Cell Stem Cell, Volume *2*

Supplemental Data

Global Transcription

in Pluripotent Embryonic Stem Cells

Sol Efroni, Radharani Duttagupta, Jill Cheng, Hesam Dehghani, Daniel J. Hoeppner, Chandravanu Dash, David P. Bazett-Jones, Stuart Le Grice, Ronald D. G. McKay, Kenneth H. Buetow, Thomas R. Gingeras, Tom Misteli, and Eran Meshorer

Supplemental Experimental Procedures

Calculating the Number of Transcripts per Cell

Cells from which RNA was extracted were counted and total RNA yield per cell was estimated. Mean total RNA yield from 10^6 cells was ~20 µg, corresponding to ~20 pg of RNA per cell, resembling textbook estimates. To quantify the number of transcripts per cell, we reverse transcribed 10 µg of RNA (i.e. 5E5 cells), out of which 5% (i.e. 25,000 cells) were used for each PCR reaction. To generate the standard curve, we used known starting amounts of plasmids (we used both GFAP- and myogenin-containing plasmids), ranging from 5 to 0.0005 pg plasmid. Based on the plasmid's molecular weight (6.7 kb x 660 ~ 4.22 MegaDalton) divided by Avogadro's number (6.022×10^{23}), we calculated the plasmid copy number per reaction multiplied by the number of cells in each reaction to estimate the copy number per cell. We found 0.25 copies per cell for GFAP and 20 copies per cell for myogenin.

Figure S1

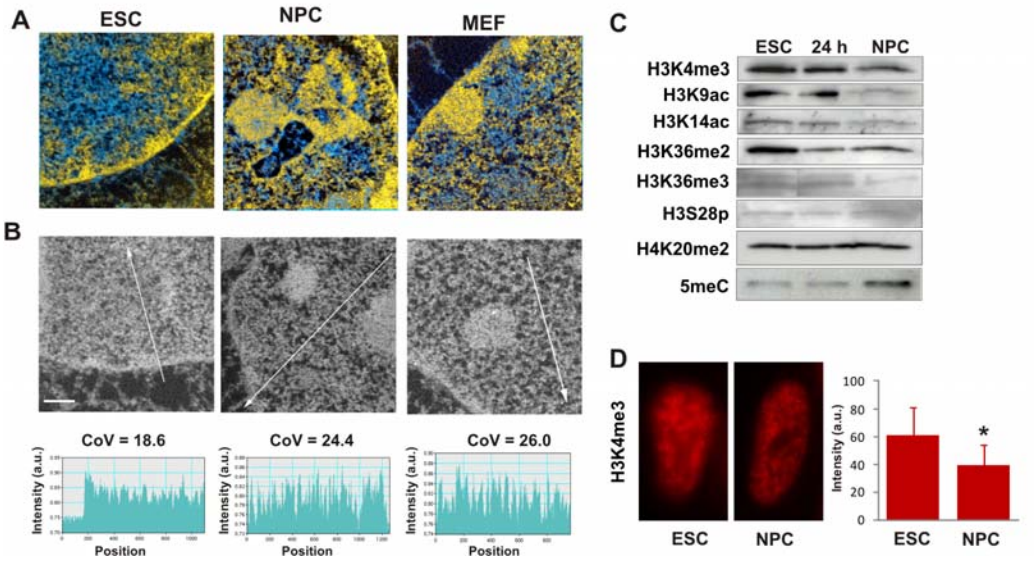


Figure S1. Decondensed and Actively Marked Chromatin in ES Cells

(A) Phosphorus (yellow) and nitrogen-phosphorus (N-P) (blue) maps in undifferentiated ES cells (ESC), ES cells-derived 7 days neuronal progenitor cells (NPC) and mouse embryonic fibroblasts (MEF). ES cells exhibit homogeneous chromatin texture and are devoid of detectable heterochromatin foci. Bar = 500 nm.

(B) Line scans of mass distribution (nitrogen maps) in ES cells (left), NPC (middle) and MEFs (right). CoV = Coefficient of Variation (% standard deviation/mean). At least 10 scans of each of 10 cells were analyzed for ESC and NPC and 4 cells for MEFs. Representative images are shown.

(C) Western blots for histone modifications as well as methylated DNA (5-meC) in undifferentiated ES cells (left), cells 24 hrs after withdrawal of LIF (middle) and NPC (right). Active histone marks are enriched in ES cells. Equal amounts of protein were loaded and histone H3 serves as a loading control.

(D) Immunofluorescence analysis for H3K4m3 in undifferentiated ES cells (ESC) and ES cell-derived neuronal progenitor cells (NPC). Fluorescence intensity quantification (n>30 cells) is shown on the right.

