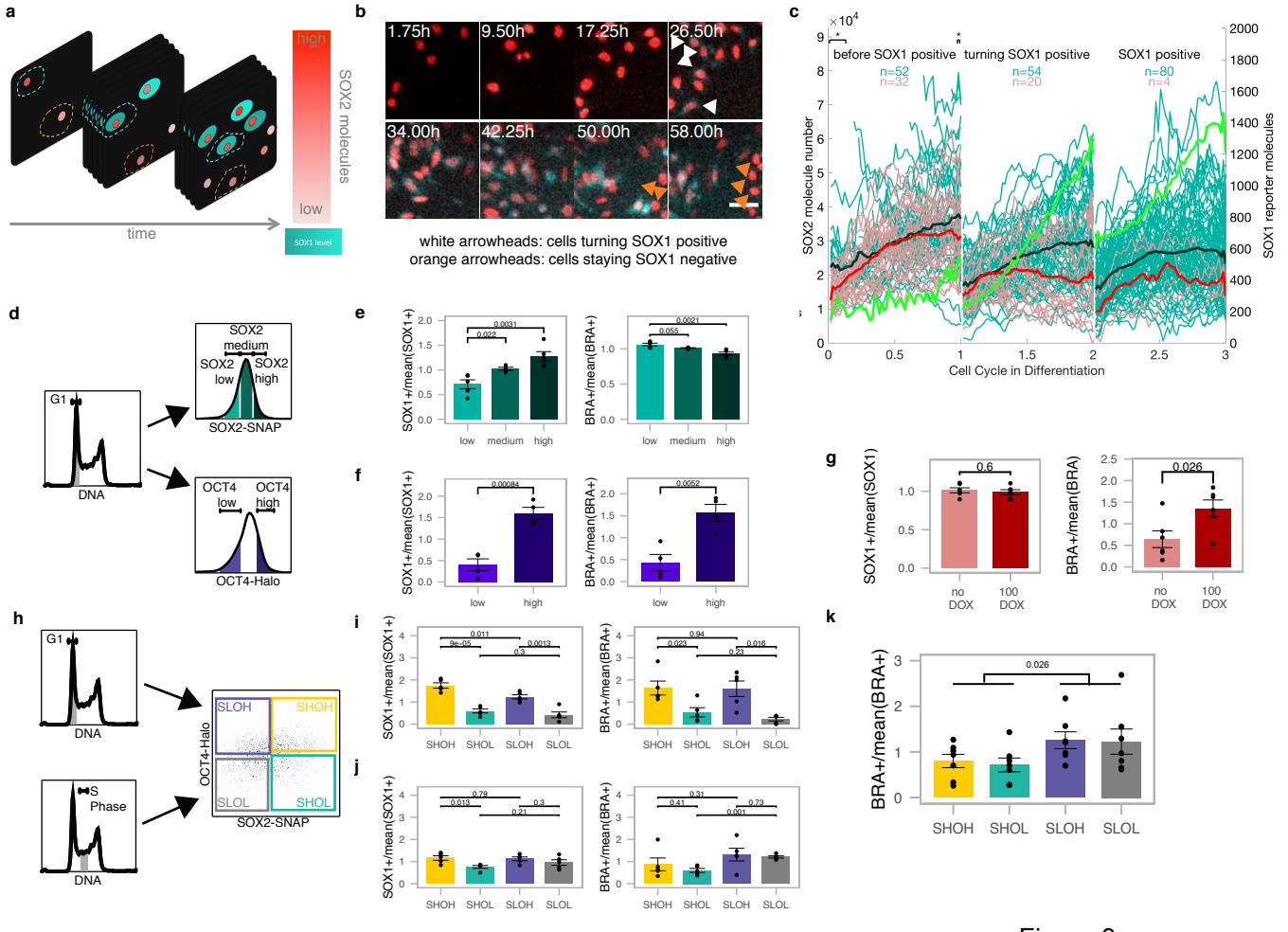
a
high
low
SOX2 molecules
time
SOX1 level
b
1.75h
9.50h
17.25h
26.50h
34.00h
42.25h
50.00h
58.00h
white arrowheads: cells turning SOX1 positive
orange arrowheads: cells staying SOX1 negative
c
before SOX1 positive
turning SOX1 positive
SOX1 positive
n=52
n=32
n=54
n=20
n=80
n=4
SOX2 molecule number
SOX1 reporter molecules
Cell Cycle in Differentiation
d
G1
DNA
SOX2 medium
SOX2 low
SOX2 high
SOX2-SNAP
OCT4 low
OCT4 high
OCT4-Halo
e
SOX1+/mean(SOX1+)
BRA+/mean(BRA+)
low
medium
high
f
low
high
g
SOX1+/mean(SOX1)
BRA+/mean(BRA)
no DOX
100 DOX
h
G1
DNA
S Phase
DNA
SLOH
SHOH
SLOL
SHOL
OCT4-Halo
SOX2-SNAP
i
SHOH
SHOL
SLOH
SLOL
j
k
BRA+/mean(BRA+)

