

## **Supplementary material Peric-Hupkes, Meuleman *et al.***

Consisting of:

- Supplementary methods
- Supplementary table
- Supplementary figure legends
- Supplementary figures

## Supplementary methods Peric-Hupkes, Meuleman *et al.*

### Cell culture and differentiation

E14Tg2A ESCs, cultured under feeder-free conditions, were maintained in a mixture of 40% Glasgow MEM (GMEM, GIBCO-BRL) plus 60% GMEM with 15% heat-inactivated ES-qualified fetal bovine serum (FBS, GIBCO-BRL) that was conditioned on Buffalo Rat Liver cells. This mixture was supplemented with 0.055 mM beta-mercaptoethanol (GIBCO-BRL), 2mM L-glutamine, 0.1mM nonessential amino acid, 5000 U/ml penicillin/streptomycin, and 1000 U/ml LIF (Chemicon). NPCs were derived from E14Tg2A as published (Ying and Smith, 2003), and eventually propagated in Dulbeccos modified Eagles medium/F12 (DMEM/F12, GIBCO-BRL) supplemented with 1:100 N2 (25 mg Insulin, 100 mg Apo-transferrin, 20 ng Progesterone, 16 mg Putrescine, 30  $\mu$ M Sodium Selenite (all Sigma), and 5 mg BSA (GIBCO-BRL) in 10 mL DMEM/F12), EGF 20 ng/mL, and FGF 20 ng/mL. ACs were obtained by culturing NPCs for at least 48 hours in DMEM/F12 with 1:100 N2 and 2% FCS(Conti et al., 2005). NIH3T3 MEFs were maintained in DMEM (GIBCO-BRL) with 10% FBS (Greiner Bio-one).

### DamID

DamID was performed as described (Vogel et al., 2007). Briefly, cells were plated on day 1, transduced with Dam-LmnB1 or Dam lentivirus on day 2, and genomic DNA was harvested on day 4. For ESCs and ACs this protocol was slightly modified: ESCs were transduced twice for a period of 6 hours on day 2 and 3, whereas ACs were first differentiated for 48 hours and then transduced. Adenine-methylated fragments were amplified from genomic DNA by methylation-specific PCR, and purified using Qiaquick columns (Qiagen).

For DamID array hybridizations 1.5  $\mu$ g of amplified methylated DNA was labeled as per Nimblegen protocol for Klenow labeling of ChIP samples with either Cy5 or Cy3. Labeling reaction time was extended to 12 hours to increase yield. Two-color hybridizations (Dam-LaminB1 versus Dam only) were performed with 80  $\mu$ g of DNA from each labeling reaction in 110  $\mu$ l hybridiza-

tion volume. Arrays were hybridized as per Nimblegen's recommendations in a Tecan HS4800 hybridization station and scanned with an Agilent G2505C scanner at dual-pass 2 micron resolution. Scans were quantified using NimbleScan 2.5 software.

## **Immunofluorescence microscopy and 3D-FISH**

Immunofluorescence microscopy was done as described (van Steensel et al., 1995) with anti-V5 (Invitrogen, cat. # R960-25), anti-Nestin (BD biosciences, cat. # 611659), or anti-GFAP (DAKO, cat. # M0761) antibodies. For 3D-FISH, amplified adenine-methylated DNA fragments from MEFs expressing Dam-LaminB1 or Dam were labeled using nick-translation with TAMRA-dUTP and FITC-dUTP respectively. 3D-FISH with mixed LaminB1-Dam and Dam probes was performed on MEFs according to standard protocol (Cremer et al., 2007).

## **Nimblegen Array design**

Candidate probe sequences were initially selected from the mouse genome (NCBI m37) using the algorithm as published (Gräf et al., 2007), with a median spacing of 250 basepairs and parameter settings  $l=50$  (probe seed length),  $m=GATC$  (disallowed motif),  $p=80$  (reject probes with palindromic content  $> 80\%$ ),  $t=74-78$  (temperature range in degree Celsius),  $w=250$  (window size). From this set we then selected probes that mapped back uniquely to the mouse genome and tiled the mouse genome at a median spacing of 1,200 bp on a single Nimblegen HD2 array.

## **Nimblegen data normalisation**

For each cell type (ESC, NPC, AC and MEF) we performed two biologically independent replicate DamID experiments. NimbleScan quantified data was first LOESS normalized, using the Ringo package for R. Subsequently, single channel data were obtained,  $\log_2$ -transformed and altogether subjected to quantile-quantile normalisation over all available cell types, replicates and channels (i.e., both Dam-LaminB1 and Dam-only).

For each experiment, resulting  $\log_2$ -transformed, normalized, single channel data were then re-combined by subtracting the Dam-only channel data from the Dam-LaminB1 channel data. This resulted in profiles

$$\text{Lam}_{\text{ESC}_1}, \text{Lam}_{\text{ESC}_2}, \text{Lam}_{\text{NPC}_1}, \dots, \text{Lam}_{\text{MEF}_2},$$

where, e.g.,  $\text{Lam}_{\text{ESC}_1}$  denotes the first replicate experiment for ESC, with

$$\text{Lam}_{\text{ESC}_1} = \{\text{Lam}_{\text{ESC}_1}(1), \text{Lam}_{\text{ESC}_1}(2), \dots, \text{Lam}_{\text{ESC}_1}(N)\}$$

a row vector where  $N$  is the number of probes on the Nimblegen array. Next, same-cell type replicates were averaged for use in further analyses, resulting in profiles  $\text{Lam}_{\text{ESC}}$ ,  $\text{Lam}_{\text{NPC}}$ ,  $\text{Lam}_{\text{AC}}$  and  $\text{Lam}_{\text{MEF}}$ , e.g.,

$$\text{Lam}_{\text{ESC}} = \frac{1}{2}(\text{Lam}_{\text{ESC}_1} + \text{Lam}_{\text{ESC}_2}).$$

## Definition of Lamina Associated Domains (LADs)

LADs were defined using the algorithm described previously (Guelen et al., 2008). Parameters were optimized to yield LADs with an FDR < 1%.

