Supplemental Material

A role for mitotic bookmarking of SOX2 in pluripotency and differentiation

Cédric Deluz, Elias T. Friman, Daniel Strebinger, Alexander Benke, Mahé Raccaud, Andrea Callegari, Suliana Manley and David M. Suter

Supplemental Methods

Supplemental Movie Legends 1-8

Supplemental Tables 1-4

Supplemental References

Supplemental Methods

Cell culture. The E14 cell line (kindly provided by Didier Trono, EPFL) was used for all ES cell experiments, except for the Luc-Oct4 knock-in and SBR cell and sub-cell lines that were generated from CGR8 ES cells (Sigma-Aldrich, Cat#07032901-1VL), and the 2TS22C cell line (obtained from Riken Bioresource Center, Ibaraki, 305-0074 Japan) that was used to generate corresponding sub-cell lines. Cells were cultured on dishes coated with 0.1% gelatin type B (Sigma), in GMEM (Sigma) supplemented with 10% ES-cell qualified fetal bovine serum, nonessential amino acids, L-glutamine, sodium pyruvate, 100μM 2-mercaptoethanol, penicillin and streptomycin, leukemia inhibitory factor, CHIR99021 at 3μM and PD184352 at 0.8μM. NIH-3T3 cells (kindly provided by Félix Naef, EPFL), and HEK 293T cells (ATCC) were cultured in DMEM supplemented with 10% fetal bovine serum, penicillin and streptomycin. Oct4-GFP MEFs (Lengner et al. 2007) (kindly provided by Matthias Lütolf, EPFL) were cultured in DMEM supplemented with 10% fetal bovine serum, penicillin and streptomycin. For reprogramming experiments, MEFs were plated on matrigel coated 6-well plates at $2*10^5$ cells per well 24 hours before transduction. Five days after transduction, cells from each well were trypsinized and replated on a 10cm cell culture dish coated with Matrigel.

DNA constructs. The Cre-recombinase (Cre) expression lentiviral vector was generated by replacing the CMV promoter and rtTA3G from pLenti CMV rtTA3 Hygro (Addgene plasmid #26730) with a PGK promoter upstream of the sequence encoding Cre recombinase. The plasmid encoding H2B-mCherry was obtained from Addgene (plasmid #21217). The vector for dox-inducible expression of each of the 16 pluripotency factors was generated as follow: A YPet sequence was PCR-amplified with primers encoding a 5' homology arm, a NdeI restriction site, a HA tag followed by two STOP codons, a LoxP site, a YPet fluorescent reporter and a 3'homology arm. This construct was inserted by infusion cloning (clontech) downstream of a gateway recombination cassette (GW). Subsequently, top and bottom oligonucleotides encoding loxP site and 2 more HA-tag sequences were annealed and ligated into the NdeI site. Finally, the cassette encoding GW- loxP-3xHA-2xSTOP-LoxP-Ypet was excised using restriction enzymes and transferred downstream of the TRE3G promoter in the pLentiCMVTRE3G Puro DEST lentiviral vector (Addgene plasmid # 27565). Pluripotency TFs were shuttled from a pENTR library generated in Deplancke lab, EPFL (Gubelmann et al. 2014) into our final inducible vector by LR-Gateway recombination. The SNAP-tagged transcription factors lentiviral constructs (Fig.S1C) were generated by digesting doxycycline-inducible pLVTRE3G-Sox2-YPet vector with AscI and EcoRV to remove YPet, which was replaced by the PCR-amplified SNAP-tag sequence with corresponding AscI and EcoRV overhangs. Subsequently, the Sox2 sequence was removed using the SalI and AscI restriction sites and then replaced by the PCR amplified coding sequences of Esrrb, Klf4 or Nanog.

