SUPPLEMENTARY INFORMATION

CTCF-Mediated Functional Chromatin Interactome in Pluripotent Cells

Lusy Handoko1,*, Han Xu1,*, Guoliang Li1,*, Chew Yee Ngan1, Elaine Chew1, Marie Schnapp1, Charlie Wah Heng Lee1, Chaopeng Ye1, Joanne Lim Hui Ping, Fabianus Mulawadi1, Eleanor Wong1,2, Jianpeng Sheng3, Yubo Zhang1, Thompson Poh1, Chee Seng Chan1, Galih Kunarso4, Atif Shahab1, Guillaume Bourque1, Valere Cacheux-Rataboul1, Wing-Kin Sung1,3, Yijun Ruan1, Chia-Lin Wei1,2,#, Yijun Ruan

# Corresponding authors

Chia-Lin Wei
Tel: 1 (925) 927-2593
Email: cwei@lbl.gov

Yijun Ruan
Tel: (65) 68088073
Email: ruanyj@gis.a-star.edu.sg

* These authors contributed equally

# Corresponding authors

Yijun Ruan
Tel: (65) 68088073
Email: ruanyj@gis.a-star.edu.sg

# Corresponding authors

Yijun Ruan
Tel: (65) 68088073
Email: ruanyj@gis.a-star.edu.sg

Current address: Joint Genome Institute, Walnut Creek, California, U.S.A.

Keyword: insulator, enhancer, chromatin organization, epigenetic regulation, nuclear lamina
Supplementary Note

1. ChIA-PET analysis
   1.1 Determine cutoff fragment size between self ligation and inter ligation PETs
   1.2 Chimeric ChIA-PETs represent non-specific random inter-molecular ligation
   1.3 Reproducibility between biological and technical replicates of ChIA-PET libraries

Supplementary Figures

Supplementary Figure 1. CTCF ChIA-PET analysis
Supplementary Figure 2. Molecular and cytogenetic validation of CTCF directed inter- and intra-chromosomal interactions
Supplementary Figure 3. Clustering of inter-chromosomal interactions
Supplementary Figure 4. Specificity of chromatin domains defined by CTCF-mediated DNA looping
Supplementary Figure 5. Probability of loop span in each category and reduction of H3K4me1 signal intensities after CTCF knock-down
Supplementary Figure 6. Histone modification and RNAP II profiles in each chromatin domain and neighboring regions
Supplementary Figure 7. Examples of loops from category I-IV
Supplementary Figure 8. p300 association with cell specific open chromatin marks and enhancer signals
Supplementary Figure 9. DNA loop brings promoter and p300 enhancer into close proximity and affects expression of the corresponding genes
Supplementary Figure 10. Characterization of Lamin B-binding regions
Supplementary Figure 11. Model of CTCF directed chromatin domains (category I-IV), their associated gene activities and sub-nuclear chromatin localizations

Supplementary Tables

Supplementary Table 1. ChIA-PET sequencing processing and mapping summary
Supplementary Table 2. CTCF binding sites (Excel file)
Supplementary Table 3. Intra-and inter-chromosomal interactions detected by CTCF ChIA-PET (Excel file)
Supplementary Table 4. 4C, FISH and 3C validation results
Supplementary Table 5. List of 5 categories assigned to intra-chromosomal interactions (Excel File)

Supplementary Table 6. RNA Pol II, p300 and LADs sites defined by ChIP-Seq (Excel file)

Supplementary Table 7. ChIP-qPCR of CTCF, Lamin B, p300

Supplementary Table 8. RNAP II interactions defined by ChIA-PET (Excel File)

Supplementary Table 9. SALL4 interactions defined by ChIA-PET (Excel File)

Supplementary Table 10. Sequences of ChIA-PET linkers, siRNA, primers (qPCR, 3C, 4C) and list of BAC clones
Supplementary Note

1. ChIA-PET analysis

1.1. Determination of a self ligation and inter ligation cutoff.

In the ChIA-PET analysis, two types of ligation products were produced: the inter-molecular ligation events and self-ligation events. Self-ligation events mean that the two ends of the same DNA fragment are ligated together. We used the inter-molecular ligation PETs to identify interactions and self-ligation PETs to define binding sites. To determine the span cutoff between self-ligation and intra-chromosomal inter-molecular ligation PETs, we used the log-log plot analysis of the span distribution on the CTCF ChIA-PETs. A mixture model was observed with two straight distribution lines (see left); clearly representing two distinct PET populations. The size distribution of the self-ligating PETs follows a power-law distribution, which is a straight line on the left in the log-log plot. A different power-law distribution was observed for the intra-chromosomal interactions. Therefore, the span cutoff can be determined by the intersection between these two lines in the log-log plot which is between 5-10 Kb. Because the PETs around the calculated cutoff are a mixture of the self-ligation PETs and inter-ligation PETs, we used 10 Kb as cutoff size to select high confidence inter-ligating PETs.

1.2. Chimeric ChIA-PETs represent non-specific random inter-molecular ligation

The chimeric libraries generate no interaction cluster with FDR=0.05 and the interaction PETs identified shared little overlap between both biological and technical replicates. We further compared the span distributions of the intra-chromosomal PETs between chimeric and non-chimeric ChIA-PET libraries. The span of a PET is the genomic distance between the head and tail of each ChIA-PET. A log-log plot of the PET span distributions for PETs with spans from 10 Kb to 1 Mb from both chimeric and non-chimeric libraries is given below. The span
distribution from the non-chimeric library shows a good correlation ($R^2 = 0.94$) to a power-law distribution, while the span distribution from the chimeric library looks quite noisy. We also generated the span distribution with the random data by shuffling the heads and tails of the PETs. The span distribution from the random data looks like a constant. This confirms the quality of the non-chimeric ChIA-PET library and indicates that the ligation in the chimeric ChIA-PET library resulted from random chromatin ligation rather than the ligation of DNA molecules proximal to one another in the nucleus.

\[
y = 3\times10^8 x^{-1.13}, \quad R^2 = 0.973
\]

\[
y = 3920 x^{-0.37}, \quad R^2 = 0.733
\]
1.3. Reproducibility between biological and technical replicates of ChIA-PET libraries

To evaluate the complexity and dynamic status of interactions, CTCF ChIP material from mouse ES cells grown separately was used to construct biological (BR1 vs. BR2) and technical replicates (TR1 vs. TR2) (Supplementary Table 1). Plotting distribution of tags across the genome in each library against its matched replicate (see left) shows a high correlation for each replicate pair.