## Test for change in LaminB1 interaction

Although we observe that the global structure of LaminB1 interactions is similar for each of the investigated cell types, we also see local changes occurring. The following section describes a method to test for statistically significant changes in LaminB1 interaction. We emphasize that our statistical test evaluates quantitative changes in the DamID signals for each gene, rather than the binary LAD versus inter-LAD status, which is defined on a coarser scale. As a consequence, some genes may show significant changes in DamID signals but not a change in their LAD status.

Here, as an example, we focus on the differentiation step from ESC to NPC, for which we define

$$\Delta\text{Lam} = \text{Lam}_{\text{NPC}} - \text{Lam}_{\text{ESC}},$$

with  $\Delta\text{Lam} = \{\Delta\text{Lam}(1), \Delta\text{Lam}(2), \dots, \Delta\text{Lam}(N)\}$  again a row vector where  $N$  is the number of probes on the Nimblegen array.

### Per-probe test

As a premise to test whether genes significantly change LaminB1 interaction, we first consider a theoretical test for changes in LaminB1 interactions of single probes.

We define  $\Phi_x$  as the technical and non-specific biological variance between biological replicates of cell type  $X$ , estimated by, e.g. for ESC,

$$\Phi_{\text{ESC}} = \frac{1}{2} (\text{Lam}_{\text{ESC}_1} - \text{Lam}_{\text{ESC}_2}). \quad (1)$$

A vector  $\Delta\Phi$  functions as the basis for a null distribution against which to test single probes from  $\Delta\text{Lam}$  for statistical significance.  $\Delta\Phi$  is constructed by concatenating differences between  $\Phi_x$  instances of the cell types to be compared. Note that there is no reason to prefer a specific order for the same cell type replicates in Equation 1, so we include all possible orderings for the two cell types being compared. If we wish to study differences between ESC and NPC,  $\Delta\Phi$  is defined as

$$\Delta\Phi = \{\Phi_{\text{NPC}} - \Phi_{\text{ESC}}, \Phi_{\text{NPC}} - (-\Phi_{\text{ESC}}), (-\Phi_{\text{NPC}}) - \Phi_{\text{ESC}}, (-\Phi_{\text{NPC}}) - (-\Phi_{\text{ESC}})\},$$

resulting in a vector of length  $4N$ .

When we view  $\Delta\Phi$  as a set of realizations of the random variable  $\phi$ , then  $\phi$  is approximately Gaussian distributed, with parameters  $\mu_\phi$  and  $\sigma_\phi^2$ , i.e.,  $\phi \sim N(\mu_\phi, \sigma_\phi^2)$  (Figure 1A in this document). The elements of  $\Delta\text{Lam}$  also follow an approximately Gaussian distribution (Figure 1B in this document).

Using the above we can, in theory, test whether a single probe shows a significant change in LaminB1 interaction by comparing its value to  $\phi$ . In the paper, we test for changes in LaminB1 interactions on a per-gene basis, which is an extension of the per-probe test described above.

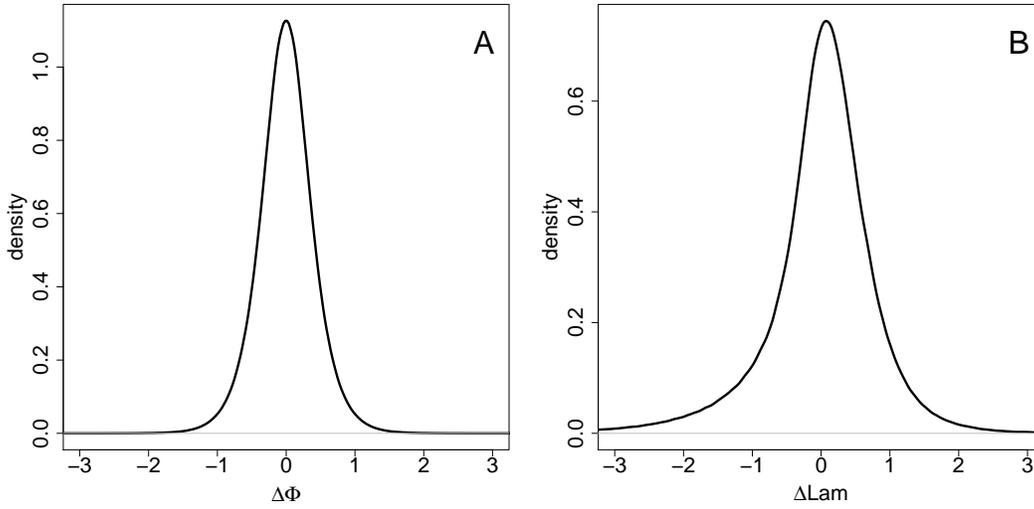


Figure 1: Density plots of (A)  $\Delta\Phi$  and (B)  $\Delta\text{Lam}$

### Per-gene test

We define the set  $G$  as all genes in the genome spanning  $k \geq 5$  or more Nimblegen probes. For each gene  $g \in G$  we define the summed change in LaminB1 interaction

$$\Delta\text{Lam}_g = \sum_{i=1}^k \Delta\text{Lam}(g_i),$$

where  $g_i$  is the genomic index of the  $i$ -th probe of  $g$ .