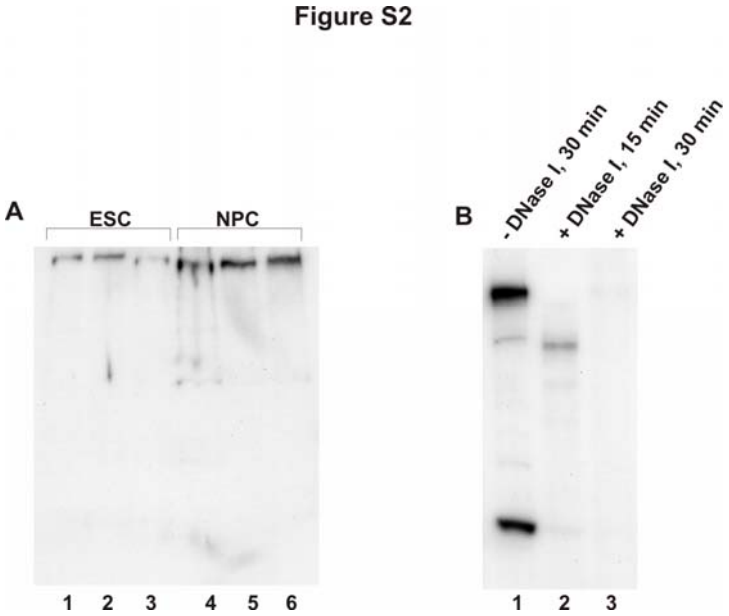


Figure S2. Detection of DNA Methylation by Western Blot

(A) Purified nuclei from three preparations of undifferentiated ES cells (lanes 1-3) and three preparations of NPC (lanes 4-6) were denatured in Laemmli buffer and electrophoresed on a 4-20% gradient SDS-PAGE gel (BioRad). Gel was blotted using a polyvinylidene fluoride (PVDF) transfer membrane (Immobilon-P, Millipore, Billerica, MA), which readily blots both proteins and nucleic acids, blocked, and incubated 2 hrs at room temperature or over-night at 4°C with a mouse monoclonal 5-methylcitidine antibody (Eurogentec) in 3% skim milk and 2% BSA (1:1000). Following standard washes and blocking, the blot was incubated with a secondary anti-mouse antibody (GE Healthcare, Buckinghamshire, UK) for 2 hrs at room temperature. Detection was performed with ECL-plus kit (GE). Note that one 'heavy' band is visible, likely corresponding to bulk DNA.

(B) Specificity of the detected band. Purified nuclei from undifferentiated ES cells were pre-incubated in DNase I buffer (10 mM Tris-HCl; 2.5 mM $MgCl_2$; 0.5 mM $CaCl_2$) in the absence (lane 1, 40 min) or presence (lane 2, 20 min and lane 3, 40 min) of DNase I enzyme (5 µg/ml). Pre-incubation with DNAse I was performed to ensure that DNA and only DNA is detected using Western blots. Note complete disappearance of detectable bands following DNase I treatment, attesting to the specificity of the reaction. Also note appearance of an additional band in lane 1, due to pre-incubation.