Among the reliable interaction clusters detected between TR1 (1,197) and TR2 (1,206), 906 (76%) clusters are shared and common. Among the reliable interaction clusters detected between BR1 (3,811) and BR2 (1,384), 533 (38%) are commonly found. Such significant degree of overlaps (empirical p-value < 1e-6) indicates that the interaction clusters defined are reliable. The larger difference between interactions detected in the biological replicates suggests that such chromatin interactions could be highly complex and dynamic among cell populations. To further determine whether the overlap between these replicates is specific and significant, we examined the level of reproducibility between the replicates of chimeric libraries. As expected, no significant interacting PET clusters were found in multiple chimeric ChIA-PET libraries even when the top 1000 interacting loci (chosen based on the FDR) were compared.

We next performed saturation analysis using the reproducibility of two biological replicates (BR1 and BR2). Based on the degree of overlapping interactions among biological replicates and assuming random sampling from the total pool by each replicate, we predict that there are 9,000-10,000 potential interactions in the ES cells. Since our cluster analysis identified ~ 5,384 PET clusters, we potentially have detected ~ 50% of the total CTCF-associated
chromatin interactions present in these cells. Upon close examination, the interactions identified are mostly involved the binding sites with higher peak intensities and thus likely to result from stronger interactions. However, the true level of comprehensiveness will probably beyond the current estimation with the improvement of the robustness of the detection method and many more bona fide but weaker or transient interactions will be found. The Venn diagram above shows the overlap between two sets of interactions. We next asked how many more sequences from BR1 were needed in order to capture most of the interactions found in BR2 To answer this question, we randomly selected 10%, 20%, …, 100% of the non-redundant reads from BR1, identified interactions using the criteria (PET2+, FDR<0.05), and counted the number of overlapping interactions in BR2 with different sequencing depths from BR1. A Hill-function was fitted to the data to estimate the level of saturation. Based on the fitted Hill function, 632 interactions in BR2 can be called if we retrieve infinite sequences from BR1. With the current sequencing depth of BR1, 533 out of 632 (84%) of the interactions are called. While only 45.7% (632 out of 1384) of the interactions in BR2 can be called even if we sequence BR1 infinitely. So, these numbers suggest that we have achieved over 80% saturation of the current library sequencing but only reached 45% saturation on the identifying all the potential interactions. This result suggests that the diversity among different biological replicates is large and library complexity is limited. Therefore, to achieve higher sensitivities, more biological replicates, rather than deeper sequencing on a single library, is needed.
Supplementary Figures

**Supplementary Figure 1.** CTCF ChIA-PET analysis

(a) Overview of CTCF ChIA-PET analysis.

Juxtaposed chromatin complexes are cross-linked and the interactions tethered by CTCF are enriched by chromatin immunoprecipitation. The resulted ChIP complexes are split into two reactions and ligated with a special linker A or B containing a MmeI restriction enzyme.
recognition site and biotin label. Intra-molecular ligation is then carried out to join the ends of proximal chromatin fragment. Three different types of ligation products are obtained: self-ligated DNA fragments (very left and very right), intra molecular ligations from chromatin fragments containing either A-A or B-B linkers (2nd to the right and left), and inter-molecular chimeric ligations obtained from two non specifically-interacting chromatin complexes carrying A-B linker (middle). Ligation products (self-circularization of individual DNA fragment or intra-molecular ligation of multiple DNA fragments within one interacting chromatin complex) are digested with MmeI to release PETs and selected by streptavidin conjugated magnetic beads. Ultrahigh throughput sequencing analysis is then performed to reveal long range chromatin interaction loci. The DNA fragments with either A-A or B-B linkers will be used as non-chimeric PETs to determine binding sites or chromatin interaction loci. The fragments with A-B linkers (chimeric PETs) are used as an indicator to estimate the level of noise in the ChIA-PET library. From the 10.1 million uniquely mapped inter-ligating PETs, we then defined the inter-ligation PET clusters. Using occurrence frequency to distinguish real interaction signals (multiple overlapping clusters) from random noise (PET singletons); we obtained 2,275 intra- and 3,109 inter-ligation PET clusters (Supplementary Table 3a, b). We further checked for homology between anchors and found that the majority of the sequences from the paired interaction anchor regions show no homology and thus are unlikely to have resulted from mapping errors of segmental duplications and homologous sequences. Furthermore, most of these interaction loci are supported by CTCF binding. 2,115 of the 2,275 intra-chromosomal (93%) and 2,648 of the 3,109 inter-chromosomal interactions (85%) harbor CTCF binding sites on either or both anchors of the interaction loci.
(b) CTCF binding sites defined by self-ligating ChIA-PETs
An example is shown here in a 92 Kb interval around chr2:106,149,842-106,241,509 surrounding the gene 4732421G10Rik. Two strong CTCF binding sites are detected by ChIP-Seq analysis (green track). Self-ligating ChIA-PETs shown as overlapping red connecting lines are also found to be accumulated at the same locations. The intensity profile resulting from these self-ligating ChIA-PETs is highly similar to the profile generated from ChIP-Seq.

(c) ChIP-qPCR validation of CTCF binding sites detected by ChIP-Seq. 21 CTCF binding sites were chosen based on the peak intensities (decreasing from the left to the right). Negative control regions (22-41) represent regions with no CTCF binding sites. The fold of enrichment is shown in the Y-axis for the list of 1-21 sites selected (Supplementary Table 7a).

(d) Correlation of binding intensities of CTCF and their involvement in the interactions. In total, 3,306 CTCF binding sites are involved in these chromatin interactions. Compared with the binding sites that are not involved in the interactions, these 3,306 sites have higher binding intensities (p<10E-308 in KS-test). The plot shows the ratio of binding sites found to anchor chromatin interactions ranked by the binding site intensities. 20% of the top 200 CTCF binding peaks are found in the interaction anchors while only 2% of the bottom 200 binding peaks are detected in the anchors.

(e) The distributions of peak heights for binding sites involved in interactions (blue) and those not involved in interactions (red). It appears that the detected interactions are mediated through the stronger binding events.
Supplementary Figure 2. Molecular and cytogenetic validation of CTCF directed inter- and intra-chromosomal interactions.

(a) Validation of intra-chromosomal interaction by 4C assay. In the Pcdhga and b locus (chr18:37792576-37895559), ChIA-PET detected 5 interactions. Using Protocadherin γ subfamily A12 promoter (chr18: 37,890,974–37,892,946) as a 4C bait region or anchor point (green triangle, ▲), 4 different intra-chromosomal interactions were detected. Among these 4 clusters, 2 confirmed the interactions detected by ChIA-PET (dashed circles). One interaction is detected between the Pcdhga12 promoter and the Pcdhgb1 promoter (chr18:37,806,000) approximately 85 kb upstream. The other interaction (~ 71 kb) is found between the promoter of Pcdhga12 and Pcdhga4 (chr18: 37,820,000). 4C also detected 2 more intra-chromosomal interactions anchored by CTCF binding which were missed by ChIA-PET analysis. One of them occurs between the anchor site, Pcdhga12 promoter, and the promoter of Pcdhga8 which are 40 kb in apart, while the other loop with 20 kb span connects the promoter of Pcdhga12 and of Pcdhga10 (see Supplementary Table 4a for a list of sites detected by 4C).
List of inter-chromosomal interaction validated by FISH (Fluorescent In Situ Hybridization)

To validate the inter-chromosomal interactions, we performed DNA FISH cytogenetic assays on the ES cells. The table shows fourteen inter-chromosomal interactions with cluster size ≥3 selected for FISH validation. 2 of them have CTCF binding site at one interaction anchor only.