In order to test whether a gene undergoes a significant change in LaminB1 interaction, we test whether  $\Delta\text{Lam}_g$ , i.e., the sum of the  $\Delta\text{Lam}$  values of  $k$  probes covering the gene, is significant. In the previous section we established that the  $\Delta\text{Lam}$  value of a probe is approximately normal under the null, i.e.,

$$\Delta\text{Lam} \sim N(\mu_\phi, \sigma_\phi^2).$$

If we assume that the  $k$  probes covering a gene are independent and identically (Gaussian)

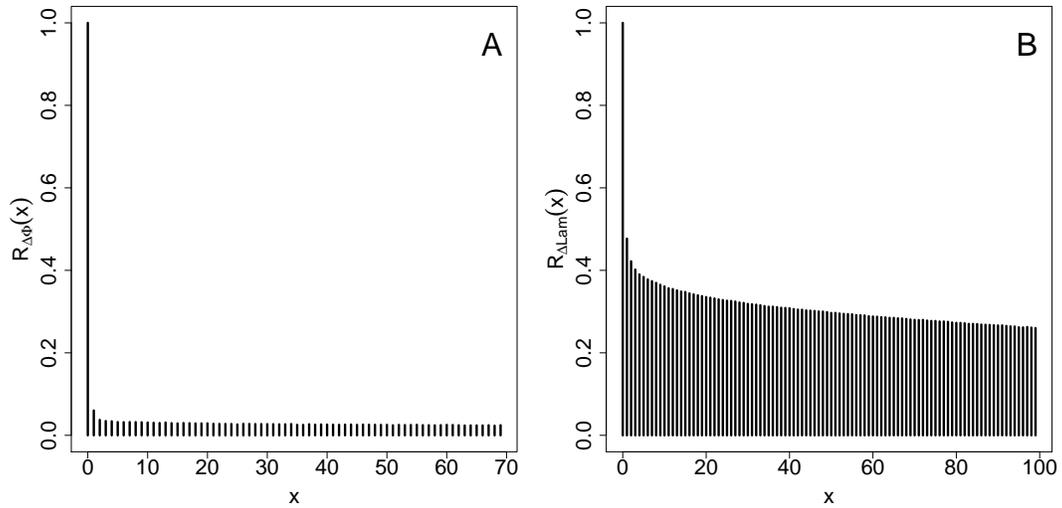


Figure 2: Autocorrelation of (A)  $\Delta\Phi$  and (B)  $\Delta\text{Lam}$  for lags ranging from 0 to 100 probes

distributed, then, under the null,

$$\sum_{i=1}^k \Delta\text{Lam}(g_i) \sim N(k\mu_\phi, k\sigma_\phi^2).$$

However, in the actual  $\Delta\text{Lam}$  signal, neighboring probes are not independent, as evidenced by an increased autocorrelation. Autocorrelation is the correlation of a variable with itself over varying lags. The autocorrelation of a signal with lag  $x$  is an estimate of the dependence present over a range of  $x$  successive values, in this case probes. Figure 2 in this document shows the autocorrelation values for (A)  $\Delta\Phi$  and (B)  $\Delta\text{Lam}$ , clearly showing an increased autocorrelation, and thus dependence, in  $\Delta\text{Lam}$ . Not accounting for this observed dependence will result in a non-conservative test. Basically, any gene showing a marginal effect in LaminB1 interaction change but with a large number of probes will yield a significant p-value.

In the general case, dependency between the random variables included in the sum is accounted for by adding an extra term to the variance, namely the covariance between the random

variables, resulting in

$$N \left( k\mu_\phi, k\sigma_\phi^2 + 2 \sum_{i=1}^k \sum_{j=i+1}^k \text{cov}(\Delta\text{Lam}(i), \Delta\text{Lam}(j)) \right).$$

This additional covariance term increases the variance for increasing dependence between variables and reduces to zero in the case of independence.

As an estimate of this covariance term, we use the autocorrelation  $R_{\Delta\text{Lam}}$  of  $\Delta\text{Lam}$ , scaled by the variance  $\sigma_\phi^2$  of  $\phi$ . This results in the following corrected null distribution of the sum of  $k$   $\Delta\text{Lam}$  measurements:

$$N \left( k\mu_\phi, k\sigma_\phi^2 + 2 \sum_{i=1}^k \sum_{j=i+1}^k \sigma_\phi^2 R_{\Delta\text{Lam}}(j-i) \right), \quad (2)$$

where  $R_{\Delta\text{Lam}}(i)$  indicates the autocorrelation of  $\Delta\text{Lam}$  over a lag of  $i$  probes.

We test whether a gene  $g$  shows a significant change in LaminB1 interaction by comparing  $\Delta\text{Lam}_g$  against the null distribution defined in Equation 2. We obtain p-values for both tails of the distribution separately, i.e., for increased and reduced LaminB1 interaction. Genes significantly changing lamina interaction are selected after Benjamini-Hochberg multiple test correction, with an estimated FDR of 5%.