Figure S3

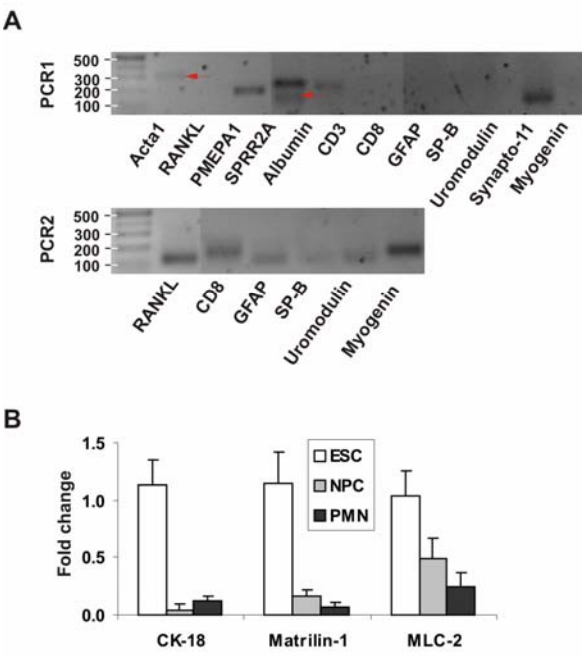
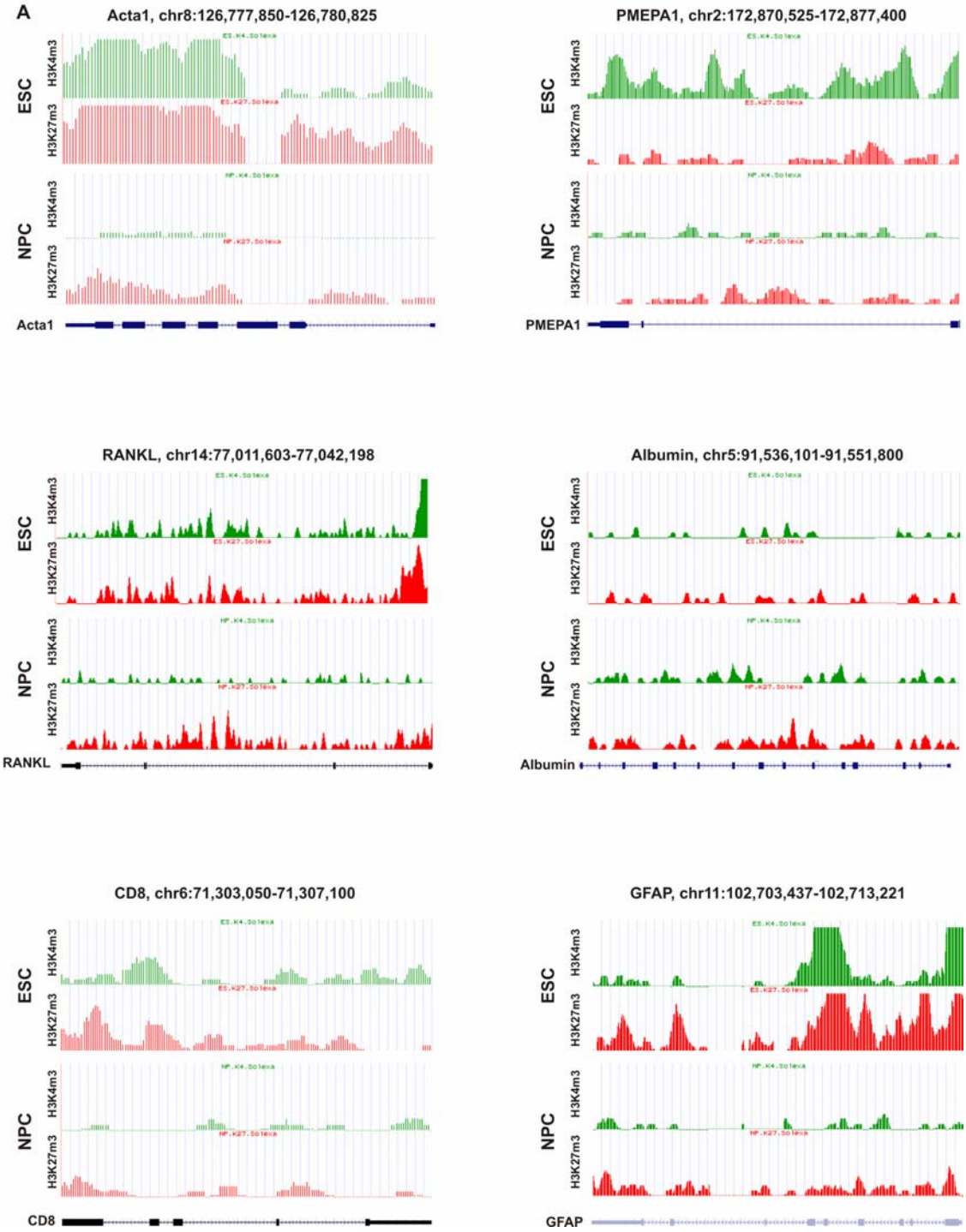