Co-localization ratio distribution (fold change between fusion of two interacting loci and fusion of the control region) among all sites validated. As a negative control region, we randomly chose a region on chr16 (chr16:52,100,818-52,400,160) which is > 1 Mb in distance from any interaction site detected by ChIA-PET. 9 of 14 inter-chromosomal interactions have p-value <0.05, co-localization ratio > 1.5 and were considered successfully validated. Interactions 1-14 indicate the interactions listed in the table (b).
(d) Example of FISH images from a validated inter-chromosomal interaction. The interaction fusion event connects *Syne1* (chr10) and *Rnps1-Abca3* gene (chr17) loci. Top panel: location of the probes for FISH. Bottom panel: co-localization of signals from two interacting chromosomes.

(e) Validation of intra-chromosomal interaction by 3C-qPCR assay.
DNA looping mediated by CTCF was detected between Acbd4 gene and Hexim1 (chr11:102918935-102975295) (top panel). Chromatin from mouse ES cells was digested with EcoRI (EcoRI digestion map is shown in the middle panel, A, 1, 2, …,6 represent the region ~100 bp from the digested sites and were used to design the primers). As expected, the interaction formed between the anchor A (green triangle, ▲) and the nearest region (A-1) which is 4 kb downstream from the anchor was found to occur at high frequency. Interaction frequency decreased with the distance to the anchor region, but increased at the region where the loop was detected (A-5, interaction frequency of 0.6) (all 3C results can be found in Supplementary Table 4c).

(f) Validation of intra-chromosomal interaction by 3C-qPCR assay on CTCF knock down cells. DNA looping mediated by CTCF was detected between the promoter region of Efna2 and 3’ end of the Mim1 gene (top panel). In this 3C assay, chromatin from mouse ES cells was digested with HindIII (middle panel). In the control cells, we observe a high interaction frequency between A-5 where the loop was detected. In contrast, the interaction between A and 5 was reduced 3 fold in the CTCF knock down cell (relative interaction frequency in CTCF kd vs. control cells= 0.000073 vs. 0.00025) (Supplementary Table 4c). This result suggests that the DNA loop detected here was indeed CTCF-specific. We did however find an overall reduction of interaction frequency in the other ligated fragments A-1, A-2, A-4 and A-7. Since we used an independent locus on another chromosome for normalization (Ercc3, chr18), we could exclude the possibility that the changes resulted from technique variation between samples (digestion and ligation efficiency). Furthermore, FACS analysis on CTCF kd cells suggests that no changes in cell cycle or cell death were observed in the CTCF knock down cells, when compared to the control cells and untreated cells (data not shown). This further ruled out the possibility that the overall reduction of interaction frequency was due to cell cycle arrest or cell death.
Supplementary Figure 3. Clustering of inter-chromosomal interactions.

(a) Normalized inter-chromosomal interaction frequency matrix between different chromosome pairs. The normalized frequency is plotted as a heat map and the enriched pairs of high frequency interactions are displayed here as intense color regions. Significant enrichments are observed above background between specific chromosome pairs.

(b) Hierarchical clustering of each chromosome pair indicates the spatial relation between chromosomes. Two or more chromosomes are assumed to be spatially closer to each other if they interact more frequently. Height represents distance between chromosomes. The clustering shows that chromosomes 8, 15, 16, 18, which belong to the same subcluster, have more
interactions and are spatially closer to each other than to the other remaining chromosomes, as indicated by having the lowest distance/height.

(c) Interaction density matrix of each pair of chromosomes (as shown in a), from two biological replicates (BR1 and BR2).
Supplementary Figure 4

a  CTCF loops
Loop Span
random CTCF loops  RNAP II loops  SALL4 loops

b
Loops < 200 kb
H3K4me1
H3K36me3

Loops > 200 kb
H3K9me3
H3K20me3

p-value
CTCF vs. simulated loops  5.18E-05
CTCF vs. SALL4 loops  5.23E-05
CTCF vs. RNAP II loops  0.02

p-value
CTCF vs. simulated loops  1.14E-06
CTCF vs. SALL4 loops  2.20E-05
CTCF vs. RNAP II loops  2.18E-05

p-value
CTCF vs. simulated loops  2.7E-05
CTCF vs. SALL4 loops  0.02
CTCF vs. RNAP II loops  0.007

p-value
CTCF vs. simulated loops  5.36E-05
CTCF vs. SALL4 loops  0.004
CTCF vs. RNAP II loops  0.001
Supplementary Figure 4. Specificity of chromatin domains defined by CTCF-mediated DNA looping.

(a) Three different sets of control loops were used to determine whether the different chromatin domains determined from the clustering analysis are specific to CTCF. Histone profiling derived from CTCF-mediated loops (left), randomly simulated loops (second from the left), SALL4 loops (second from the right) and RNAP II loop (right). 1,622 of RNAP II-associated intra-chromosomal interactions with cluster size $\geq$ 5 (PET5+ RNAP II) (Supplementary Table 8), 1,636 SALL4-associated intra-chromosomal interaction loops (PET-4+ SALL4) (Supplementary Table 9) and simulated loops randomly paired by CTCF binding sites spanning between 10 Kb to 1 Mb were selected. Loops were sorted in ascending order of span, and we examined the histone pattern associated with different span. Each column corresponds to an aligned bin, and each row corresponds to a loop. A window containing 100 loops was moved vertically to average the signal. CTCF Loops with span < 200k are mostly active domains (indicated by H3K4me1, and to lesser degree H3K36me3 enrichments. CTCF loops with span > 200k has clearly different pattern (K9 and K20 me3). Loops > 200K are mostly repressive domains indicated by H3K9 and K20 me3 enrichment. As shown, the histone signal intensity patterns defined from CTCF interactions are unique to the CTCF and not found in RNAP II and SALL4-associated interactions. Furthermore, the loop span correlation, particularly around 200 Kb, is not observed in the RNAP II, SALL4 and simulated control interactions.