The H2B-CerFP construct was purchased from Addgene (plasmid#51006). The plasmid expressing rtTA3G was constructed by replacing tTR-KRAB and DsRed from pLV tTR-KRAB-Red (kindly provided by Didier Trono, EPFL), with rtTA3G that was PCR-amplified from Addgene plasmid # 26730, upstream of an internal ribosomal entry site and a blasticidin resistance. To construct the doxycycline-inducible expression lentiviral vectors, we first removed the Gateway cassette from pLenti CMVTRE3G Puro DEST (Addgene, plasmid#27565) and replaced it with a multicloning site by oligo annealing. We then inserted the different Halo-Tag, YPet, and TagRFP-T fusion proteins by restriction cloning. To generate the YPet-Sox2delHMG construct, 5' and 3' parts of the mouse Sox2 coding sequence (omitting amino acids 64 to 106 encoding the HMG DNA-binding domain) were PCR-amplified with primers flanked by restriction sites, ligated together, and cloned downstream and in frame to YPet. The following primers were used: Sox2delHMG5'XbaIF: 5'-CACTCTAGAATGTATAACATGATGGAGACGGAG-3'; Sox2delHMG5'EcoRI: 5'-CAACTGAATTCCTGGGCCATCTTACGCC-3'; Sox2delHMG3'EcoRIF: 5'-CCAGAGAATTCCACCCGGATTATAAATACCGG-3'; Sox2delHMG3'AscIR: 5'-ACAGGCGCGCCTCACATGTGCGACAGGGG-3'. To generate the YPet-HMG construct, the sequence of the Sox2 HMG domain was PCR-amplified with primers flanked by restriction sites, ligated together, and cloned downstream and in frame to YPet. We also generated a truncated YPet-HMG lacking the NLS by skipping the first three amino acids of the HMG domain. The following primers were used: HMGXbaIF: 5'-ACATCTAGAGTCAAGAGGCCCATGAACG-3'; HMGAscIR: 5'-CACGGCGCGCCTTATTTATAATCCGGGTGCTCCT-3'; HMGNoNLSXbaIF: 5'-ACATCTAGACCCATGAACGCCTTCATG-3'. To generate the YPet-Oct4delPou$_S$ construct, 5' and 3' parts of the mouse Oct4 coding sequence were PCR-amplified (omitting amino acids 131 to 205 encoding the Pou$_S$ DNA binding domain) with primers flanked by restriction sites, ligated together, and cloned downstream and in frame to YPet. The following primers were used: Oct4delPH5'XbaIF: 5'-AAGTCTAGAATGGCTGGACACCTGGCT-3'; Oct4delP5'EcoRIR: 5'-GAGCAGAATTCCTGGGACTCCTCGGGAGT-3'; Oct4delP3'EcoRIF: 5'-AGACAGAATTCAACAATGAGAACCTTCAGGAGATATG-3'; Oct4delPH3'AscIR: 5'-CCAGGCGCGCCTCAGTTTGAATGCATGGGAG-3'. To generate the YPet-Oct4delPou$_H$ construct, the 5' and 3' parts of mouse Oct4 coding sequence were PCR-amplified (omitting amino acids 228 to 262 encoding the Pou$_H$ DNA binding domain) with primers flanked by restriction sites, ligated together, and cloned

downstream and in frame to YPet. We also generated a YPet-Oct4delPou$_H$ construct with an additional NLS by insertion of 15 nucleotides encoding the amino acid sequence RKRKR in place of the Pou$_S$ domain sequence. The following primers were used: Oct4delPH5'XbaIF: 5'-AAGTCTAGAATGGCTGGACACCTGGCT-3'; Oct4delH5'EcoRIR: 5'-ACATTGAATTCTCGCTTTCTCTTCCGGG-3'; Oct4delH3'EcoRIF: 5'-ACATCGAATTCATTGAGTATTCCCAACGAGAAGAG-3'; Oct4delPH3'AscIR: 5'-CCAGGCGCGCCTCAGTTTGAATGCATGGGAG-3'; Oct4delH3'EcoRIF+NLS: 5'-ACATCGAATTCCGGAAGAGAAAGCGAATTGAGTATTCCCAACGAGAAGAG-3'. To generate the YPet-Pou$_S$ construct, the Pou$_S$ domain was PCR-amplified with primers flanked by restriction sites, ligated together, and cloned downstream and in frame to YPet. The following primers were used: Pou$_S$XbaIF: 5'-GGATCTAGAGACATGAAAGCCCTGCAGA-3'; Pou$_S$AscIR: 5'-ACAGGCGCGCCTTAGTCGGCTTCCTCCACC-3'.
The lentiviral vectors encoding SOX2 fusions to YPet or to a mitotic degron (MD) or point mutant thereof (MD*) were generated by PCR amplification with primers flanked by restriction sites, ligated together, and cloned downstream of either TRE3G promoter using pLV-TRE3G-MCS, or downstream of the PGK promoter of pLV tTR-KRAB-Red after excision of tTR-KRAB and DsRed. The MD coding sequence was synthesized by GeneArt and encodes an amino acid sequence identical to the one reported in (Kadauke et al. 2012). The MD* coding sequence was generated by site-directed mutagenesis on the MD sequence to mutate the arginine at position 42 into an alanine (R42A), allowing to inactivate the MD as previously reported (Kadauke et al. 2012).
The lentiviral vectors allowing expression of Oct4, Klf4, cMyc and either Sox2-YPet-MD, Sox2-YPet-MD* were constructed by excising Sox2 from pFuW-TetO-OSKM (addgene plasmid # 20321), and ligating PCR-amplified Sox2-YPet-MD and Sox2-YPet-MD*.
All constructs were verified by Sanger sequencing.