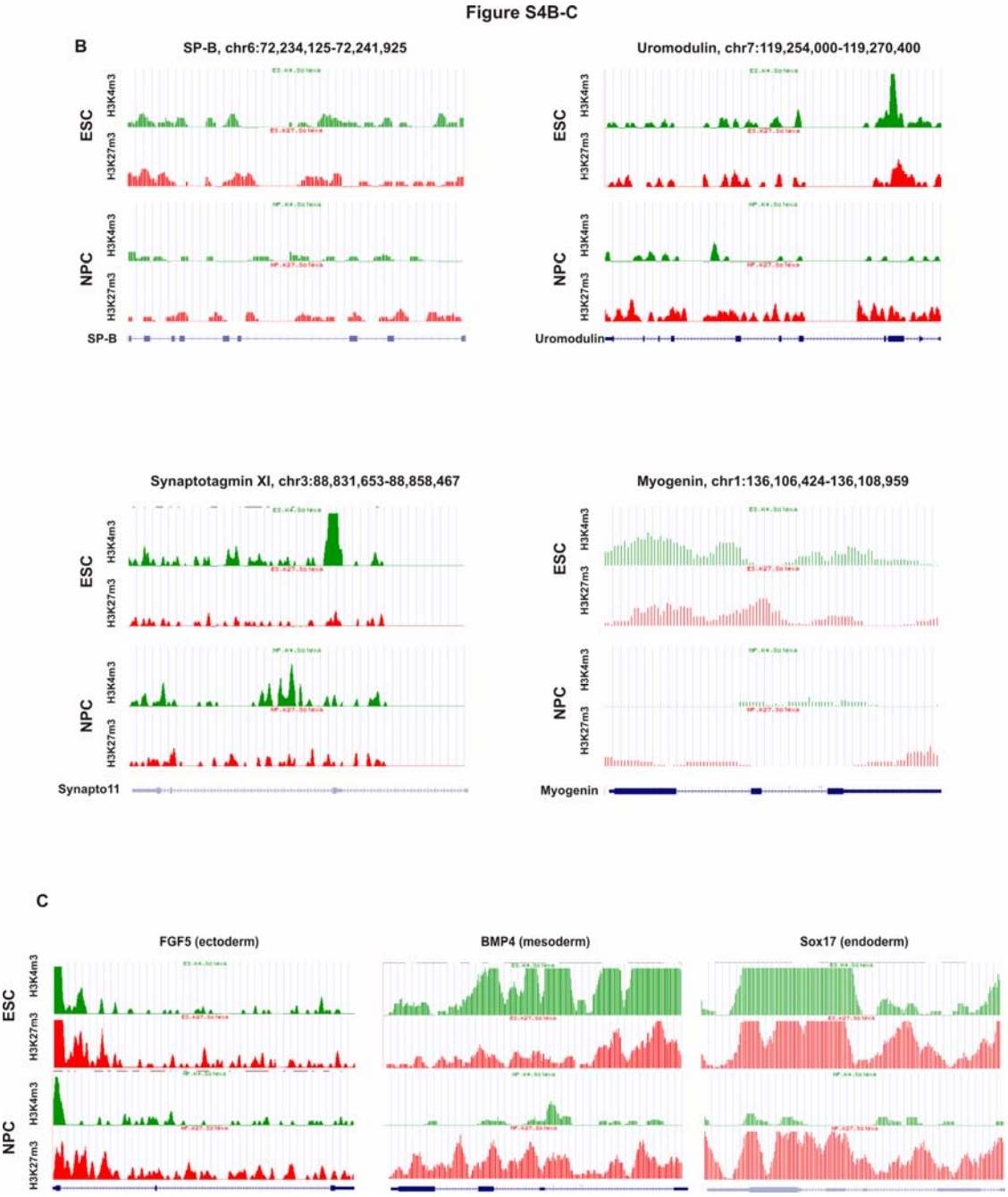
Figure S3. Detection of Lineage-Specific Genes in ES Cells

(A) RT-PCR detection. Top: RT-PCR products for 12 lineage-specific genes. Five genes (Acta1, PMEPA1, SPRR2A, Albumin and Synaptotagmin-11) were detected in the first RT-PCR reaction. Bottom: second PCR was performed for the seven genes that failed to yield a product in the first RT-PCT reaction. Of these, the six noted genes were detected in the second reaction, suggesting low-level transcription.

(B) Real-time RT-PCR for three lineage-specific genes, the transcription level of which is reduced following differentiation. ESC: undifferentiated ES cells; NPC: ES cells-derived neuronal progenitor cells; PMN: ES cells-derived pot-mitotic neurons.

Figure S4A


A
Acta1, chr8:126,777,850-126,780,825
PMEPA1, chr2:172,870,525-172,877,400
RANKL, chr14:77,011,603-77,042,198
Albumin, chr5:91,536,101-91,551,800
CD8, chr6:71,303,050-71,307,100
GFAP, chr11:102,703,437-102,713,221
ESC
NPC
H3K4m3
H3K27m3
Acta1
PMEPA1
RANKL
Albumin
CD8
GFAP



Figures S4. ChIP-Seq Data for the Selected Lineage-Specific Genes

(A-B) Chromatin immunoprecipitation (ChIP) followed by single-molecule-based sequencing (ChIP-seq) using Illumina/Solexa technology was performed and described by Mikkelsen et al (2007). In this approach, DNA molecules that are pulled-down by ChIP are spotted on an array and amplified locally. Then, successive rounds of primer-mediated single-base extension are applied, using fluorescently labeled reversible terminators, and imaged after every round. ChIP was performed with the noted

antibodies: H3K4m3 is shown in green and H3K27m3 in red for both ES cells (ESC) and ES cell-derived neuronal progenitor cells (NPC). Gene names and their genomic locations appear at the top of each panel and complete gene annotations at the bottom. While genes are bivalently marked in ESC, the H3K4m3/H3K27m3 ratio is always higher in ESC except for Synaptotagmin which is expressed in NPC. These results support low-level transcription of these transcripts in ES cells.
(C) ChIP-seq data for classic markers of the three germ layers are shown. FGF5 is a marker for ectoderm, BMP4 for mesoderm and Sox17 for endoderm. In all cases, the H3K4m3/H3K27m3 ratio is higher in ESC supporting transcriptional activation at low levels. All data in this figure is from Mikkelsen et al, 2007.