(b) Histone modification patterns within the chromatin domains are CTCF loop-specific. Top panel: in CTCF loops less than 200 Kb in size, K4me1 shows significant enrichment (left) and K36me3 shows significant depletion (right) relative to three different sets of control interactions (p-values are shown below each graph). Bottom panel: In the CTCF interactions > 200 Kb, H3K9 (left) & K20 me3 (right) are found to be significantly enriched inside of the loops compared with the signals found in other control interactions. Therefore, we conclude the histone modification patterns and chromatin domains uncovered here are unique to the CTCF-associated interactions.
**Supplementary Figure 5.** Reduction of H3K4m1 signal intensities within the loops after CTCF knock-down.
**(a)** Reduction of CTCF led to decrease in signal intensities of H3K4me1 within active domain defined by CTCF-associated DNA loop. The accumulated normalized intensity of H3K4me1 is plotted along the CTCF demarcated chromatin loops and their neighboring regions. The blue line represents the signals from the CTCF kd cells and the red line represents the signals from the control siRNA transfected cells.

**b)** Examples of the active domains with decreased H3K4me1 level in the CTCF kd cells. Top panel: DNA looping detected by CTCF, H3K4me1 signals represented by ChIP-seq tag count distribution are plotted in the middle (control cells) and in the bottom (CTCF kd cells) panels. The scale was normalized based on the sequencing depth.
Supplementary Figure 6. Histone modification and RNAP II profiles in each chromatin domain and neighboring regions.

Cumulative histone modification signals and RNAP II intensities within (center) and outside (to the upstream and downstream) of the CTCF demarcated loops (see model on the top) for each category. The X-axis shows the relative location of loops and the Y-axis shows the normalized cumulative intensities.
**Supplementary Figure 7**. Examples of loops from category I-IV.

Examples of loops found in category I-IV at genomic coordinates chr14:53,899,359-54,135,506, chr19:43,484,753-43,776,079, chr6:83,807,453-83,917,825 and chr7:99,844,767-99,896,702, respectively. The categories are labeled on the top left corner. The order of the tracks shown from the top is: genes, CTCF binding peaks, observed interactions and active histone marks (H3K4m1, m3, H3K36m3), RNAP II profiles and repressive histone marks (H3K27m3 and H3K9m3).
Supplementary Figure 8. p300 association with cell specific open chromatin marks and enhancer signals.
(a) ChIP-qPCR validation of p300 binding sites. The level of enrichment is shown. 21 p300 sites and 25 negative control regions were chosen (Supplementary Table 7b).

(b) ChIP-qPCR validation of Lamin associated domains (LADs). LADs were selected based on fold change. 16 of 17 LADs were successfully validated. Nine sites outside LAD were used as negative controls (Supplementary Table 7c).

(c) Genomic distribution of p300 binding relative to gene locations. The genome was divided into 4 distinct regions: proximal promoter (± 2.5 kb from well-annotated transcription start sites/TSS), distal promoter (a region lies between 2.5 kb – 20 kb upstream from TSS), gene body or intragenic region (2.5 kb downstream from TSS – 2.5 kb downstream from transcription stop sites) and intergenic region (>20 kb distal from TSS or transcription stop site). ~60% of the p300 sites are associated with gene regions, but largely (>80%) located distal from the proximal promoters. Only 16% of p300 binding sites occur in proximal promoter regions.

(d) FAIRE, an open chromatin indicator, signal intensities from ES cells are plotted ± 1Kb of ES specific p300 binding peaks. ES specific p300 sites were associated with cell specific FAIRE signals. p300 associated genomic regions exhibited cell specific open chromatin states as indicated by FAIRE signal.

(e) Venn diagrams of the overlaps between p300, H3K4me1 and me2 sites in ES cells. Majority of the p300 binding sites overlap with a subset of enhancer marks H3K4me1 and me2 marks. 70% (3,526/5,033) of ESC-p300 sites overlap with only 9% and 12% of the H3K4me1 sites found in ES cells. Similarly, only 8% of the H3K4me2 sites found overlap with 62% (3,127/5,033) of the p300 sites. Therefore, it appears that the repertoire of whole genome p300 sites only represents a subset of enhancers defined by H3K4 methylations. When dissecting which state(s) of H3K4 methylation best correlate with p300 occupancy, we found that the majority of the p300 sites overlap with cell specific H3K4me1&me2 co-modified regions. Out of 5,033 p300 sites, 3,837 (76%) overlap with either me1 or me2 and 2,816 (56%) overlap with regions modified by both H3K4me1&me2.

(f) The enrichment of p300 binding sites in active and enhancer loops

The # of p300 binding sites found per Mb among different categories of CTCF tethered intrachromosomal interaction loops. The genomic length of each different category of loop is normalized.
Supplementary Figure 9. DNA loop brings promoter and p300 enhancer into close proximity and affects expression of the corresponding genes. (a) Tmem170 (chr8:114,748,089-114,793,183) and (b) Crtac1 (chr19:42,298,499-42,553,831). Top panel: the associated genes and CTCF-associated DNA loops. Middle panel: RNAP II
Handoko L*, Han X*, Li G* et. al., Chromatin interactome in ES cells

binding was detected at the promoter and the p300 site in the normal cells. Reduced CTCF levels decreased the binding intensities of RNAP II at the p300 sites and the relative gene expression (right top panel). Bottom panel: the box shows a detailed view of RNAP II binding at p300 in the normal control and CTCF kd cells.
Supplementary Figure 10. Characterization of Lamin B-binding regions.

(a) An example of LADs within Chr1:132,356,654-164,713,307. DNA loops formed by CTCF (represented by CTCF loop track) are mostly found outside or between LADs (light blue track). As a comparison, LADs detected by the Dam ID technology are shown. Histone profiles were also shown. Overall, the genomic features of LADs determined here in ES cells using sequencing exhibit good agreements with the earlier analysis in human fibroblast cells and mouse ES cells.
using DamID technology; indicating that ChIP-Seq can result in equivalent resolution and should be feasible to apply for genome wide Lamin study in other cells.

(b) CTCF signals distribution across the LAD borders. Strong enrichment of CTCF signal is found at the borders of LADs.

(c) A Circos map of inter-chromosomal interactions among subcluster of chromosomes 8-15-16 and 18. The purple lines indicate the inter-chromosomal interactions and the color intensity is proportion to the cluster size. The orange bars depict the LADs and the green peaks show the p300 binding sites.

