Supplemental Materials and Methods

List of primers used for qPCR.

| Gene | Forward 5→3'           | Reverse 5→3'                               |
|------|------------------------|--------------------------------------------|
| Nanog| TTCTTGCTTACAAGGGTCTGC  | AGAGGAAGGGCGAGGAGA                         |
| Foxd3| GTCGGCTGGGAATAACTTTCGTA| ATGTACAAAGAATGTCCCTCCCCACCC               |
| Rai2 | CGGTCAATTAGTGGAAGTGAG  | GAGGCTGGATTTCGCGTG                        |
| Zfp451| CTGGTGGCACAGCCACACACTT| TCCATCAGAGCTTTAAACGAG                     |
| Satb1| AGTGCCCCCTTTCACAGAG   | TGCTGCTGAGACATTATCAT                     |
| Satb2| ATGAACCCCAATGTGAGCAT  | GTTGTCGCGTGAGGTTTT                        |
| Oct4 | GTTGGAGAGGTGGAACAAA   | CTCCCTTCTCGAGGGCTTTC                      |
| Nestin| GAGCTGGAGCGCGAGTTGA  | GCCACTCCAGACTAGGA                         |
| T    | TACCCAGCCTATGCTCA     | GGCACTCCAGAGCTAGGCA                       |
| Gata6| AACTGCGGCTCCATCCAGAC  | TGGCACAGGACAGTCCA                         |

List of antibodies used for Immunoblot.

| Protein | Source  | Identifier |
|---------|---------|------------|
| Flk-1   | Santa Cruz | sc-6251    |
| Nanog   | Abcam   | ab80892    |
| Nestin  | Santa Cruz | sc-23927  |
| Oct3/4  | Santa Cruz | sc-5279   |
| Pias1   | Abcam   | ab32219    |
| Rex1    | Abcam   | ab50828    |
| Satb1   | Abcam   | ab92307    |
| Satb2   | Abcam   | ab51502    |
| Tubulin | Abcam   | ab7291     |
| Zfp451  | Sigma   | SAB2108741 |
| Sumo1   | Raised in house | 21C7      |
| Sumo2   | Raised in house | 8A2       |
| FLAG    | Sigma   | M2AB       |
| HA      | Biolegend| 637302     |
| Lsd1    | Abcam   | 17721      |
| CoRest  | Abcam   | 32631      |
| Hdac1   | Abcam   | 7028       |
| Hdac2   | Cell signalling | 5113      |
| Mi-2b   | Bethyl  | A301-081A  |
| Ty1     | Invitrogen | MA5-23513 |

List of antibodies used for ChIP.

| Protein | Source   | Cataloge  |
|---------|----------|-----------|
| Ty1     | Invitrogen | MA5-23513 |
| Satb1   | Abcam    | 92307     |
| Lsd1    | Abcam    | 17721     |
| H3K4me2 | Millipore| 07-030    |

List of primers used for 3C assays at the Nanog locus.

| Site               | Primer                                      |
|--------------------|---------------------------------------------|
| SatbBS (a)         | 5’-GGGCACAGGAGCAGAGAC-3’                    |
| enhancer (b)       | 5’-CCAGTTCTCTGGACTCCTCCC-3’                 |
| promoter (c)       | 5’-AATGGAAGAGGAACCTCAGATCC-3’               |
| +12kb element (d)  | 5’-GGGCTAGGAAGACATCTTGTTG-3’                |

Protein extraction ± NEM

Cells were washed twice with ice cold PBS and resuspended in lysis buffer (20 mM Hapes pH 7.6, 0.1% NP-40, 10 mM Sodium fluoride, 2 mM magnesium chloride, 150 mM sodium chloride, 10% glycerol, 1X Protease inhibitor mix (Sigma), 1mM PMSF, 1 mM sodium orthovanadate). The isopeptidase inhibitor N-Ethylene Maleimide (NEM, Sigma) was added to the PBS and lysis buffer to a final concentration of 10mM where indicated. The suspension was sonicated at 4°C and then diluted with 2x Laemmli Sample buffer for western blot analysis.