Figure S5

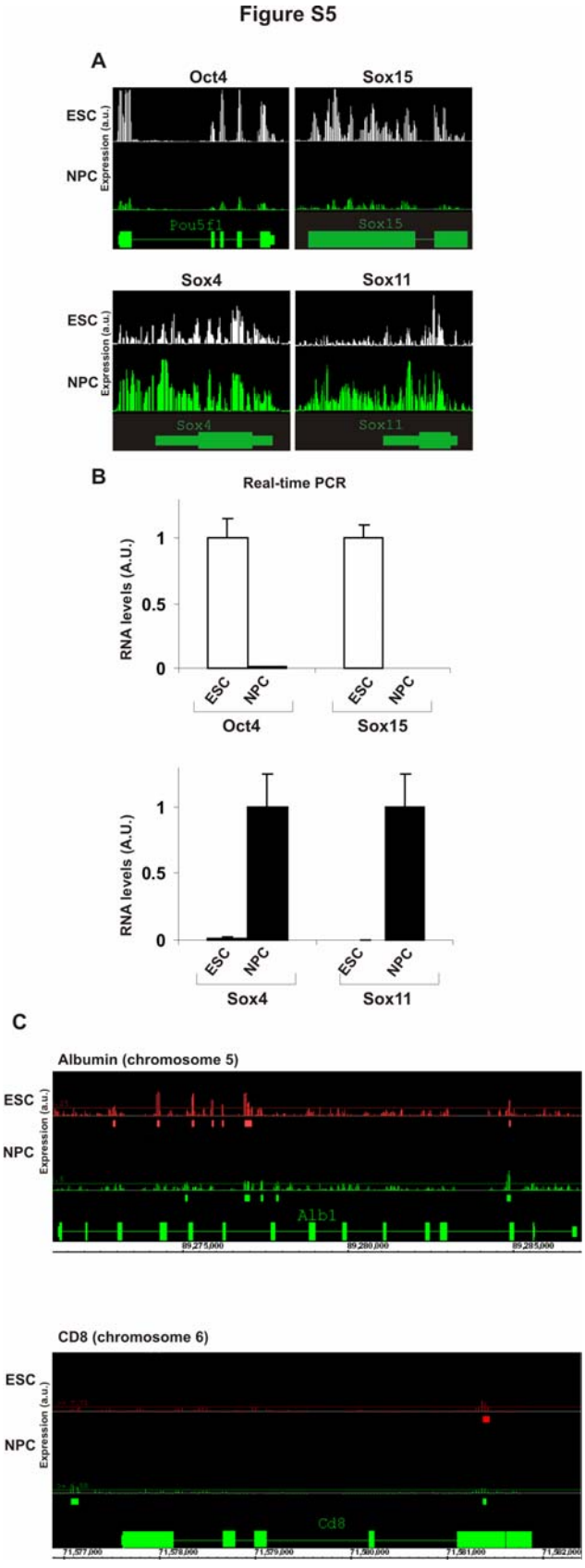

A
Oct4
Sox15
ESC
NPC
Expression (a.u.)
Pou5f1
Sox15
Sox4
Sox11
Sox4
Sox11
B
Real-time PCR
RNA levels (A.U.)
1
0.5
0
ESC
NPC
Oct4
Sox15
Sox4
Sox11
C
Albumin (chromosome 5)
Alb1
CD8 (chromosome 6)
Cd8

Figure S5. Microarray Data Validation

(A) Tiling microarray expression pattern of the stem cell markers Oct4 (Pou5f1) (top left) and Sox15 (top, right) as well as the neuronal progenitor markers Sox4 (bottom, left) and Sox11 (bottom, right) in ES cells (ESC, white) and in ES cells-derived neuronal progenitor cells (NPC, green). Each panel represent composite graphs describing signal intensity form the three independent biological replicas to represent probe intensity per genomic coordinates. All coordinates are in the mm.NCBIv33 version of the mouse genome.

(B) Semi-quantitative real-time RT-PCR for Oct4, Sox15, Sox4 and Sox11. The following primers were used:

Oct4-F: 5-AGCTGCTGAAGCAGAAGAGG-3; Oct4-R: 5-CCTGGGAAAGGTGTCCTGTA-3

Sox4-F: 5-CGCCTTGGTGATTTCTTGTT-3; Sox4-R: 5-ATCACCCCAGAGCCTTCTTT-3

Sox11-F: 5-CAGGACCTGTTGCTGTCAGA-3; Sox11-R: 5-GGTCACGAAGCTGGTTTGAT-3

Sox15-F: 5-CGGCGTAAGAGCAAAAACTC-5; Sox15-R: 5-TGGGATCACTCTGAGGGAAG-3

(C) Microarray detection. Shown are the annotated regions of albumin (top) and CD8 (bottom). Both genes display only background levels of transcription on the microarray and most exons do not generate transfrags.