Additionally, we obtain a  $Z$ -score to be used for, e.g., GO analyses, defined as

$$Z = \frac{\Delta\text{Lam}_g - k\mu_\phi}{k\sigma_\phi^2 + 2 \sum_{i=1}^k \sum_{j=i+1}^k \sigma_\phi^2 R_{\Delta\text{Lam}}(j-i)},$$

which describes the  $\Delta\text{Lam}$  value for gene  $g$ , translated and scaled by the mean and variance of the null distribution (cf. Equation 2).

## **Expression data**

For all Affymetrix data, we computed Robust Multichip Average (RMA) expression values (Irizarry et al., 2003) using the affy R/Bioconductor-package (Gentleman et al., 2004). Differential expression analyses were done using the limma R/Bioconductor package (Smyth, 2004).

## **Preparation of gene list**

Gene annotation data were obtained from the NCBI FTP site (<ftp://ftp.ncbi.nlm.nih.gov>, downloaded March 2nd, 2009). We only analyzed genes that map to chromosomes 1-19 and X of the mouse reference genome (C57BL/6J, build NCBI m37). In case of overlapping genes, we retained the longest gene. We furthermore removed genes that overlap with fewer than 5 DamID probes, to allow for sufficient power in the statistical test for changes in NL interactions. This yielded a total of 17,266 genes for subsequent analyses.

## **Analysis of clustered genes**

P-values for the significance of the observed degree of clustering are obtained by randomizing the order of genes and intergenic regions 10,000 times (while keeping the interlaced structure of genes and intergenic regions) and counting the number of genes found to be clustered using the procedure described above. The number of clustered genes in real data is then compared to the distribution of counts in permuted data to estimate a p-value.

## **Gene Ontology Analysis**

As the basis for the GO analyses we use the list of 17,266 genes prepared earlier. Annotation data linking genes to GO categories have been obtained from the NCBI FTP site (downloaded March 2nd, 2009). The GO repository consists of a number of hierarchical trees, of which we use the 'Biological Processes' tree. Genes can be linked to multiple GO categories, at various levels

in the tree, but are typically annotated only at the lowest-available level. We therefore propagate annotations upwards through the tree, up to the top node. After this, we remove GO categories with five or fewer annotated genes, yielding a total of 2586 categories to be used for subsequent analyses. The total set of genes used for GO analyses consists of all genes that are annotated to at least one of these GO categories.

We perform Gene Ontology (GO) analyses by comparing a statistic for all genes in a GO category against a distribution of statistics based on the total set of genes. This shows a resemblance to previously reported methods (Efron and Tibshirani, 2007; Knijnenburg et al., 2008). In the case of  $\Delta$ Lam the used statistic is the  $Z$ -score, yielded by the test for changes in LaminB1 interaction, and in the case of  $\Delta$ Expr it is the  $t$ -statistic for differential expression, yielded by the *limma* R-package.

We define a GO category  $\Gamma \subset G$ , i.e., a subset of genes. For any such  $\Gamma$ , we calculate the sum statistic, i.e., either

$$S_Z = \sum_{\forall g \in \Gamma} Z(g) \text{ or } S_t = \sum_{\forall g \in \Gamma} t(g).$$

This sum statistic is then compared to a distribution given by  $N(m\mu, m\sigma^2)$  where  $\mu$  and  $\sigma^2$  are the mean and variance of either all  $Z$ -scores or all  $t$ -statistics and  $m$  is the number of genes in  $\Gamma$ . One-tailed tests yield separate p-values for enrichment and depletion. Significant GO categories are selected after Benjamini-Hochberg multiple test correction, with an estimated FDR of 5%.

To select neural-related GO categories, we selected GO categories with names containing one of the following keywords: "neur, nerv, axon, dendri, synap, brain or learn (158 out of 2586 categories). For cell cycle-related GO categories, we used the keywords "cell cycle", "M phase", "replication", "mitosis" or "division" (39 out of 2586 categories).

## GeneAtlas tissue type selection

We chose the following 10 tissue types from the GeneAtlas data (Lattin et al., 2008) to represent the set of central nervous system (CNS) tissue types (GEO accession numbers indicated between

parentheses):

- amygdala (GSM258617 & GSM258618)
- cerebellum (GSM258633 & GSM258634)
- cerebral cortex (GSM258635 & GSM258636)
- cerebral cortex prefrontal (GSM258637 & GSM258638)
- dorsal striatum (GSM258653 & GSM258654)
- hippocampus (GSM258671 & GSM258672)
- hypothalamus (GSM258673 & GSM258674)
- microglia (GSM258721 & GSM258722)
- nucleus accumbens (GSM258733 & GSM258734)
- olfactory bulb (GSM258735 & GSM258736)

As a control analysis, we selected another four datasets (GSM258655, GSM258656, GSM258657 & GSM258658) from the GeneAtlas data, representing (duplicate measurements of) two independent ESC cultures.

We used the remaining tissue types, minus two peripheral nervous system tissues ("dorsal root ganglia" and "spinal cord"), as a set of 77 non-neural tissues.

## References

Conti, L., Pollard, S.M., Gorba, T., Reitano, E., Toselli, M., Biella, G., Sun, Y., Sanzone, S., Ying, Q.L., Cattaneo, E., et al. (2005). *PLoS Biol* 3, e283.

Cremer, M., Muller, S., Kohler, D., Brero, A., and Solovei, I. (2007). *Cold Spring Harbor Protocols*, pdb.prot4723-.