Immunoprecipitation

Cells grown to ~80% of confluency were washed with PBS and harvested by trypsinization. Every 5x10^7 cells were resuspended in 50ml of PBS + 1mM DSP and incubated for 30mins in rotation at 4°C; then glycine was added to a final concentration of 50mM to quench the fixative during 15 min at 4°C. The cell pellet was washed with cold PBS containing PIM (Sigma). Every 1x10^6 cells were resuspended in 100 µl of IP lysis buffer: 20 mM Hapes pH 7.6, 0.1% NP40, 10 mM NaF, 2 mM MgCl2, 150 mM NaCl, 10% glycerol, 1 mM PMSF, 1x PIM, 1 mM Na3V3O4 and 10 mM NEM. Next, the lysate was sonicated two times with 30sec pulse of output 20 and duty cycle 90 (Bransom sonifier 450). Following centrifugation, the supernatant was recovered and quantified via Bradford assay. The lysate was pre-cleared with mouse Ig-G Agarose beads (Sigma) for 1h at 4°C. For FLAG IP, every 15 µl of anti-Flag M2 affinity gel beads (Sigma) per each 100ug of protein were equilibrated with 30 µl of lysis buffer. For HA and Satb2-IP, a similar amount of Ig-G Agarose beads was equilibrated with lysis buffer. After pre-clearing, every 100 µg of lysate was mixed with 15 µl of equilibrated beads and the final volume was adjusted to 1 ml with lysis buffer. The samples were rotated overnight at 4 °C. The resin was centrifuged at 8200 g for 30s and the supernatant was discarded. The beads were washed 3 times with lysis buffer containing 300 mM of NaCl, and the immunoprecipitated proteins were eluted at 95 °C for 10 min with 15 µl of SDS Laemmli sample buffer for western blot analysis.
Cell proliferation and annexin V analyses
Cell proliferation was determined by measuring EdU incorporation with the Click-iT EdU flow cytometry assay kit containing Alexa Fluor 647 (Invitrogen). The percentage of apoptotic cells in culture was measured with the PE Annexin V Apoptosis Detection Kit I (BD Pharmingen). The cells were analyzed via flow cytometry in a FACSCalibur (BD Biosciences) machine.

Alkaline phosphatase staining
An alkaline phosphatase detection kit was used (Millipore). ES cells were first plated in ES cell medium on 0.1% gelatine. 24h later the media was switched to differentiation medium when indicated. 3 days later the cells were stained for AP.

RNA-seq analysis
Depth coverages were generated with deeptools computeMatrix bamCoverage (Ramirez et al., 2014) using -bs 10 --skipNonCoveredRegions --effectiveGenomeSize 2611812472 --blackListFileName mm9-blacklist.bed --normalizeUsing RPKM -of bigwig as parameters. Gene set enrichment analyses (GSEAs) were performed using the GSEA package (v2.2.4) (Subramanian et al., 2005), using geneset permutation, weighted statistic, difference of classes on log2+1 transformed Fragments per Kilobase of transcript per million (FPKM). Venn diagrams were generated using BioVenn (Hulsen et al., 2008). Gene ontology analyses were conducted using Homer findGO (v4.7) (Heinz et al., 2010), retrieving enrichment values on the union of the top 5 wikipathways annotations for all combinations of $\text{Satb2}^{+/R}$ and $\text{Zfp451}^{-/-}$ differentially regulated genes, plotting scores using heatmap.2 in R.