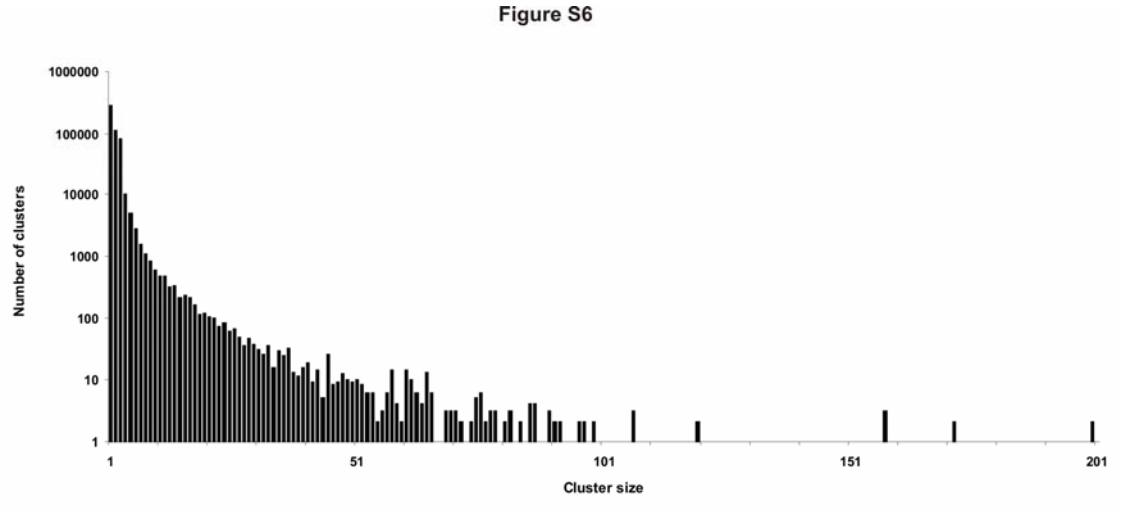


Figure S6. Cluster Size Distribution Maps of the Probes Present in ESC and Absent in NPC

In order to determine the contiguity of the detected transcription in ESC cells, probes selected to be positive in ES cells but absent in NPC were clustered into bins of varying sizes. The frequency of each cluster is represented in the ordinate and the size (# of adjacent probes) in the abscissa. For example, a frequency of 100,000 clusters of size 3 that fit the criteria. This means 100,000 instances of three adjacent probes that were both present in ES cells and absent in NPC across the whole genome (both intragenic and intergenic regions). Since each probe encompasses 25 bp of genomic DNA with a gap of 5 bp between two adjacent probes, a cluster of 3 probes represent an approximate genomic size of 100 bp. Cluster size of 200 corresponds to 6 kb.