Efron, B., and Tibshirani, R. (2007). *The Annals of Applied Statistics*, 1:107-129.

Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., et al. (2004). *Genome Biol* 5, R80.

Gräf, S., Nielsen, F.G., Kurtz, S., Huynen, M.A., Birney, E., Stunnenberg, H., and Flicek, P. (2007). *Bioinformatics* 23, i195-204.

Guelen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M.B., Talhout, W., Eussen, B.H., de Klein, A., Wessels, L., de Laat, W., et al. (2008). *Nature* 453, 948-951.

Irizarry, R.A., Bolstad, B.M., Collin, F., Cope, L.M., Hobbs, B., and Speed, T.P. (2003). *Nucleic Acids Res* 31, e15.

Knijnenburg, T.A., Wessels, L.F.A., and Reinders, M.J.T. (2008). *International Journal of Bioinformatics Research and Applications* 4, No.3 306-323.

Lattin, J.E., Schroder, K., Su, A.I., Walker, J.R., Zhang, J., Wiltshire, T., Saijo, K., Glass, C.K., Hume, D.A., Kellie, S., et al. (2008). *Immunome Res* 4, 5.

Smyth, G.K. (2004). *Stat Appl Genet Mol Biol* 3, Article3.

van Steensel, B., Brink, M., van der Meulen, K., van Binnendijk, E.P., Wansink, D.G., de Jong, L., de Kloet, E.R., and van Driel, R. (1995). *J Cell Sci* 108 ( Pt 9), 3003-3011.

Vogel, M.J., Peric-Hupkes, D., and van Steensel, B. (2007). *Nat Protoc* 2, 1467-1478.

Ying, Q.L., and Smith, A.G. (2003). *Methods Enzymol* 365, 327-341.

## Supplementary Tables Peric-Hupkes, Meuleman, et al.

	Number of LADs	Median size of LADs (bp)	Coverage of genome (%)
ESC	1180	454,011	40.22
NPC	1258	354,196	40.31
AC	1441	339,991	41.88
MEF	1189	534,499	44.39
Human fibroblasts (Guelen et al., 2008)	1344	552,876	42.76

**Table S1.** LAD statistics for all cell types.

**Table S2.** LAD definitions for ESC, NPC, AC and MEF.  
(Excel spreadsheet)

**Table S3.**  $\Delta\text{Lam}^{\text{down/up}}$  genes for ESC→NPC, NPC→AC and ESC→MEF transitions.  
(Excel spreadsheet)

**Table S4.** Gene Ontology enrichment analysis results for ESC→NPC and NPC→AC transitions.  
(Excel spreadsheet)

## Supplementary Figure legends Peric-Hupkes, Meuleman, *et al.*

**Figure S1.** (A, B) Examples of the localization of Dam-LaminB1 fusion protein transiently expressed in NIH3T3 fibroblasts and detected by immunofluorescent labeling of the V5 epitope tag that is present between Dam and Lamin B1 (green). DNA was counterstained by TOPRO3 (red). A single confocal microscopy section is shown in each panel. (C, D) Detection of NPC and AC marker proteins by immunofluorescence microscopy. (C) Expression of Nestin (green) in NPCs; 149 of 161 counted cells (92%) were Nestin positive. (D) Expression of the AC marker protein GFAP (green in ACs); 219 of 237 cells (92%) were GFAP positive. DNA was counterstained by DAPI (blue). (E, F, G) Hierarchical clustering dendrogram (complete linkage, Pearson correlation based distance measure) and heatmap showing Pearson correlation between replicate experiments for (E) unsmoothed data and (F) data smoothed using a running median function with a window size of 19 probes. The indicated clustering coefficient is the ratio between mean correlations of same and non-same cell-type replicates. Clustering dendrogram and heatmap for average of two replicates are shown in (G). (H) In situ hybridization in mouse fibroblast nuclei with fluorescently labeled adenine-methylated DNA fragments amplified from MEFs that expressed either Dam-LaminB1 or unfused Dam. DNA was stained with DAPI. A single confocal section is shown. Distribution of LaminB1 and Dam DNA was assessed on mid sections through fibroblast nuclei (n=30) acquired using a confocal microscope. Images were evaluated using ImageJ software (Colins, 2007). The border of the nuclei was defined using DAPI counterstain and two successive outer shells with the width of 640 nm (8 pixels) were defined. Mean signal intensities of LaminB1 and Dam signals were measured in each nucleus for both shells and for the remaining inner portion of the nucleus (i1, i2, i3). (I, J) Relations between the shell position and mean signal intensity are highly statistically significant ( $p < 0.001$ , Friedman test) and all shells are significantly different from one other ( $P < 0.05$ , paired Wilcoxon tests) for both LaminB1 (I) and Dam (J). Error bars show standard error around the mean. (K) LaminB1 DamID profiles for all chromosomes in ESCs (orange), NPCs (blue), ACs (magenta) and MEFs (green). Each track shows the average of two independent experiments.

**Figure S2.** Profiles of gene expression across LAD borders in ACs (magenta) and MEFs (green). Graph was constructed as in Figure 2b. Microarray expression data for MEFs are from (Mikkelsen et al., 2007) and for ACs from (Meissner et al., 2008).