Lentiviral vector production and generation of stable cell lines. Lentiviral vector production was carried out by calcium phosphate transfection of HEK 293T cells with the envelope (PAX2) and packaging (MD2G) constructs together with the lentiviral vector of interest, and concentrated 120-fold by ultracentrifugation as described previously (Suter et al. 2006), except for the intial screen of 16 pluripotency TFs, for which lentiviral particles were generated using X-tremeGENE 9 DNA Transfection Reagent (Roche). Target E14 cells, NIH-3T3 cells or HEK 293T cells were then seeded at 50,000 cells per well of a 24-well plate (E14 cells) or of a 6-well plate (NIH-3T3 and HEK 293T cells) and transduced with 50µl of concentrated lentiviral vector. Selection of transduced cells was performed by addition of the respective antibiotics 48-72 hours after transduction, and the antibiotics were maintained in the cell culture medium throughout passaging. For blasticidin selection, we used 8µg/ml (E14 cells) or 5µg/ml (NIH-3T3 and HEK 293T cells); for puromycin selection, we used 2µg/ml (all cell lines). For the initial screen for mitotic chromosome binding of 16 pluripotency factors, we first generated a cell line constitutively expressing H2B-mCherry and Cre recombinase. This cell line was then used to generate 16 sub-cell lines allowing dox-inducible expression of each of 16 transcription factors, by transducing 4000 cells in 96-well plates twice (on two consecutive days) with non-concentrated lentiviral vector particles. Constitutive expression of Cre allowed excising a cassette containing 3 HA repeats and each transcription factor coding sequence to be in frame with a C-terminal YPet tag (Supplemental Methods). Selection with puromycin was started 48hrs after the first transduction.

Confocal microscopy. The acquisition of confocal images (Fig.S1) was performed on a Leica TCS SP8 confocal laser scanning microscope, with a 63x oil-immersion objective, using a 405nm laser to excite CerFP and a 514nm laser to excite YPet, and the pinhole diameter was set to obtain slices < 1µm thick in both channels.

Live luminescence microscopy. 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% CO_2) for 24-48 hours. One day before the experiment, cells were seeded on fluorodishes (WPI, FD35-100) coated with E-Cadherin as described (Nagaoka et al. 2006). The next day, the medium was exchanged and supplemented with 1mM luciferin (NanoLight Technology, Cat#306A). Images were acquired every 90s with a 30s exposure time.

Single Molecule Microscopy. Single molecule imaging was performed on an inverted TIRF microscope (Axio Observer.D1, Zeiss) with a Plan-Apochromat, 100x, numerical aperture =1.46, oil objective (Zeiss). The microscope was equipped with a piezo objective scanner (N-725, Physik Instrumente), a motorized stage (MFC-2000, Applied Scientific Instrumentation) and with a live-cell environmental chamber providing constant 37°C temperature and 5% CO2 concentration (Live Imaging Services). Samples were excited using a 561 nm laser (CrystaLaser) with irradiance of 1 kW/cm^2. To decrease out of focus fluorescence we performed imaging using highly inclined illumination (HILO) conditions (Tokunaga et al. 2008). Fluorescent light was collected by the objective and directed through a dichroic mirror (89100BS, Chroma) and an emission filter

(ET630/75m, Chroma) to an EMCCD camera (iXon 897+, Andor) with 160nm pixel size. Samples were incubated one hour before imaging with 5nM Halo-TMR substrate (Promega) for 30 minutes, washed and incubated again without substrate, and then the medium was changed to DMEM without phenol red to reduce background fluorescence, supplemented with 50ng/ml of doxycycline. To allow accurate measurements of long DNA residence times, we used a time-lapse illumination strategy that we described before (Gebhardt et al. 2013). Briefly, this consists of directing the excitation laser light through an acousto-optic tunable filter (AOTF) device (AA Optoelectronics) acting as a shutter controlled by the LabView software (National Instruments), to introduce a variable dark time, or gap between images. We typically acquired between 200-1000 image frames (2s exposure time) per cell for a given gap duration, with gaps between frames ranging between 0-8s (Gebhardt et al. 2013).