Figure 3

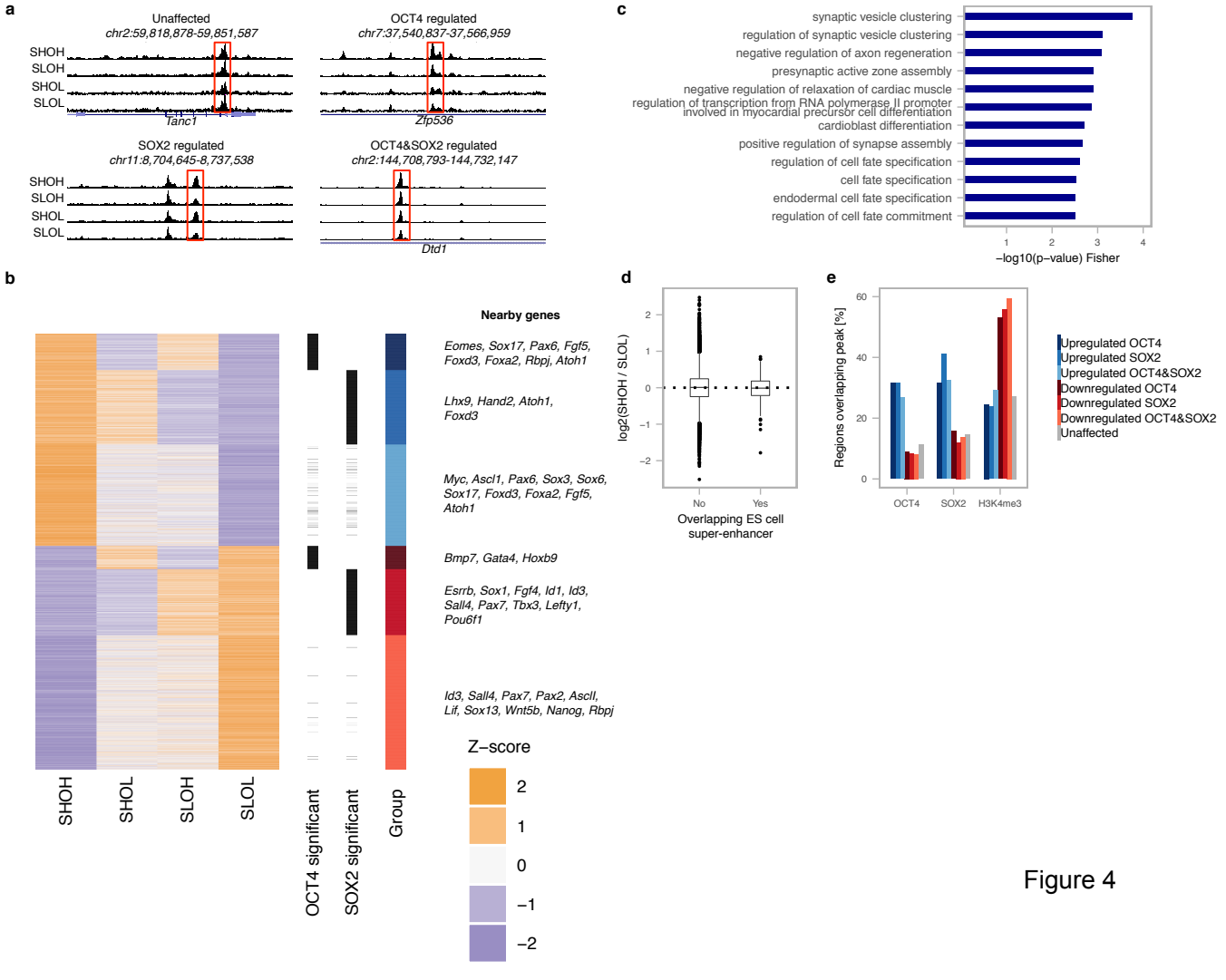
a
Unaffected
chr2:59,818,878-59,851,587
SHOH
SLOH
SHOL
SLOL
Tanc1
OCT4 regulated
chr7:37,540,837-37,566,959
Zfp536
SOX2 regulated
chr11:8,704,645-8,737,538
Dtd1
OCT4&SOX2 regulated
chr2:144,708,793-144,732,147
b
Nearby genes
Eomes, Sox17, Pax6, Fgf5, Foxd3, Foxa2, Rbpj, Atoh1
Lhx9, Hand2, Atoh1, Foxd3
Myc, Ascl1, Pax6, Sox3, Sox6, Sox17, Foxd3, Foxa2, Fgf5, Atoh1
Bmp7, Gata4, Hoxb9
Esrrb, Sox1, Fgf4, Id1, Id3, Sall4, Pax7, Tbx3, Lefty1, Pou6f1
Id3, Sall4, Pax7, Pax2, Ascll, Lif, Sox13, Wnt5b, Nanog, Rbpj
SHOH
SHOL
SLOH
SLOL
OCT4 significant
SOX2 significant
Group
Z−score
2
1
0
−1
−2
c
synaptic vesicle clustering
regulation of synaptic vesicle clustering
negative regulation of axon regeneration
presynaptic active zone assembly
negative regulation of relaxation of cardiac muscle
regulation of transcription from RNA polymerase II promoter involved in myocardial precursor cell differentiation
cardioblast differentiation
positive regulation of synapse assembly
regulation of cell fate specification
cell fate specification
endodermal cell fate specification
regulation of cell fate commitment
−log10(p−value) Fisher
d
log2(SHOH / SLOL)
No
Yes
Overlapping ES cell super-enhancer
e
Regions overlapping peak [%]
OCT4
SOX2
H3K4me3
Upregulated OCT4
Upregulated SOX2
Upregulated OCT4&SOX2
Downregulated OCT4
Downregulated SOX2
Downregulated OCT4&SOX2
Unaffected

Figure 4