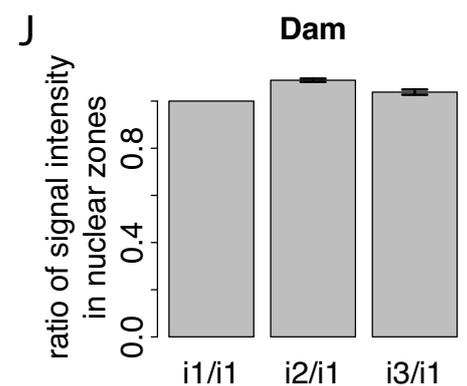
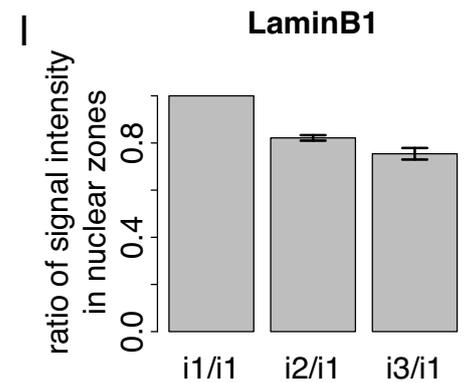
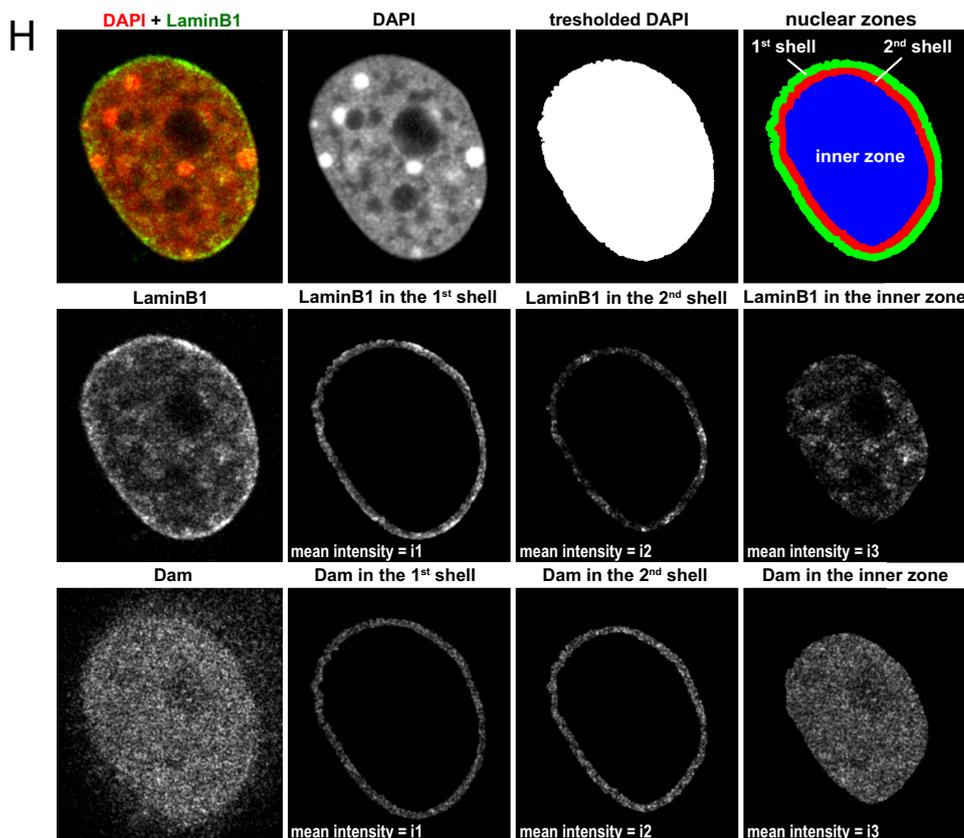
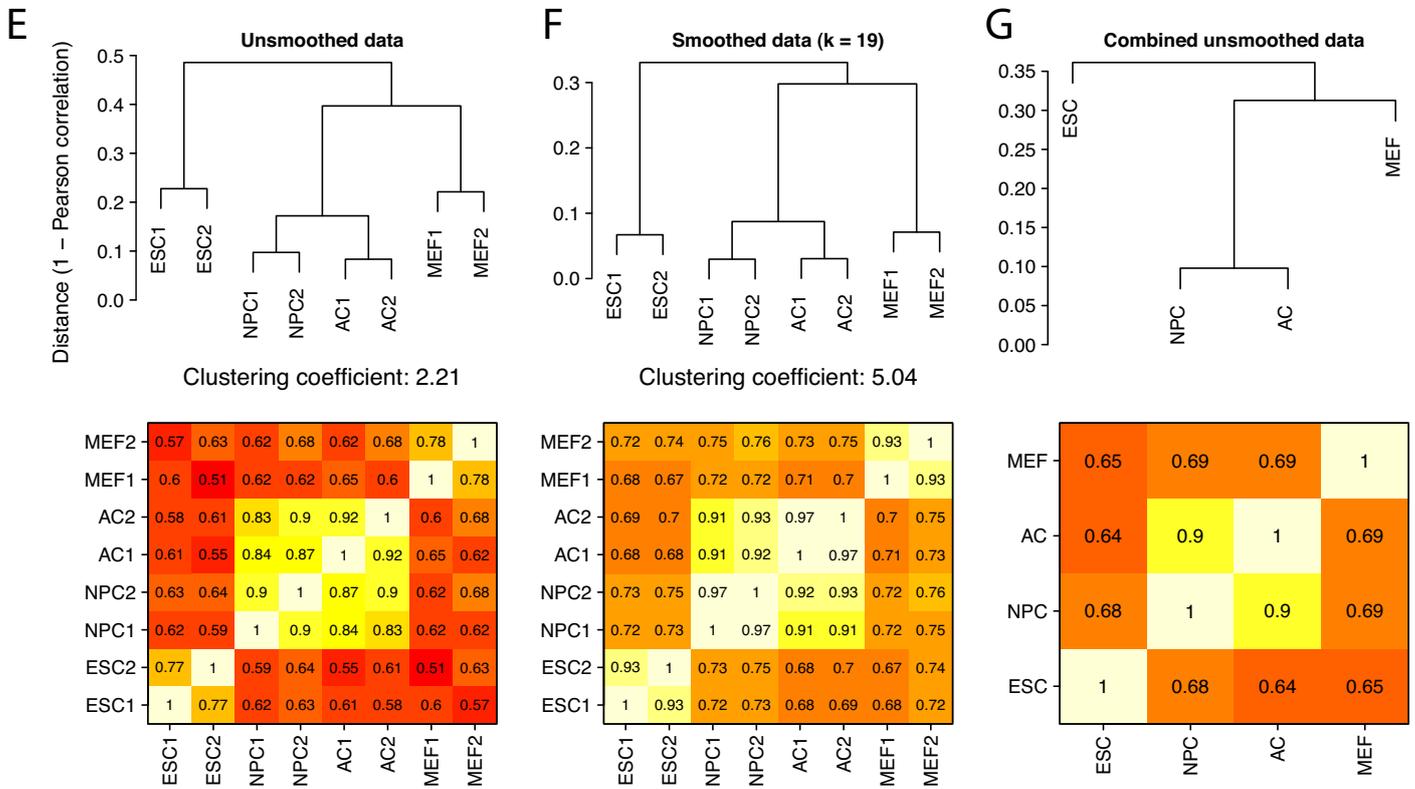
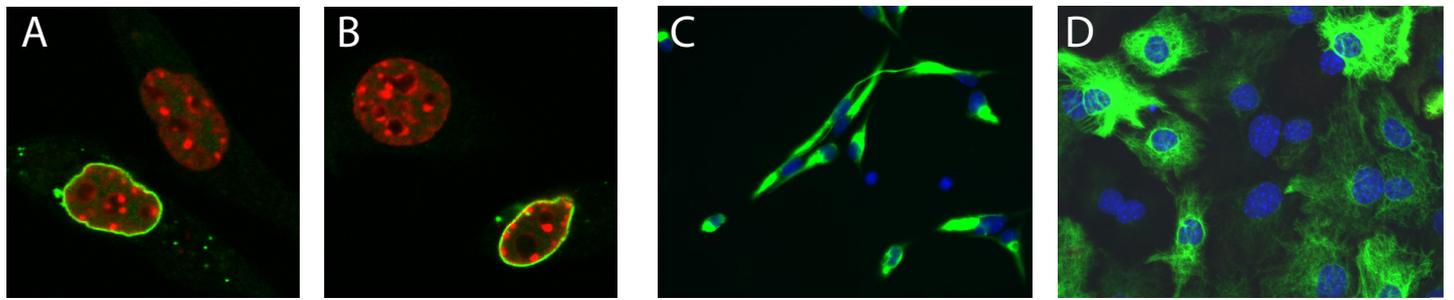
**Figure S3.** (A,B) Size distribution of relocating units during NPC→AC transition, calculated as the number of neighboring genes and inbetween intergenic regions with concordant significant decreases (A) or increases (B) in NL interaction levels. (C,D) Average profiles of the change in LaminB1-interaction along singleton  $\Delta\text{Lam}^{\text{down}}$  and  $\Delta\text{Lam}^{\text{up}}$  genes. (C,D) Grey areas mark estimated 95% confidence intervals.