(d) Profiles of active histone modification marks, H3K4me1, H3K4m2, and H3K36me3 across LADs and neighboring regions. LADs are depleted of active histone marks. In particular, active chromatin signals, H3K4me1&2 marks, are mildly enriched in the LAD borders and then devoid within LADs; while the active transcription H3K36me3 mark is depleted sharply in the boundaries and the depletion is further extended inside the LADs. We did not observe any significant enrichment of the heterochromatin marks such as H4K20me3 and H3K9me3.
(e) LADs are enriched with repeats. 48.5% of LADs contain repeat sequences. Among the 53 known repeat families, the L1 repeats are significantly enriched within LADs (Z-score 24, 30% of LADs). The L1 repeat family is one of the largest and most common repeats in the genome. As a comparison, the repeat distribution in overall genome is shown. Repeat sequences enriched in LADs could also be involved in regulating the dynamics of transcription factor binding or gene regulation.
Supplementary Figure 11. Model of CTCF directed chromatin domains (category I-V), their associated gene activities and sub-nuclear chromatin localizations.
Handoko L*, Han X*, Li G* et al., Chromatin interactome in ES cells

Supplementary Tables

Supplementary Table 1

| Supplementary Table 1 | Biological replicate 1 | Biological replicate 2 | Combined Mega library |
|-----------------------|------------------------|------------------------|-----------------------|
|                       | Technical replicate 1  | Technical replicate 2  |                        |
|                       | BR1/BR2                | TR2                    | TR2                   |
| Chimeric (A-B)        | 14,749,779             | 23,203,374             | 19,071,841            |
| Non-chimeric (A-A, B-B) | 69,001,929             | 64,727,733             | 46,141,215            |
|                       | 3,715,267              | 10,179,880             | 5,828,688             |
|                       | 19.7%                  | 32.50%                 | 30.62%                |
| # PETs                |                        |                        |                       |
| # of unique PET       |                        |                        |                       |
| sequences             |                        |                        |                       |
| % chimerism           |                        |                        |                       |
| No alignment          | 207,846                | 386,944                | 1,381,506             |
| >10 mapping location  | 1,514,792              | 5,683,773              | 6,366,360             |
| 10 mapping location   | 199,120                | 8,634,624              | 5,102,014             |
| Uniquely mapped PETs | 1,434,300              | 6,192,480              | 3,560,601             |
| rescued from multiple | 168,365                | 881,346                | 455,110               |
| mapped PETs           |                        |                        | 1,167,091             |
| Total # PETs w/usable | 1,602,745              | 7,073,826              | 4,014,711             |
| mapping locations     |                        |                        | 8,440,438             |
| H-T* paired           | 168,779                | 3,665,541              | 1,313,641             |
| H-T* non-paired       | 1,495,966              | 3,488,285              | 2,728,070             |
| Define chromatin      |                        |                        |                       |
| interactions          |                        |                        |                       |
| Intra-chromosomal PETs| 249,123                | 304,655                | 197,355               |
| (clusters FDR < 5%)   | 1,246                  | 1,271                  | 562                   |
| w/ binding site support | 942 (70%)               | 961 (75%)              | 340 (56%)             |
| Inter-chromosomal PETs| 3,159,162              | 3,999,538              | 2,826,927             |
| (clusters FDR < 5%)   | 1,137                  | 1,307                  | 802                   |
| w/ binding site support | 255 (22%)               | 255 (19%)              | 183 (23%)             |

Uniquely mapped PETs: PETs that have unique mapping location when allowed either 0 mismatch or 1 mismatch
H-T*: Head and tail tag mapping locations were paired if they are within 10Kb in distance

32
## Supplementary Table 4

### a 4C Validation

**Protocadherin** gene locus

| cluster ID | chrom | start  | end    | strand | sequence ID                                                                 | length | match | mismatch |
|------------|-------|--------|--------|--------|------------------------------------------------------------------------------|--------|-------|----------|
| anchor site | chr16 | 37890074 | 37899240 | +      | GME006\_bulk\_454S\_GS001R003224SR\_GFYV3A01DE17X09                       | 77     | 77    | 0        |
| 1          | chr16 | 37890033 | 37890111 | +      | GME006\_bulk\_454S\_GS001R003224SR\_GFYV3A01CA1D0V                       | 130    | 124   | 2        |
| 2          | chr16 | 37820604 | 37820952 | -      | GME006\_bulk\_454S\_GS001R003224SR\_GFYV3A01E87L                       | 31     | 30    | 0        |
| 2          | chr16 | 37820712 | 37820736 | +      | GME006\_bulk\_454S\_GS001R003224SR\_GFYV3A01A87H                       | 27     | 26    | 1        |
| 2          | chr16 | 37820673 | 37820720 | +      | GME006\_bulk\_454S\_GS001R003224SR\_GFYV3A01E7C7E                       | 111    | 111   | 0        |
| 2          | chr16 | 37852500 | 37852574 | +      | GME006\_bulk\_454S\_GS001R003224SR\_GFYV3A01EQ10L                       | 92     | 67    | 3        |
| 3          | chr16 | 37873405 | 37873429 | -      | GME006\_bulk\_454S\_GS001R003224SR\_GFYV3A01CK18WQ                       | 28     | 23    | 1        |
| 3          | chr16 | 37873416 | 37873440 | -      | GME006\_bulk\_454S\_GS001R003224SR\_GFYV3A01D38TZ                       | 33     | 22    | 0        |

**Cyp2 gene locus**

| cluster ID | chrom | start  | end    | strand | sequence ID                                                                 | length | match | mismatch |
|------------|-------|--------|--------|--------|------------------------------------------------------------------------------|--------|-------|----------|
| anchor site | chr7  | 25625700 | 25626961 | +      | GME006\_454S\_GS001R003232SR\_GFHLHLRO134HTQ                        | 63     | 47    | 1        |
| 1          | chr17 | 25697076 | 25697126 | +      | GME006\_454S\_GS001R003232SR\_GFHLHLRO11G1QQC                        | 53     | 47    | 1        |
| 1          | chr17 | 25697145 | 25697184 | -      | GME006\_454S\_GS001R003232SR\_GFHLHLRO12H94T4                        | 51     | 48    | 2        |
| 1          | chr17 | 25697102 | 25697160 | -      | GME006\_454S\_GS001R003232SR\_GFHLHLRO12HERY2                       | 64     | 56    | 1        |
| 1          | chr17 | 25697108 | 25697151 | +      | GME006\_454S\_GS001R003232SR\_GFHLHLRO11G1Q4H                       | 48     | 41    | 1        |
| 1          | chr17 | 25697141 | 25697184 | -      | GME006\_454S\_GS001R003232SR\_GFHLHLRO11H9B74                       | 43     | 30    | 0        |
| 1          | chr17 | 25697117 | 25697256 | -      | GME006\_454S\_GS001R003232SR\_GFHLHLRO10GF2VS                       | 109    | 126   | 4        |
| 1          | chr17 | 25697120 | 25697154 | -      | GME006\_454S\_GS001R003232SR\_GFHLHLRO10F80X0                       | 34     | 34    | 0        |
| 1          | chr17 | 25697133 | 25697164 | -      | GME006\_454S\_GS001R003232SR\_GFHLHLRO12HCLV                       | 34     | 30    | 1        |
| 2          | chr17 | 26381452 | 26381504 | +      | GME006\_454S\_GS001R003232SR\_GFHLHLRO11G5DN2                       | 161    | 143   | 9        |
| 2          | chr17 | 26381461 | 26381620 | +      | GME006\_454S\_GS001R003232SR\_GFHLHLRO12H9G99                       | 43     | 35    | 0        |
| 2          | chr17 | 26381509 | 26381601 | +      | GME006\_454S\_GS001R003232SR\_GFHLHLRO12H7O5                       | 46     | 46    | 0        |
| 2          | chr17 | 26381562 | 26381567 | +      | GME006\_454S\_GS001R003232SR\_GFHLHLRO12H6O50                       | 36     | 33    | 1        |
| 2          | chr17 | 26381582 | 26381587 | +      | GME006\_454S\_GS001R003232SR\_GFHLHLRO12H7C5R                        | 27     | 25    | 0        |
| 2          | chr17 | 26381587 | 26381631 | +      | GME006\_454S\_GS001R003232SR\_GFHLHLRO12H7I18                       | 35     | 34    | 0        |
| 3          | chr17 | 26447705 | 26447736 | -      | GME006\_454S\_GS001R003232SR\_GFHLHLRO11GZN5Y                       | 35     | 30    | 0        |
| 3          | chr17 | 26447714 | 26447743 | -      | GME006\_454S\_GS001R003232SR\_GFHLHLRO12H7F9F                       | 35     | 35    | 0        |
| 4          | chr17 | 26575805 | 26575900 | -      | GME006\_454S\_GS001R003232SR\_GFHLHLRO12H2V94                       | 43     | 35    | 0        |
| 4          | chr17 | 26575876 | 26575905 | -      | GME006\_454S\_GS001R003232SR\_GFHLHLRO11GZCO3                       | 29     | 27    | 0        |
| 4          | chr17 | 26575918 | 26575944 | -      | GME006\_454S\_GS001R003232SR\_GFHLHLRO12HUXUZ                       | 37     | 26    | 0        |