ChIP protocol and library preparation
The ChIP protocol for pluripotent and differentiating ES cells was adapted from (Lee et al., 2006) and optimized as follows. Cells were trypsinized, resuspended in 40 ml of ES cell medium containing 1% Methanol-free Formaldehyde (Thermo/Pierce) and incubated for 10 min on a rotor at room temperature (RT). Then, Glycine solution was added to a final concentration of 125 mM and incubated for 5 min on a rotor at RT. Next, the cells were centrifuged at 2000 rpm for 10 min at 4°C and washed two times with 50 ml of cold PBS. The cell pellet was resuspended in 10 ml of lysis buffer 1 (50mM HEPES-KOH, pH 7.5, 140mM NaCl, 1mM EDTA, 10% glycerol, 0.5% NP-40, 0.25% Triton X-100, 1× protease inhibitors) and then incubated on a rotor for 10 min on 4 °C. The cells were then pelleted, resuspended in 10 ml of lysis buffer 2 (10 mM Tris-HCl, pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 1× protease inhibitors) and incubated at RT for 10 min on a rotor. The cells were pelleted again and resuspended in 2ml of sonication buffer (20mM Tris-HCl pH8, 150mM NaCl, 2mM EDTA, 0.1% SDS, 1× Triton X-100, 1× protease inhibitors). The suspension was transferred to 1ml AFA tubes (Covaris) and sonicated with the following settings: 140 peak incidence, 5 duty factor, 200 cycles/burst, 1200 secs in a E220 Ultrasonicator (Covaris). The sonicated lysate was recovered into DNA LoBind tubes (Eppendorf) and centrifuged at 16,000 g for 10 mins at 4°C to eliminate debris. Meanwhile, 100ul of GammaBind G Sepharose (Sigma) beads were washed twice and equilibrated for 15h at 4°C in blocking solution (0.5% BSA in PBS). The sonicated lysate was pre-cleared for 3h at 4°C with 100 µl of blocked beads equilibrated in sonication buffer. The samples were centrifuged
for 3 mins at 16,000 g to recover the precleared lysate, 10% of the total volume was set aside as input control. Each ChIP was set up by adding 10 µg of antibody (Supplemental table S9) to the lysate and incubating it 15h at 4°C on a rotator. The beads were then washed once with the sonication buffer, once with washing buffer 1 (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 2 mM EDTA, 0.1% SDS, 1% Triton X-100), once with washing buffer 2 (10 mM Tris-HCl pH 8.0, 250 mM LiCl, 2 mM EDTA, 1% NP40), once with TE buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA) containing 50 mM NaCl, and twice with TE buffer. Following washing, the beads were resuspended in 100 µl of elution buffer (50 mM Tris-HCl, pH 8.0, 10 mM EDTA, 1.0% SDS) and incubated at 65°C for 30 min at 800 rpm. The supernatant was transferred to a new tube and further 100 µl of elution buffer were added to the beads to repeat the elution and recover any remaining complexes. The input and eluates were then incubated first with 0.1 mg/ml RNase A (Thermo) at 37°C for 2h, and then with 0.1 mg/ml of Proteinase K (NEB) at 55°C for 2h. Finally, the DNA was purified using the QIAquick PCR purification kit (Qiagen) according to the manufacturer instructions, the samples were eluted in 50 µl of milliQ water and stored at -20°C until analysis.

ChIP-seq data processing

Reads were aligned in paired-end mode to the mm9 genome using bowtie2 v2.2.8 (Langmead and Salzberg, 2012) with -very-sensitive-local as a parameter. Peak calling and read coverage generation were performed using macs2 v2.1.2 (Zhang et al., 2008) using the input reads as control for the -c switch, using -f BAMPE -g mm --keep-dup auto -B --trackline --SPMR SPMR as parameters. Read coverages were retrieved over merged peak summits as count matrices derived using Homer annotatePeaks (Heinz et al., 2010) with -hist 10 -ghist -bedGraph -size 400 as parameters. To ascertain no Input signal was present in count matrices, sites showing ≥6 total Input signal ±100 bp of the peak center were discarded (corresponding the average of Input signal in regions visibly depleted of Input). Read coverages around differential and shared regions were computed using deeptools computeMatrix v2.4.2 (Ramirez et al., 2014) with reference-point --referencePoint center -b 3000 -a 3000 as parameters. Heatmaps and average profiles were generated using deeptools plotHeatmap and plotProfile. For correlation clustering of experiments, correlation matrices were derived from read coverages retrieved on ranked merged summits using annotatePeaks, with -hist 400 -ghist -size 400 -bedGraph, and subsequently clustered and output as heatmaps using gplots heatmap.2 in R (R Core Team, 2015; Warnes et al., 2016).

Motif discovery

For motif average profiles, motifs resulting from motif discovery results as well as other known motifs were mapped to regions using Homer annotated Peaks, with -m -size 6000 -mbed, and subsequently output as bedGraph files using piping of bedtools sort to bedtools genomcov with -bga -g mm9.chrom.sizes as parameters. BedGraph files were subsequently converted to bigwig files using bedGraphToBigWig (Kent et al., 2010) for use with deeptools computeMatrix. Average profiles were plotted using deeptools plotProfile. Motif co-occurrence clustering in LIF- and RA-specific regions was performed as previously described (Cauchy et al. 2016) using 1000 iterations of random sampling from mESC ATAC-seq peaks (Wu et al., 2016) to compute background co-occurences.
HiC