**Figure S4.** Behaviour of ESC marker genes during ESC→AC and (hypothetical) ESC→MEF transitions. (A)  $\text{Log}_2$  changes in LaminB1 interaction ( $\Delta\text{Lam}$ ) and gene expression levels ( $\Delta\text{Expr}$ ) for the ESC→AC step. ESC specific (orange) (Takahashi and Yamanaka, 2006), neuron specific (blue) and housekeeping (red) genes. (B) Same analyses for ESC→MEF (note that MEFs are not directly derived from ESCs in our experiments, hence this comparison addresses a virtual transition between the cell types).

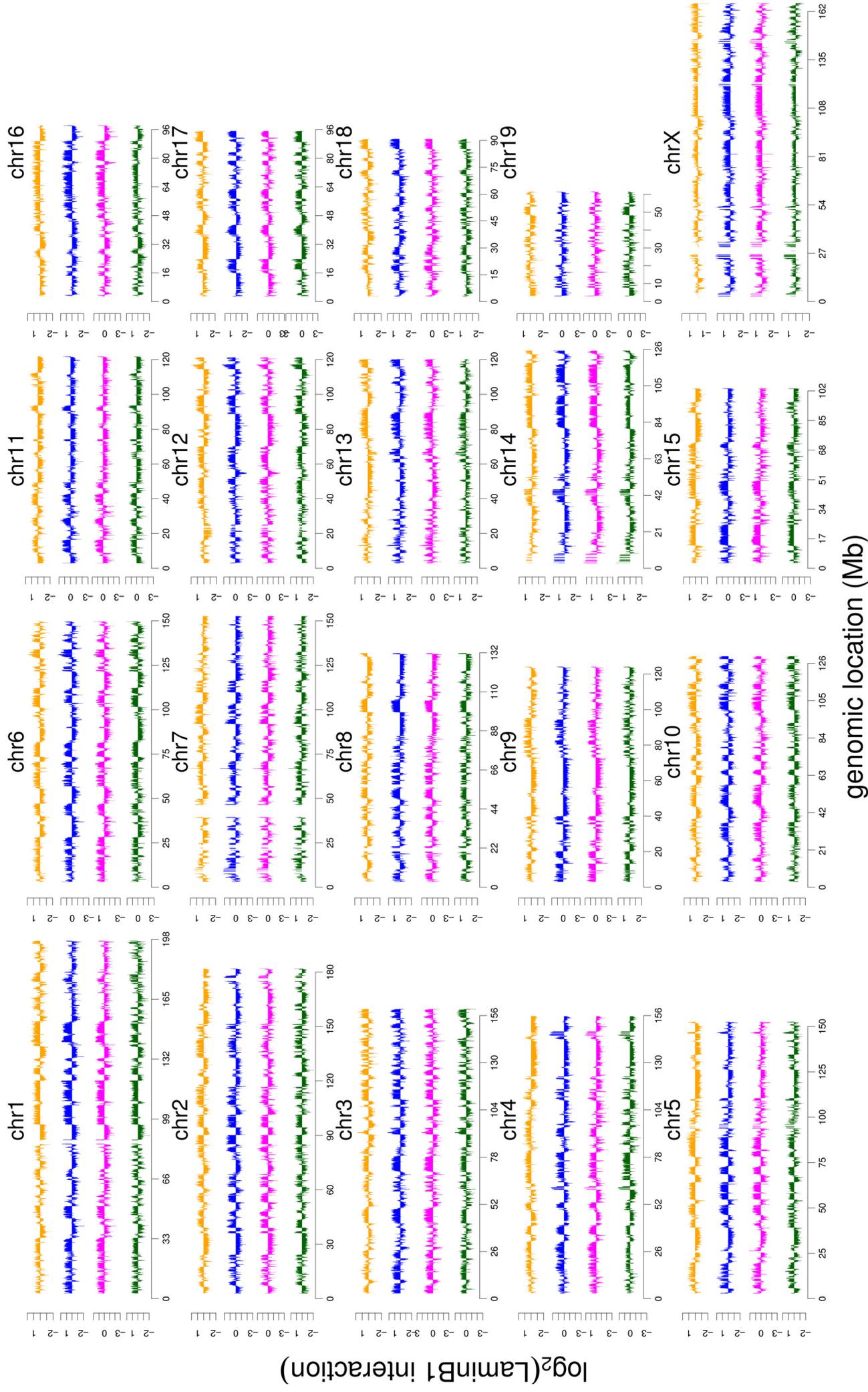
**Figure S5.** (A, B) Pol II levels as determined by CHIP (Mohn et al., 2008) at promoter regions of genes that detach from the NL ( $\Delta\text{Lam}^{\text{down}}$  genes) and either remain silent (A) or become active (B) in NPCs. Even though the  $\Delta\text{Lam}^{\text{down}}$  genes that remain silent are unlocked for activation at a later stage, they do not yet have detectable amounts of Pol II at their promoters in NPCs, ruling out a "polymerase poisoning" mechanism. (C, D) Comparison of  $\text{log}_2$  gene expression levels (measured with Affymetrix microarrays) to  $\text{log}_2$  H3K36me3 levels (average per gene, measured by CHIP-seq) scatterplots for ESCs (C) and NPCs (D). Density plots above and beside each graph visualize the bimodal distributions of gene expression and H3K36me3 values. Blue lines mark the local minima in each distribution. The lower quadrant demarcated by the blue lines was used to define silent genes. Colored dots highlight  $\Delta\text{Lam}^{\text{up}}$  (yellow) and  $\Delta\text{Lam}^{\text{down}}$  (brown) genes that are silent in both ESCs and NPCs and show no significant differential expression between the two cell types. Expression and H3K36me3 data are from (Mikkelsen et al., 2007). (E) Heatmap showing which of the  $\Delta\text{Lam}^{\text{down}}$  genes that are silent in both ESCs and NPCs become active (black box) in which tissue type. Red labels indicate neural tissues.

## REFERENCES

- Collins, T.J. (2007). ImageJ for microscopy. *Biotechniques* 43(1 Suppl), 25-30.
- Guelen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M.B., Talhout, W., Eussen, B.H., de Klein, A., Wessels, L., de Laat, W., *et al.* (2008). Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* 453, 948-951.
- Meissner, A., Mikkelsen, T.S., Gu, H., Wernig, M., Hanna, J., Sivachenko, A., Zhang, X., Bernstein, B.E., Nusbaum, C., Jaffe, D.B., *et al.* (2008). Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* 454, 766-770.
- Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.K., Koche, R.P., *et al.* (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* 448, 553-560.
- Mohn, F., and Schubeler, D. (2009). Genetics and epigenetics: stability and plasticity during cellular differentiation. *Trends Genet* 25, 129-136.
- Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126, 663-676.



Peric-Hupkes, Meuleman *et al.*, Supplementary Figure S1K



Peric-Hupkes, Meuleman *et al.*, Supplementary Figure S2