Number of binding events/total fluorescence analysis. The bound fraction (BF) – that is, the ratio of bound versus free TF molecules – contains valuable biological information about the association and dissociation rates of the TF to chromatin. In fact, $BF \stackrel{\text{def}}{=} k_{on}^{*}/(k_{on}^{*} + k_{off})$ where k_{on}^{*} and k_{off} are the pseudo-association and dissociation rates, respectively (Mueller et al. 2013). The pseudo on-rate k_{on}^{*} is the only experimentally accessible parameter linked to the actual on-rate k_{on} by the expression $k_{on}^{*} = k_{on}[S_{eq}]$, where $[S_{eq}]$ represents the concentration of free binding sites at the thermodynamic equilibrium. Since $K_D^{*} = k_{off}/k_{on}^{*}$ is the pseudo-dissociation equilibrium constant of the TF, then $BF = 1/(1 + K_D^{*})$. Also, for a given binding reaction where the TF binds DNA, at the equilibrium K_D^{*} = [TF]·[DNA] / [TF-DNA], where brackets refer to the concentration of each species. In particular, [TF] / [TF-DNA] is the ratio of unbound over specifically-bound TF molecules. Assuming that [TF-DNA] is negligible as compared to the total amount of molecules [TF] that are freely diffusing and non-specifically bound, we have [TF] ≫ [TF-DNA]. Under this assumption, K_D^{*} will scale inversely with the number of binding events n_i divided by the total fluorescence F_i of a given image i and $BF \propto \frac{n_i}{F_i}$. We then estimated BF by computing the ratio $\frac{n_i}{F_i}$ for a given image. Binding events (n_i) were localized running the Fiji plugin Octane and counted by visual inspection per each considered frame. The peak intensity threshold was set individually for each sample to account for variations in the image background level. Total nuclear fluorescence F_i was calculated by manual thresholding of the nucleus (interphase cells) or the whole cell (mitotic cells) using the Fiji software and corrected for the background noise, B_i. Irradiance and exposure time were kept constant for all samples as well as bleaching conditions to achieve single-molecule regime (~1 minute at maximum laser power). For each sample, we estimated the bound fraction (BF) by calculating the ratio of the number of individual binding molecules to total nuclear fluorescence in arbitrary units (a.u.). Ratios are averaged over 10 consecutive time points (frames 110-119; 2 minutes and 40 s after imaging started), so that the estimated bound fraction was given by:

$$BF = \frac{1}{10} \cdot \sum_{i=1}^{10} \frac{n_i}{F_i - B_i}$$

Ratiometric BFs were calculated out of 20-21 different interphase or mitotic cells either expressing Sox2 or Oct4. For comparing distributions of BF values, 25th-75th percentile box-plots were used. To appreciate additional information about the spread of the data, whiskers were drawn from the upper and lower hinges of the box to the maximum and minimum values. Differences in data distributions between mitotic and interphase cells were scored by calculating the P-value through a Mann-Whitney non-parametric statistical test, assuming that distributions did not differ each other as null-hypothesis.

Mitotic cell sorting. E14 ES cells were synchronized using 200 ng/ml Nocodazole for 12 hours, trypsinized, fixed with 1% formaldehyde for 10 minutes and quenched with 200 mM Tris-HCl pH 8.0. Cells were permeabilized with cold 70% EtOH for 15 minutes, stained with 1:1000 anti-H3S10p antibody CMA312 (EMD Millipore) for 1 hour at 4°C, and subsequently stained with 1:80 anti-mouse IgG1 PE antibody (eBioscience) for 1 hour at 4°C. PE-positive cells were sorted, spun down, and stored at -80°C. For the DAPI control staining, sorted cells were stained with 1mg/ml DAPI for 15 minutes, imaged on a Zeiss Axioplan, and cells with condensed chromatin were counted.

ChIP-seq experiments. Asynchronous E14 ES cells ($11x10^6$, $13x10^6$, and $14x10^6$ cells per replicate) were fixed with 1% formaldehyde for 10 minutes at room temperature and quenched with 200 mM Tris-HCl pH 8.0, washed with PBS, spun down, and stored at -80°C. The cell pellet was resuspended and 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% Triton X-100), incubated 10 min at 4°C, spun down, and resuspended in 1.5 ml LB2 (10 mM Tris-HCl pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA) incubated 10 min at 4°C. The pellet was spun down, rinsed twice with SDS shearing buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA, 0.15% SDS), and finally resuspended in 0.9 ml SDS shearing buffer. All buffers contain contain 1:100 diluted Protease Inhibitor Cocktail in DMSO (Sigma). Mitotically sorted cells were directly resuspended in SDS shearing buffer. 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, 140W, and input sample aliquots were taken. To pull down