Cells were crosslinked for 10 mins at room temperature in 1% methanol-free formaldehyde at a density of 10^6 cells/ml, quenched for 5 mins with 125mM glycine at room temperature. 1.5 10^7 crosslinked cells were resuspended in 500 μL Hi-C Lysis Buffer (10 mM Tris-HCl pH 8.0, 10 mM NaCl, 0.2% NP40, 1X protease inhibitors) for 30 mins at 4°C on a rotating wheel, then centrifuged 5 mins at 2500g, and subsequently washed with 500 μL Hi-C Lysis Buffer and resuspended in 100 μL 0.5% SDS. Cells were then incubated at 62°C for 10 minutes, after which 285 μL of H2O and 50 μL of 10% Triton X-100 were added to quench the SDS, and incubated at 37°C for 15 mins. 50 μL 10X NEB Buffer 2 and 375 U of MboI restriction enzyme (NEB, R0147) were added to digest chromatin for 2 hours at 37°C with rotation, followed by MboI inactivation at 62°C for 10 minutes. Restriction fragment overhang filling and biotin marking of DNA-ends was carried out in 52 μL Master Mix (0.4 mM biotin-dATP (Thermo 19524016), 10 mM dCTP/dGTP/dTTP, 5U/μL DNA Polymerase I, Large (Klenow) Fragment (NEB, M0210)), mixed and incubated at 37 for 1 hour with rotation. 948 μL ligation master mix was subsequently added (10X NEB T4 DNA ligase buffer with 10 mM ATP (NEB, B0202), 10% Triton X-100, 50 mg/mL BSA, 400 U/μL T4 DNA Ligase (NEB, M0202)), incubated at room temperature for 4 hours with rotation. Nuclei were then centrifuged at 2500g for 5 mins. For sonication, pellets were resuspended in 880 μL in Nuclear Lysis Buffer (50 mM Tris-HCl pH 7.5, 10 mM EDTA, 1% SDS, 1X protease inhibitors), and sonicated in a Covaris millitube using a Covaris E220 sonicator (fill level: 10, duty cycle: 5, PIP: 140, cycles/burst: 200, time: 4 mins). Biotin pull-down was carried out with 5 μL Streptavidin C-1 beads washed with Tween Wash Buffer (5 mM Tris-HCl pH 7.5, 0.5 mM EDTA, 1M NaCl, 0.05% Tween-20) in 10 μL 2X Biotin Binding Buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA, 2 M NaCl). Beads were incubated at room temperature with 50ng chromatin for 15 minutes with rotation, magnet-separated and washed twice with 500 μL Tween Wash Buffer at 55°C for 2 minutes with shaking and subsequently washed in 100 μL of 1X TD Buffer (10 mM Tris-HCl pH 7.5, 5 mM MgCl2, 10% dimethylformamide), and resuspended in 25 μL 2X TD Buffer, to which 2.5 μL Tn5 transposase was added, and subsequently incubated at 55°C with interval shaking for 10 minutes. Beads were magnet-separated and incubated in 50 mM EDTA at 50°C for 30 mins, washed twice in 50 mM EDTA at 50°C for 3 mins, and twice in Tween Wash Buffer at 55°C for 2 minutes, as well as once in 10 mM Tris. For PCR and Post-PCR Size Selection, beads were resuspended in 50 μL of PCR master mix (Phusion HF 2X, Nextera Ad1.1 (Universal) 12.5 µM, Nextera Ad2.x (Barcoded) 12.5 µM), with 15 cycles (98°C 15 s, 63°C 30 s, 72°C 1 min). Libraries were cleaned up with 1.8X and subsequently 0.6X Ampure XP beads, and finally eluted in 10 μL water. Sequencing was carried out on an Illumina HiSeq 2500 sequencer.

HiC data analysis

Read 1 and 2 fastq files were aligned separately to the mm9 genome with bwa (v 0.7.16a) (Li and Durbin, 2009) using the following command: bwa mem -A1 -B4 -E50 -L0. Interaction matrices were created using hicBuildMatrix from the HiCExplorer (v2.1.4) package (Ramirez et al., 2018) using both aligned read bam files, with --binSize 40000 --restrictionSequence GATC --threads 4 --inputBufferSize 400000 as parameters. For quality check visualization, matrix bins were merged into groups of 100 bins using hicMergeMatrixBins with --numBins 100 as the parameter. Matrices (both original and merged into 100 bins for quality check)
were interrogated for filter thresholds using hicCorrectMatrix diagnostic_plot, and subsequently corrected using hicCorrectMatrix correct with --filterThreshold as the parameter using the obtained lower and upper filter thresholds. Matrices were correlated using hicCorrelate using --method=pearson --log1p --range 5000:200000 as parameters. For interaction calling, the Homer makeTagDirectory command was used, first using -tbp 1 as the parameter and then filtered using the following arguments: -update -genome mm9 -removePEbg -restrictionSite GATC -both -removeSelfLigation -removeSpikes 10000 5. Interactions were called using findHiCInteractionsByChr.pl WT_RA_filtered -res 2000 -superRes 10000. Matrices from HiCExplorer were converted to the Homer format using hicConvertFormat --inputFormat h5 --outputFormat homer --outFileName as parameters. The union of interactions was computed by merging the output of bedtools pairtopair commands (Quinlan and Hall, 2010) to identify specific and shared interactions (using -v and alternating between inputs to identify specific populations, without -v for shared interactions). The union of interactions was intersected with the union of Satb2 d0 LIF and d4 RA peaks, with intersecting matrix bins recovered into a separate matrix for each condition. Resulting interaction matrices were plotted using Java TreeView (v1.1.6r2) (Saldanha, 2004). Venn diagrams of interaction anchor points were plotted using ChIPpeakAnno (v3.11) (Zhu et al., 2010). Sørensen-dice coefficients (Dice, 1945) were clustered and plotted using the heatmap.2 function of the R gplots package (Warnes et al., 2009).