Figure S7

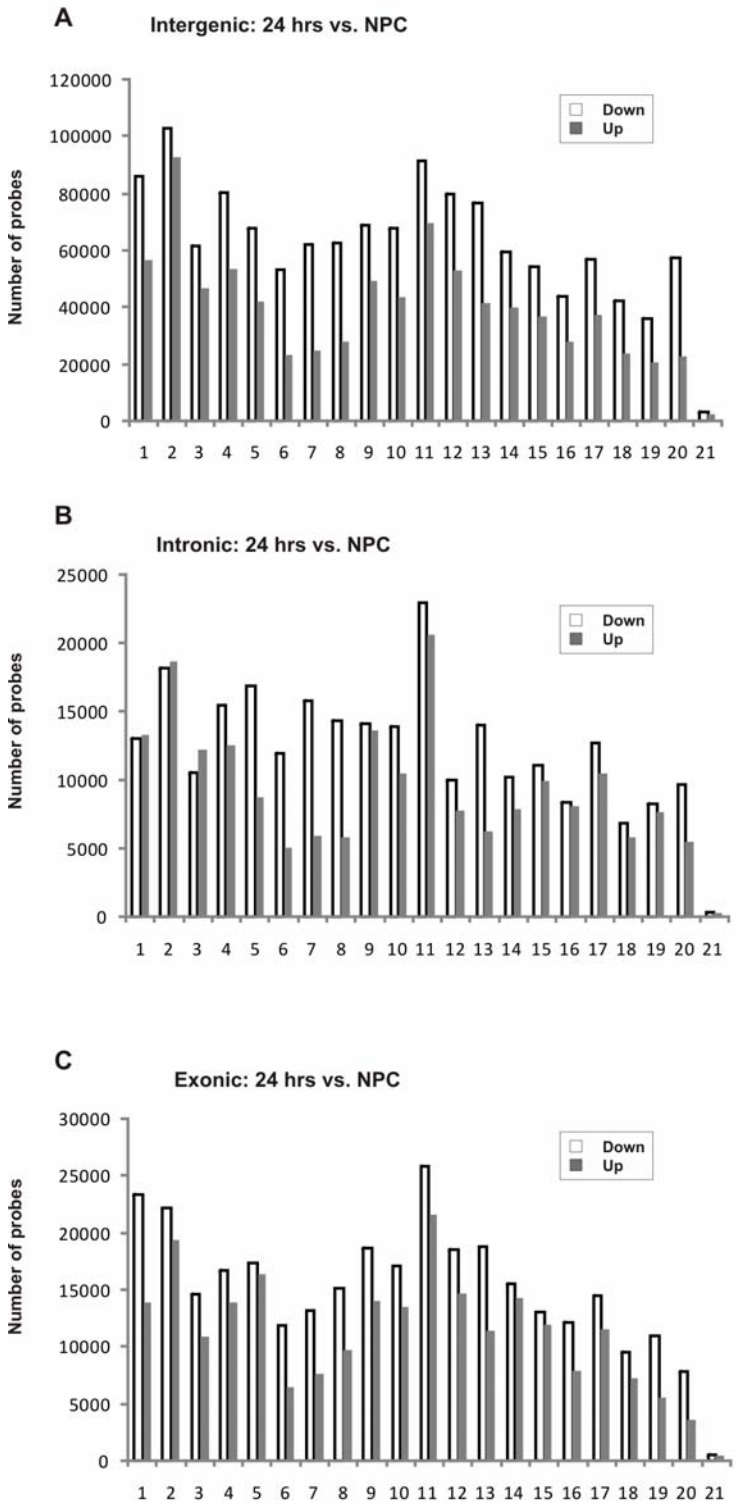
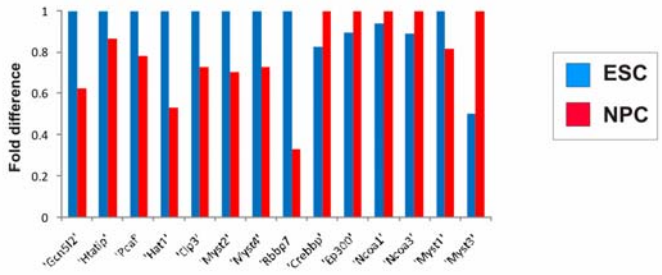

A
Intergenic: 24 hrs vs. NPC
Number of probes
Down
Up
B
Intronic: 24 hrs vs. NPC
Number of probes
Down
Up
C
Exonic: 24 hrs vs. NPC
Number of probes
Down
Up

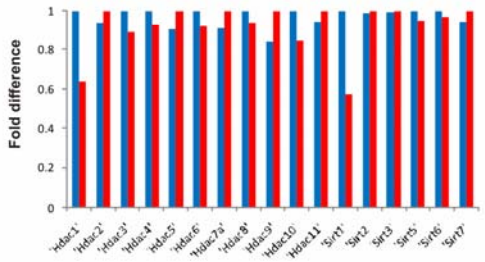
Figure S7. Global Expression Changes during ES Cell Differentiation – 24 Hours versus NPC

(A-C) Comparison of positive probes between cells 24 hrs after LIF withdrawal and between ES cells-derived neuronal progenitor cells (NPC). Total numbers of down-regulated and up-regulated probes are depicted as white and gray bars respectively for intergenic regions (A), intronic regions (B) and exonic regions (C) for all mouse chromosomes. Only probes which were positive in both time points were used for this analysis. Data represents the average of three independent experiments.