SOX2-bound DNA, sonicated chromatin was incubated with 500 mg anti-SOX2 antibody Y-17 (Santa Cruz) per $1x10^6$ cells overnight at 4°C. 2.5 ml Protein G Dynabeads (Thermo Fischer) per $1x10^6$ cells were added to the chromatin and incubated 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 finally with TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA, 1 mM PMSF). Beads were finally 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 1100 rpm overnight. Input samples were treated with 100 mg/ml RNase A and 400 ng/ml Proteinase K and reverse crosslinked at 65°C 1100 rpm overnight. Sonicated chromatin samples were purified using Qiagen MinElute PCR purification kit, and sequencing libraries were prepared using NEBNext ChIP-seq Library Prep Master Mix Set (NEB), either without size selection (replicate one) or with size selection to obtain average size chromatin fragments of ~400 bp (replicates two and three). Libraries were sequenced using 100 bp single end reads on an Illumina HiSeq 2500.

Western Blotting after Sox2 ChIP

$9*10^6$ asynchronous ES cells were fixed and $8.5*10^6$ mitotic E14 cells were fixed, stained, and sorted as described in "ChIP-seq experiments". Cells were lysed and sonicated, and SOX2 was pulled down as described in "ChIP-seq experiments" using 25µg antibody. The negative control sample contained SDS shearing buffer and 25µg anti-SOX2 antibody and was treated as samples during SOX2 IP. After washing with TE buffer, proteins were eluted from the Protein G beads by addition of 1X Laemmli buffer (62.5 mM Tris-HCl, pH 6.8, 10% glycerol, 1% LDS, 0.0045% Bromophenol blue, 10% beta-mercaptoethanol; BioRad) and incubation at 95°C for 5 min. 1/10 of each sample was loaded on a 4-20% Tris-glycine gel (BioRad) together with a PageRuler Plus protein ladder (Thermo). After gel separation, proteins were transferred to a MeOH-activated PVDF membrane (Immobilon) using semi-dry transfer. The membrane was blocked with 5% milk in PBS with 0.05% Tween-20 (Milk/PBS-T) for 30 minutes at RT, followed by incubation with Rabbit anti-SOX2 antibody (Cat. 48-14800; Thermo) in Milk/PBS-T at 1:250 overnight at 4°C. The next day, the membrane was washed once with PBS-T and then incubated with anti-Rabbit HRP Conjugate (Promega) at 1:10'000 for 45 minutes at room temperature. The membrane was washed five times in PBS-T, developed with Clarity Western ECL Substrate (BioRad), and detected on a Witec Fusion imaging system.

Generation of Sox2-Luc and Luc-Oct4 knock-in cell lines. The E14 Sox2-Luc and CGR8 Luc-Oct4 cell lines were generated using CRISPR-Cas9 mediated homology directed repair (HDR), as previously described (Cong et al. 2013). The repair templates for HDR were designed to contain a knock-in cassette flanked by 1kb homology arms with the target sequence. For the Luc-Oct4 knock-in, the repair template construct contained the 5'UTR of Oct4 upstream of a cassette encoding a blasticidin-TagRFP-T fusion protein followed by a F2A peptide allowing to separate the polypeptide chain encoding blasticidin-TagRFP-T from the downstream sequence by ribosomal skipping. Firefly luciferase was inserted downstream and in frame to this cassette and fused to the 5' part of the Oct4 coding sequence. For the Sox2-Luc knock-in, the repair template construct consisted of the 3' part of the Sox2 coding sequence fused to firefly luciferase, upstream of a cassette encoding a P2A peptide allowing ribosomal skipping followed by a fusion between blasticidin resistance and sfGFP. The guide RNAs were designed by using the Zhang Lab toolbox (www.genome-engineering.org/crispr) and identified guides were subsequently cloned into a pX330 backbone (Cong et al. 2013) to target the Start codon of Oct4 (guide sequence: 5'-GCA GGT GTC CAG CCA TGG GGA-3') or the Stop Codon of Sox2 (guide sequence: 5'-GCA GCC CTC ACA TGT GCG ACA-3'), respectively. Next, $2.4*10^6$ ES cells were co-transfected with pX330 and the corresponding repair template in a 1:3 ratio (20µg total DNA), grown for two days before selection with blasticidin (10µg/ml) was started. 10 days later, colonies were picked manually and allowed to grow out. Genomic DNA was extracted from the candidate cell lines using a Genomic DNA extraction kit (Sigma). The knock-ins were verified by PCR of the target loci, using primers listed in Table S1.