*the red color: the sites identified by ChiA-PET*
b

**FISH validation on CTCF knock down ES cells**

| Interaction | CTCF siRNA | control siRNA | p-value | CTCF siRNA | control siRNA | p-value |
|-------------|------------|---------------|---------|------------|---------------|---------|
|            | experimental mix % colocalization | control mix % colocalization | fold change | experimental mix % colocalization | control mix % colocalization | fold change |
| chr13-chr15 | 5.69% | 4.62% | 1.23 | 2.37E-01 | 7.51% | 4.18% | 1.87 | 5.15E-08 |
| chr10-chr17 | 4.63% | 4.46% | 1.04 | 8.49E-01 | 9.01% | 9.27% | 1.63 | 1.47E-03 |
| chr3-chr14  | 3.62% | 3.15% | 1.15 | 5.79E-01 | 5.79% | 4.49% | 1.81 | 3.05E-03 |
| chr14-chr19 | 4.93% | 4.99% | 0.99 | 1.00E+00 | 4.77% | 2.93% | 1.63 | 3.20E-02 |

C

**3C validation**

chr11:102,806,377-103,010,000, cut by EcoRI

| Interaction | relative interaction frequency | st error |
|-------------|--------------------------------|----------|
| A-1(chr11:102918177-102921854) | 1.56 | 0.55 |
| chr12(chr11:102918177-102914053) | 0.62 | 0.09 |
| A-3(chr11:102918177-102950583) | 0.42 | 0.17 |
| A-4(chr11:102918177-102971753) | 0.22 | 0.09 |
| A-5(chr11:102918177-102972701) | 0.63 | 0.23 |
| A-6(chr11:102918177-103005836) | 0.19 | 0.04 |

chr10:79,564,519-79,700,518, cut by Hind III

| Interaction | normal mESC | control siRNA | CTCF siRNA |
|-------------|-------------|---------------|------------|
|              | relative interaction frequency | st error | relative interaction frequency | st error | relative interaction frequency | st error |
| A-1(chr10:79572904-79581195) | 6.2E-04 | 2.0E-04 | 4.0E-04 | 2.35E-04 | 2.5E-04 | 5.4E-05 |
| A-2(chr10:79581195-79615146) | 1.4E-04 | 3.3E-05 | 1.56E-04 | 5.91E-05 | 4.67E-05 | 1.11E-05 |
| A-3(chr10:79581195-79628450) | 3.6E-05 | 8.6E-06 | 1.66E-05 | 6.46E-06 | 8.87E-06 | 2.49E-06 |
| A-4(chr10:79581195-79630483) | 1.66E-04 | 3.3E-05 | 8.22E-05 | 5.23E-05 | 3.03E-05 | 0.34E-06 |
| A-5(chr10:79581195-79653030) | 2.20E-04 | 6.0E-05 | 2.50E-04 | 1.08E-04 | 7.30E-05 | 2.64E-05 |
| A-6(chr10:79581195-79675471) | 6.35E-05 | 1.4E-05 | 1.90E-05 | 5.18E-06 | 1.03E-05 | 2.30E-06 |
| A-7(chr10:79581195-79696000) | 6.47E-05 | 1.35E-05 | 8.03E-05 | 8.14E-05 | 2.84E-05 | 1.42E-05 |

A is the anchor region
A-5 is the interaction region found by ChIA-PET
Handoko L*, Han X*, Li G* et. al., Chromatin interactome in ES cells