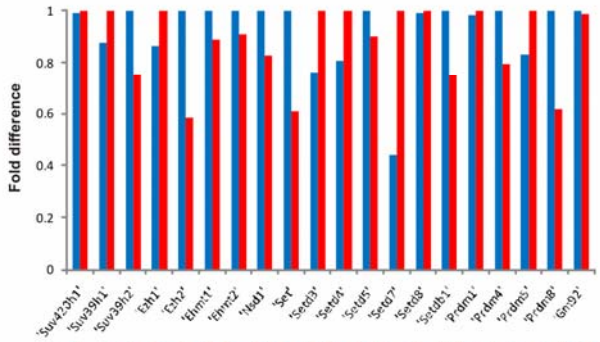
HATs



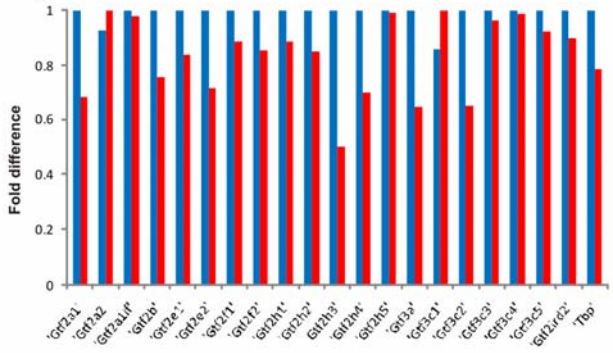
HDACs



HMTs



GFTs



Remodelers

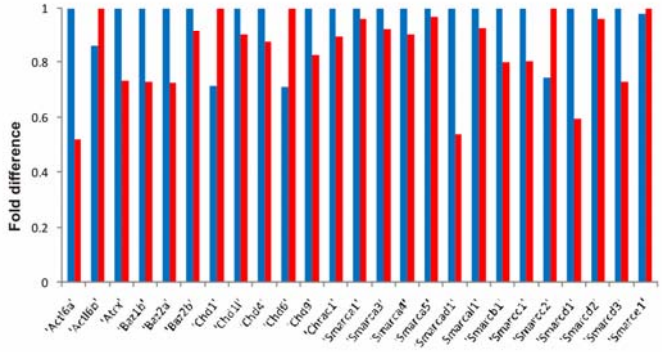


Figure S8. Comparison of Expression Levels of Specific Classes of Genes in Differentiating ES Cells

Shown are relative expression levels of microarray data in ES cells (ESC, blue) and ES cells-derived neuronal progenitor cells (NPC, red) of histone acetyl transferases (HATs), histone deacetylases (HDACs), histone methyl transferases (HMTs), general transcription factors (GTFs) and chromatin remodeling factors (Remodelers). Data is the same as presented in Figure 5 but displayed as bar graphs rather than heat-maps.

Figure S9

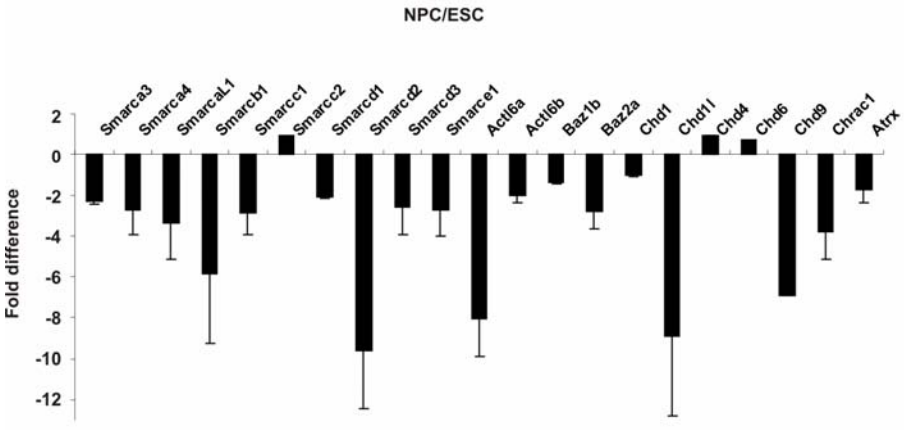


Figure S9. Real-Time qRT-PCR Analysis of Chromatin Remodeling Genes

Validation of microarray expression data for chromatin remodelers by quantitative real-time RT-PCR. From a total of 21 genes for which signal could be obtained from both ESC and NPC states, 17 displayed reduced expression in NPC, most of which correlated with the array data. Of the four genes that displayed either slight elevation or no change (Chd1, Chd4, Chd6, Smarcc2) three (Chd1, Chd6, Smarcc2) were also in agreement with the microarray data ($p < 0.05$). An additional 4 genes (Smarca1, Smarca5, Smarcad1 and Baz2b) which were detected only in the ES cells qRT-PCR measurements were excluded. All measurements were normalized against Cyclophilin B. Values shown are fold difference representing averages ± SD from 3 independent experiments. These results strongly corroborate the microarray analysis and reconfirm the abundance of chromatin remodeling factors and general transcription factors in undifferentiated ES cells.

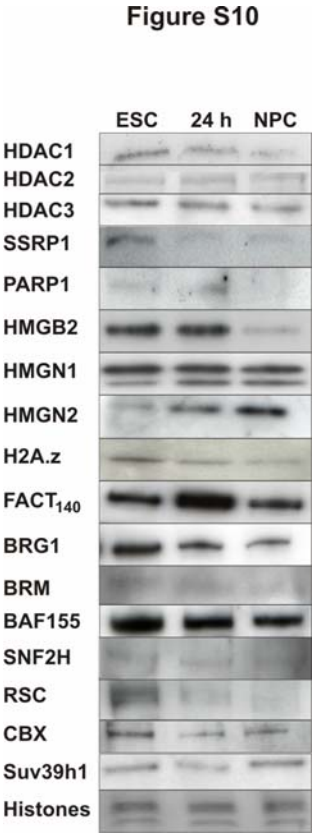


Figure S10. Protein Levels for Various Chromatin-Related Factors during ES Cell Differentiation

Western blots for a variety of chromatin proteins from nuclear extracts of undifferentiated ES cells (ESC), 24 hrs following LIF withdrawal (24 h) or 7 d after induction of differentiation (NPC) were probed for the indicated proteins. Total histones served as a loading control.