Generation of the Sox1/eGFP Bra/mCherry (SBR) knock-in reporter cell line. Knock-in to express fluorescent proteins (FPs) from the Sox1 and Bra loci was performed with homology directed repair (HDR) as described previously (Cong et al. 2013). The HDR templates were designed to contain a knock-in cassette flanked by homology arms (HAs) spanning 1 kb surrounding the stop codons of Sox1 and Bra without the stop codon. The HDR templates were designed as: 5'HA-P2A-FP-STOP-loxP-selection cassette-loxP-3'HA. For Sox1, the FP was eGFP and the selection cassette PGK-Hygromycin resistance. For Bra, the FP was mCherry and the selection cassette PGK-Puromycin resistance–sfGFP to enable monitoring of loss of sfGFP expression following Cre recombination. Guide RNAs targeting the Sox1 and Bra loci were designed using the Zhang Lab toolbox (www.genome-engineering.org/crispr) and cloned into the pX335 vector expressing Cas9D10A (nickase) and the guide RNA. For pX335-Sox1, the guide RNA binds in a region spanning the STOP codon, which is removed in the HDR template (guide sequence: 5'-GAC GCA CAT CTA GCG CCG CG). For

pX335-Bra, the guide RNA binds after the STOP codon (guide sequence: 5'-GTG CTG AGA CTT GTA ACA AC-3'). To avoid cuts of the HDR template by Cas9, one nucleotide mismatch was incorporated in the untranslated 3'HA.
CGR8 mouse ES cells (Sigma Aldrich) were transfected with pX335-Sox1 and the Sox1 HDR template at a 1:3 ratio. After two days, selection was started with 200 ug/ml Hygromycin B. After 14 days of selection, the pool of cells was transfected with pX335-Bra and the Bra HDR template at a 1:3 ratio. After 2 days, selection was started with 2 ug/ml Puromycin. After 13 days of selection, cells were sorted for sfGFP expression and transfected with pLVEF1a-Cre constitutively expressing Cre recombinase. After 4 days, single sfGFP negative cells were sorted into wells of 96-well plates. Single colony outgrowths were split into four conditions, without selection, selection with either Hygromycin or Puromycin, and into N2B27 medium for differentiation. A clone that was sensitive to both antibiotics and expressed eGFP and mCherry upon differentiation was picked and verified using PCR of DNA extracted using a Genomic DNA extraction kit (Sigma) using the primers listed in Table S2.

Reverse transcription followed by Quantitative PCR (RT-QPCR). Total RNA was extracted using a GenElute™ Mammalian Total RNA Miniprep Kit Q-PCR (Sigma-Aldrich), and reverse transcription was performed using an oligoDT primer using superscript II (Life Technologies). QPCR was performed on a 7900HT Fast Real-Time PCR System (Thermofischer) with SYBR green reagent (Roche). Rps9 cDNA was used for data normalization. Primers used for RT-QPCR are listed in table S3.

ChIP-QPCR. ChIP-QPCR was performed on library preparations of ChIP and input samples, using the primers listed in Table S4, on a 7900HT Fast Real-Time PCR System (Thermofischer) with SYBR green reagent (Roche).

Sorting of YPet-tagged and SNAP-tagged Sox2 cell lines. We developed two different strategies to sort for similar expression levels between MD and MD* cell lines. In the first strategy, we aimed at obtaining similar expression levels for the MD and MD* cell lines once expression levels have recovered from mitotic degradation. For convenient testing of this strategy, we used the dox-inducible cell lines, which allow expression levels that are high enough to be reliably measured by live cell imaging. We observed that sorting for the same, large fluorescence window (Fig.S10a-b) of the TRE3G-Sox-YPet-MD and TRE3G-Sox2-YPet-MD* cell line, allowed obtaining cell populations which the median fluorescence intensity was slightly different between MD and MD* cell lines (about 1.4 fold higher for the TRE3G-Sox2-YPet-MD* as compared to the TRE3G-Sox2-YPet-MD cell line, Fig.S10a-b). We then measured YPet fluorescence intensity from time-lapse series in single cells, 5 frames before observable chromatin condensation, and observed very similar average intensities (Fig.S10i). Thus, even though these two cell lines have different median fluorescence intensities over the whole cell cycle, they have similar expression levels once the effect of mitotic degradation has disappeared. Conversely, we reasoned that sorting for a much narrower cell population should yield similar median fluorescence intensities, which should then be at the expense of slightly higher expression levels in the MD cell lines in cell cycle stages at which expression levels have recovered from mitotic degradation. We thus sorted a much narrower fluorescence window of the TRE3G-SNAP-MD-Sox2 and TRE3G-SNAP-MD*-Sox2 cell lines, which resulted in virtually identical median fluorescence intensities (Fig.S10c-d). We then performed measurements of SNAP-SiR647 fluorescence intensities in cells 5 frames before observable chromatin condensation, and observed about 1.4-fold higher fluorescence intensities in the TRE3G-SNAP-MD-Sox2 than the TRE3G-SNAP-MD*-Sox2 cell line (Fig.S10j). In that case, the two cell lines have similar median fluorescence intensities, but the SNAP-MD fluorescence intensity is higher when the effect of the MD has disappeared, thus compensating for the lower levels at the end of the M phase and in early G1. For sorting of the 2TS22C cell lines expressing the four different tagged versions of Sox2 under the control of the PGK promoter, we used the "narrow window strategy" for the Sox2-YPet fusion (Fig.S10e-f) and the "large window strategy" for the SNAP-Sox2 fusions (Fig.S10g-h).