Supplementary Table 7

a

CTCF binding sites validation by ChIP-qPCR

| Primer ID | peak location | peak intensity | intensity in control | local FDR | qPCR enrichment | SD |
|-----------|---------------|----------------|----------------------|-----------|-----------------|----|
| chr1:12700060-1270180 | 392.06 | 3.39 | 0.000042 | 154.35 | 1.51 |
| chr2:9810970-9816030 | 237.57 | 2.19 | 0.000042 | 123.67 | 3.64 |
| chr2:12061830-12061924 | 160.25 | 0.90 | 0.000042 | 121.14 | 4.75 |
| chr1:113001370-113061800 | 105.84 | 0.00 | 0.000042 | 153.81 | 0.75 |
| chr1:123852020-12506480 | 127.94 | 2.00 | 0.000042 | 67.90 | 8.30 |
| chr1:59024300-59024810 | 113.25 | 4.33 | 0.000115 | 121.66 | 8.34 |
| chr1:100069650-100070270 | 106.54 | 2.09 | 0.000046 | 62.68 | 0.31 |
| chr10:53476060-53477060 | 91.03 | 2.90 | 0.000075 | 50.68 | 4.06 |
| chr11:94381240-94381660 | 87.67 | 3.51 | 0.000266 | 32.85 | 3.22 |
| chr13:59702220-59720060 | 76.25 | 3.76 | 0.000704 | 24.70 | 0.24 |
| chr11:115758050-115758470 | 70.40 | 0.87 | 0.000044 | 31.28 | 2.30 |
| chr16:78174770-78175260 | 67.25 | 1.33 | 0.000064 | 46.45 | 3.67 |
| chr16:85789450-85790607 | 57.26 | 2.26 | 0.000321 | 65.79 | 11.34 |
| chr17:78149960-78150420 | 54.25 | 2.73 | 0.000993 | 21.98 | 2.04 |
| chr12:67326260-6722030 | 47.71 | 4.70 | 0.013118 | 31.24 | 0.62 |
| chr11:100073330-100073750 | 41.77 | 4.49 | 0.020969 | 37.04 | 4.71 |
| chr10:27765450-27766890 | 38.94 | 2.03 | 0.002294 | 31.25 | 1.99 |
| chr19:120784026-1207844150 | 38.02 | 1.82 | 0.001576 | 17.64 | 0.69 |
| chr11:34375170-34375620 | 31.13 | 2.84 | 0.014119 | 26.37 | 1.67 |
| chr12:421240-4212450 | 27.76 | 0.85 | 0.002293 | 9.42 | 0.14 |
| chr19:1974170-16742120 | 10.39 | 0.81 | 0.022148 | 26.84 | 0.42 |
| chr13:23690760-23691420 | negative control | negative control | 7.21 | 0.21 |
| chr15:87378500-87378920 | negative control | negative control | 4.95 | 0.36 |
| chr10:35344060-35344550 | negative control | negative control | 3.27 | 0.38 |
| chr16:87353560-87354050 | negative control | negative control | 1.19 | 0.28 |
| chr16:82591720-82592190 | negative control | negative control | 1.03 | 0.01 |
| chr19:49892020-49896700 | negative control | negative control | 1.04 | 0.11 |
| chr16:64900850-64901150 | negative control | negative control | 1.10 | 0.20 |
| chr11:77302970-77303450 | negative control | negative control | 0.97 | 0.12 |
| chr18:63680870-63681150 | negative control | negative control | 1.22 | 0.07 |
| chr12:93181330-93181630 | negative control | negative control | 1.99 | 0.47 |
| chr11:107550150-107550600 | negative control | negative control | 3.15 | 0.03 |
| chr11:119690100-119690580 | negative control | negative control | 1.23 | 0.11 |
| chr11:119870930-119871420 | negative control | negative control | 1.45 | 0.21 |
| chr16:71506030-71561100 | negative control | negative control | 1.13 | 0.27 |
| chr13:15303810-15304300 | negative control | negative control | 1.23 | 0.14 |
| chr16:9102790-91063260 | negative control | negative control | 0.89 | 0.19 |
| chr13:60081020-60081550 | negative control | negative control | 0.73 | 0.05 |
| chr10:89024600-89025120 | negative control | negative control | 1.39 | 0.18 |
| chr16:89962250-89962760 | negative control | negative control | 1.21 | 0.32 |
| chr11:74948370-74948860 | negative control | negative control | 1.45 | 0.16 |
Handoko L*, Han X*, Li G* et al., Chromatin interactome in ES cells

b

p300 ChIP-Seq validation by ChIP-qPCR

| Primer ID | peak location | peak intensity | intensity in control | local FDR | qPCR enrichment | SD |
|-----------|---------------|----------------|-----------------------|-----------|-----------------|----|
| chr4:48419470-48820010 | 143.696322 | 2.32696933 | 0.002178 | 63.36237233 | 2.483823 |
| chr7:2609180-2695050 | 16.430279 | 0.98641431 | 0.00701645 | 2.680450153 | 2.496365 |
| chr7:132953210-132953750 | 84.486506 | 0.52191229 | 0.002178 | 50.92875618 | 2.624846 |
| chr8:30187670-30188820 | 73.262048 | 2.2749 | 0.002172 | 25.5690169 | 1.252312 |
| chr5:135221280-135221850 | 17.84606 | 3.183267 | 0.00245552 | 1.65963514 | 0.081311 |
| chr5:103964860-103965290 | 76.011952 | 0.696414 | 0.002178 | 33.018616 | 0.089093 |
| chrd:57784800-57785270 | 165.310757 | 2.940239 | 0.002178 | 31.906549 | 3.588678 |
| chrd:82288110-82292490 | 598.713147 | 57.410359 | 0.0044327 | 13.7369722 | 2.278647 |
| chrd:68652010-68653530 | 85.330677 | 2.876494 | 0.002172 | 11.2574286 | 0.00188 |
| chrd:30951060-30951690 | 54.330677 | 1.087649 | 0.002178 | 30.1073583 | 0.436911 |
| chrd:262743500-262743800 | 72.625498 | 2.61753 | 0.0022244 | 14.7784687 | 0.586935 |
| chrd:183026410-3026590 | 72.282869 | 1.406375 | 0.002178 | 53.9001328 | 4.222565 |
| chrd:174060540-45065880 | 70.76494 | 4.119522 | 0.0025288 | 20.8515962 | 7.782903 |
| chrd:1695875920-95876450 | 62.095618 | 5.808797 | 0.0044124 | 13.5566822 | 0.664156 |
| chrd:169695650-96956890 | 90.25498 | 1.450199 | 0.002178 | 16.22985772 | 0.636204 |
| chrd:127897030-47391280 | 187.25996 | 4.769824 | 0.002178 | 84.4809672 | 3.311678 |
| chrd:1274904660-7405270 | 64.298058 | 2.697211 | 0.0022208 | 13.9298085 | 0.4773913 |
| chrd:118469290-38498120 | 80.7251 | 1.796813 | 0.002178 | 12.8813950 | 1.009132 |
| chrd:116276330-182761800 | 182.406375 | 4 | 0.002178 | 34.5084645 | 5.672415 |
| chrd:116272300-182725880 | 132.557766 | 4.869414 | 0.0022168 | 32.5864957 | 5.241262 |
| chrd:115577600-155678400 | 47.796813 | 4.10952 | 0.0066122 | 18.4454005 | 4.5254959 |
| chrd:31974080-31974600 | negative control | negative control | negative control | negative control | negative control | negative control |
| chrd:1197634550-97635860 | negative control | negative control | negative control | negative control | negative control | negative control |
| chrd:517562430-51653020 | negative control | negative control | negative control | negative control | negative control | negative control |
| chrd:541893490-108938180 | negative control | negative control | negative control | negative control | negative control | negative control |
| chrd:541251126-125211240 | negative control | negative control | negative control | negative control | negative control | negative control |
| chrd:115616103-161449383 | negative control | negative control | negative control | negative control | negative control | negative control |
| chrd:119641345-334136436 | negative control | negative control | negative control | negative control | negative control | negative control |
| chrd:7947321656-72322183 | negative control | negative control | negative control | negative control | negative control | negative control |
| chrd:109364940-36490323 | negative control | negative control | negative control | negative control | negative control | negative control |
| chrd:1278460920-74660510 | negative control | negative control | negative control | negative control | negative control | negative control |
| chrd:2181720540-181725580 | negative control | negative control | negative control | negative control | negative control | negative control |
| chrd:107705365-70572795 | negative control | negative control | negative control | negative control | negative control | negative control |
| chrd:688521887-85822126 | negative control | negative control | negative control | negative control | negative control | negative control |
| chrd:92798086-92798367 | negative control | negative control | negative control | negative control | negative control | negative control |
| chrd:1085272870-82573203 | negative control | negative control | negative control | negative control | negative control | negative control |
| chrd:1632975890-32976320 | negative control | negative control | negative control | negative control | negative control | negative control |
| chrd:1641300264-13012724 | negative control | negative control | negative control | negative control | negative control | negative control |
| chrd:1053641403-46414688 | negative control | negative control | negative control | negative control | negative control | negative control |
| chrd:1731100130-71105570 | negative control | negative control | negative control | negative control | negative control | negative control |
| chrd:65573003-55374217 | negative control | negative control | negative control | negative control | negative control | negative control |
| chrd:47495360-474953840 | negative control | negative control | negative control | negative control | negative control | negative control |
| chrd:156273620-92733220 | negative control | negative control | negative control | negative control | negative control | negative control |
| chrd:1248152-134481526 | negative control | negative control | negative control | negative control | negative control | negative control |
| chrd:1813500820-350097010 | negative control | negative control | negative control | negative control | negative control | negative control |
| chrd:1332621080-132621540 | negative control | negative control | negative control | negative control | negative control | negative control |
**c**