Public dataset processing

ChIP-seq data for H3K9me2 in wt mESCs cultured in LIF and d2.5 RA (Liu et al., 2015) were retrieved from GEO series GSE54412 mESC ATAC-seq data (Wu et al., 2016), from GEO series GSE66581 and d4 RA ATAC-seq data (Rhee et al., 2016) from GEO series GSE79561. Data was processed with exactly the same methodology and parameters as used for the ChIP-seq data generated in this study.

Supplemental References

Dice, L.R. (1945). Measures of the Amount of Ecologic Association Between Species. Ecology 26, 297-302.

Hailesellasse Sene, K., Porter, C.J., Palidwor, G., Perez-Iratxeta, C., Muro, E.M., Campbell, P.A., Rudnicki, M.A., and Andrade-Navarro, M.A. (2007). Gene function in early mouse embryonic stem cell differentiation. BMC Genomics 8, 85.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Molecular cell 38, 576-589.

Hulsen, T., de Vlieg, J., and Alkema, W. (2008). BioVenn - a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. BMC Genomics 9, 488.

Kent, W.J., Zweig, A.S., Barber, G., Hinrichs, A.S., and Karolchik, D. (2010). BigWig and BigBed: enabling browsing of large distributed datasets. Bioinformatics 26, 2204-2207.

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

Lee, T.I., Johnstone, S.E., and Young, R.A. (2006). Chromatin immunoprecipitation and microarray-based analysis of protein location. Nature protocols 1, 729-748.
Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754-1760.

Liu, N., Zhang, Z., Wu, H., Jiang, Y., Meng, L., Xiong, J., Zhao, Z., Zhou, X., Li, J., and Li, H. (2015). Recognition of H3K9 methylation by GLP is required for efficient establishment of H3K9 methylation, rapid target gene repression, and mouse viability. Genes & development 29, 379-393.

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

R Core Team (2015). R: A Language and Environment for Statistical Computing (Vienna, Austria: R Foundation for Statistical Computing).

Ramirez, F., Bhardwaj, V., Arrigoni, L., Lam, K.C., Gruning, B.A., Villaveces, J., Habermann, B., Akhtar, A., and Manke, T. (2018). High-resolution TADs reveal DNA sequences underlying genome organization in flies. Nat Commun 9, 189.

Ramirez, F., Dundar, F., Diehl, S., Gruning, B.A., and Manke, T. (2014). deepTools: a flexible platform for exploring deep-sequencing data. Nucleic acids research 42, W187-191.

Rhee, H.S., Closser, M., Guo, Y., Bashkirova, E.V., Tan, G.C., Gifford, D.K., and Wichterle, H. (2016). Expression of Terminal Effector Genes in Mammalian Neurons Is Maintained by a Dynamic Relay of Transient Enhancers. Neuron 92, 1252-1265.

Saldanha, A.J. (2004). Java Treeview--extensible visualization of microarray data. Bioinformatics 20, 3246-3248.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 102, 15545-15550.

Warnes, G.R., Bolker, B., Bonebakker, L., Gentleman, R., Huber, W., Liaw, A., Lumley, T., Maechler, M., Magnusson, A., and Moeller, S. (2009). gplots: Various R programming tools for plotting data. R package version 2, 1.

Warnes, G.R., Bolker, B., Bonebakker, L., Gentleman, R., Huber, W., Liaw, A., Lumley, T., Maechler, M., Magnusson, A., Moeller, S., et al. (2016). gplots: Various R Programming Tools for Plotting Data.

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

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

Zhu, L.J., Gazin, C., Lawson, N.D., Pages, H., Lin, S.M., Lapointe, D.S., and Green, M.R. (2010). ChIPPeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. BMC Bioinformatics 11, 237.