Imaging of dox-inducible SNAP-tag cell lines. Time-lapse fluorescence recordings were performed on an inverted Olympus IX81 motorized microscope or an InCell Analyzer 2200. Cells were pre-incubated overnight in the presence of dox, and then the medium was exchanged to fluorobrite DMEM supplemented with 10% ES-cell qualified fetal bovine serum, nonessential amino acids, L-glutamine, sodium pyruvate, 100μM 2-mercaptoethanol, penicillin and streptomycin, leukemia inhibitory factor, CHIR99021 at 3μM and PD184352 at 0.8μM and 0.5μg/ml of dox. For imaging of the Sox2-SNAP-MD and Sox2-SNAP-MD* ES cell lines, the SNAP SiR dye was added at a final concentration of 3nM and left throughout time-lapse imaging. For imaging of the Nanog-SNAP, Klf4-SNAP and Esrrb-SNAP, cells were incubated with the SNAP SiR dye at a final concentration of 0.6nM for 30 minutes and Hoechst at 50μM for 15 minutes, washed twice in complete fluorobrite medium, and imaged.

Supplemental Movie Legends

All movies were performed at a time resolution of 5 minutes.

MovieS1. Time-lapse imaging of a dividing ES cell expressing YPetSox2.
MovieS2. Time-lapse imaging of a dividing ES cell expressing YPetOct4.
MovieS3. Time-lapse imaging of dividing Sox2-luciferase knock-in ES cells.
MovieS4. Time-lapse imaging of a dividing Luciferase-Oct4 knock-in ES cell.
MovieS5. Time-lapse imaging of dividing cells expressing Sox2-YPet-MD.
MovieS6. Time-lapse imaging of dividing cells expressing Sox2-YPet-MD*.
MovieS7. Time-lapse imaging of dividing cells expressing SNAP-MD-Sox2.
MovieS8. Time-lapse imaging of dividing cells expressing SNAP-MD*-Sox2.

Supplemental Tables

Table S1

Sequence (5' to 3')	Target sequence position
ACCTCCGGGACATGATCA	In Sox2 5' HA (sense)
CGTTTCGCTGCGGAGATT	In Sox2 3' HA (antisense)
CGGTAAGACACTGGGTGTGA	In Firefly Luciferase (sense)
GCATGCTAGCCACAAAGAAA	Downstream of Sox2 3' HA (antisense)
GAGGATGAACACCGGAGTCC	Upstream of Oct4 5' HA (sense)
CTAATTGCTTACACTTGCTCC AGA	Downstream of Oct4 3'HA (antisense)
CTTGTACAGCTCGTCCATGC	In sfGFP (antisense)

Primers used to verify genomic integration of the Sox2-Fluc and Fluc-Oct4 knock-in cassettes. HA: homology arms.

Table S2

Sequence (5' to 3')	Target sequence position
AAAAACGTTGTGAGCAAGGGCGAGGAG	eGFP start
GTGCCCCTGACGCACAT	Inside Sox1 5' HA
AGTATGCAAAAGGCAGAACCAC	Downstream of Sox1 3' HA
CATGGATCCATGGTGAGCAAGGGCGAG	mCherry start
GGGCTGCTACAGTCACAGAC	Inside Bra 5' HA
TGCACATTCTCCAAGAAGGA	Downstream of Bra 3' HA

Primers used to verify genomic integration of the Sox1-P2A-eGFP and Brachyury-P2A-mCherry knock-in cassettes. HA: homology arms.

Table S3

Sequence (5' to 3')	Target sequence
GCTCCTCAGCATGACTACTACT	Dlx3 cDNA (forward)
CTGCGAGCCCATTGAGATTG	Dlx3 cDNA (reverse)
GAAAACGCGGAATCTGAAACAA	Esx1 cDNA (forward)
TCCTGGTAGTTTGCTTCTTCCA	Esx1 cDNA (reverse)
TTGTCGCAAAACCTATGTGACC	Rps9 cDNA (forward)
GCCGCCTTACGGATCTTGG	Rps9 cDNA (reverse)

Primers used for RT-QPCR.

Table S4

Sequence (5’ to 3’)	Target sequence
CAATACCGCCCATTTCCAGG	Rif1_F
TGGCAGCACATTTCTTAGCG	Rif1_R
GGTGGAGAGTGCTGTCTAGG	Oct4_F
CTGCCCAGAACTCTCAACCT	Oct4_R
GCTGAGTCTAATTGAGGGCC	NF-kB1F
ACACAGTACTGACAGGGCTC	NF-kB1R
GGAGGAGGAGAACAAGGTCC	Nanog_F
TGAGAGAACACAGTCCGCAT	Nanog_R
GGCACCACCCAATCGAAAG	Nrarp_F
ACAATAGACGCTCTTGACGC	Nrarp_R
GAAAGGGGCTGCATGACAAA	Cnpy1_F
CGTGTGAAAGGGGACAAAGG	Cnpy1_R
CAATGAGGTTCGTGGTCTGC	Ube2r2_F
GTGAACTTCCAGACCTGTGC	Ube2r2_R
AAAGGCAGAACAAAGGAGCC	Mitotic/Interphase peak1_F
CCCGAGGTAGCTTGAGATCA	Mitotic/Interphase peak1_R
GCAAAAGTGTGGCAATCCTAC	Mitotic/Interphase peak2_F
TGGAAACCCAGCCTGTGTG	Mitotic/Interphase peak2_R
GCTAGTCTTGCCAGGTCCTA	Mitotic peak1_F
TAAAGGAGAGTGGGGTCTGG	Mitotic peak1_R
CTTTCATCGGAGCAGCGAAG	Mitotic peak2_F
GTGAGGGAGGACAGACAGAG	Mitotic peak2_F
CTTTCCATCTCCCGTCCTGT	Mitotic peak3_F
CAGTGGTCCAGAAGTGACGA	Mitotic peak3_R
GCTCGGCTAAGAAAGGGC	Mitotic peak4_F
CCCGCCCAATCAGCTGTA	Mitotic peak4_R
TAGAACACAGGGCTTGGTGT	Mitotic peak5_F
GCTCACACAGGTTCCCATTC	Mitotic peak5_R
ACCAGTCAGTCCAGCTAGAG	Mitotic peak6_F
CTGACTCTGCTGCCTACCAT	Mitotic peak6_R
GTGGTATAGCTGGGTCCTCA	Negative Region1_F
ACACACCAATGGACGAGTGT	Negative Region1_R
CACATGGACCTTTCAAGCTGT	Negative Region2_F
GACAGGAGATGGACCATGGG	Negative Region2_R
TTGAAATCCGTTGGTCCCAA	Negative Region3_F
GTTCCTGTCCTCCTCAGCAT	Negative Region3_R

Primers used for ChIP-QCPR.

Supplemental References

Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA et al. 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**: 819-823.

Gebhardt JC, Suter DM, Roy R, Zhao ZW, Chapman AR, Basu S, Maniatis T, Xie XS. 2013. Single-molecule imaging of transcription factor binding to DNA in live mammalian cells. *Nature methods* **10**: 421-426.

Gubelmann C, Schwalie PC, Raghav SK, Roder E, Delessa T, Kiehlmann E, Waszak SM, Corsinotti A, Udin G, Holcombe W et al. 2014. Identification of the transcription factor ZEB1 as a central component of the adipogenic gene regulatory network. *eLife* **3**: e03346.

Kadauke S, Udugama MI, Pawlicki JM, Achtman JC, Jain DP, Cheng Y, Hardison RC, Blobel GA. 2012. Tissue-specific mitotic bookmarking by hematopoietic transcription factor GATA1. *Cell* **150**: 725-737.

Lengner CJ, Camargo FD, Hochedlinger K, Welstead GG, Zaidi S, Gokhale S, Scholer HR, Tomilin A, Jaenisch R. 2007. Oct4 expression is not required for mouse somatic stem cell self-renewal. *Cell stem cell* **1**: 403-415.

Mueller F, Stasevich TJ, Mazza D, McNally JG. 2013. Quantifying transcription factor kinetics: at work or at play? *Critical reviews in biochemistry and molecular biology* **48**: 492-514.

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.

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

Tokunaga M, Imamoto N, Sakata-Sogawa K. 2008. Highly inclined thin illumination enables clear single-molecule imaging in cells. *Nature methods* **5**: 159-161.