**Lamin B ChIP-Seq validation by ChIP-qPCR**

| Primer ID  | Lamin region                  | local FDR | qPCR enrichment | SD  |
|------------|--------------------------------|-----------|-----------------|-----|
| ESC-LAD-Primer1 | chr6:133157900-133686560 | 0.10101   | 13.11           | 1.03|
| ESC-LAD-Primer2 | chrX:165031880-165334980   | 0         | 0.91            | 0.15|
| ESC-LAD-Primer3 | chr14:105364250-106320960  | 0         | 10.14           | 0.70|
| ESC-LAD-Primer4 | chr2:140119310-140300050   | 0         | 11.20           | 2.03|
| ESC-LAD-Primer5 | chr17:93808780-93932800    | 0         | 10.77           | 2.15|
| ESC-LAD-Primer6 | chr14:109788790-110546270  | 0         | 8.20            | 0.52|
| ESC-LAD-Primer7 | chr16:72167190-72472270    | 0         | 12.49           | 4.49|
| ESC-LAD-Primer8 | chr8:52896820-53033740     | 0.01005   | 7.91            | 0.85|
| ESC-LAD-Primer9 | chr12:101749700-102015280  | 0.199005  | 11.35           | 0.28|
| ESC-LAD-Primer10 | chr11:111036970-111345490 | 0.017699  | 3.55            | 1.47|
| ESC-LAD-Primer11 | chr13:86146570-86383580   | 0         | 9.11            | 1.34|
| ESC-LAD-Primer12 | chr2:116978130-117508070  | 0         | 17.72           | 1.47|
| ESC-LAD-Primer13 | chr5:67507050-87671780    | 0         | 8.30            | 2.09|
| ESC-LAD-Primer14 | chr1:124499710-124748270  | 0         | 9.62            | 0.33|
| ESC-LAD-Primer15 | chr9:76357860-76654150    | 0.01005   | 8.60            | 0.93|
| ESC-LAD-Primer16 | chr7:88619710-88708280    | 0.01005   | 8.67            | 1.53|
| ESC-LAD-Primer17 | chr4:89295750-89596550    | 0.017699  | 7.29            | 2.82|
| Lamin-negative region1 | chr19:47382500-47387500 | 0.780477  | 0.60            | 0.27|
| Lamin-negative region2 | chr19:53045000-53050000 | 0.780477  | 0.57            | 0.03|
| Lamin-negative region3 | chr19:61277500-61282500 | 0.780477  | 0.95            | 0.25|
| Lamin-negative region4 | chr19:43780000-43785000 | 0.780477  | 0.67            | 0.06|
| Lamin-negative region5 | chr19:50800000-50850000 | 0.780477  | 0.69            | 0.09|
| Lamin-negative region6 | chr16:17440000-17445000 | 0.780477  | 0.72            | 0.17|
| Lamin-negative region7 | chr15:76972500-76977500 | 0.780477  | 0.77            | 0.02|
| Lamin-negative region8 | chr14:121390000-121395000 | 0.780477 | 1.04            | 0.25|
| Lamin-negative region9 | chr10:76642500-76647500 | 0.780477  | 1.03            | 0.23|
**Supplementary Table 10**

Sequences of ChIA-PET linkers, CTCF siRNA, Primers for ChIP-Seq, 4C and 3C validations, BAC clones

### ChIA-PET linker and siRNA sequences

| Name          | Sequence                        |
|---------------|---------------------------------|
| ChIA-PET linkers |                                |
| AA Linker     | GTTGGATCGGATACCGCGG CGCGATACGGGATCCAAAC |
| BB Linker     | GTTGGATCGGATACCGCGG CGCGATACGGGATCCAAAC |
| AB Linker     | GTTGGATCGGATACCGCGG CGCGATACGGGATCCAAAC |
| CTCF siRNA    | CUGUGUUAUAGAGACG | GUGUACAUAAAGUCGCUCA |
| (SMART Pool,  |                                  |
| Dhharmacon)   | GCUAUAAACAUACUGAAGCC | CCAACAUACUGAAGACGA |

### BAC clones for FISH validation

| Site 1        | BAC         | Site 2        | BAC         |
|---------------|-------------|---------------|-------------|
| chr1:175171487-175178287 | RP23-370C6   | chr9:40161665-40168708 | RP23-114G13 |
| chr10:4790220-4793536    | RP24-459C3   | chr10:24137721-24144513 | RP24-162G18 |
| chr13:13656736-136626929 | RP24-423F5   | chr15:74601084-74916772 | RP24-456G86 |
| chr14:25135306-25138072   | RP24-92A2    | chr19:5568002-5570147  | RP23-389K5  |
| chr16:13096489-13096078   | RP23-314Q21  | chr19:61316108-61321184 | RP24-318N5  |
| chr18:13096489-13096078   | RP23-314Q21  | chr19:106941688-106943794 | RP23-109G14 |
| chr3:122629570-122631866  | RP23-1A12    | chr13:100038281-100038753 | RP23-467G99 |
| chr3:153608586-153614193   | RP24-186A10  | chr9:198783230-198787800 | RP24-233B16 |
| chr3:58532869-58542425     | RP23-413C3   | chr14:44693459-44696310 | RP24-289I17 |
| chr4:133811217-133818925   | RP24-132E8   | chr8:81951426-81922213 | RP23-134A7  |
| chr5:147421759-147425653   | RP23-129N7   | chr17:25220457-25226803 | RP23-42123  |
| chr6:43734276-43736441     | RP23-181A6   | chr17:29526735-29526841 | RP23-447G21 |
| chr12:87868837-87883388    | RP23-205D14  | chr17:28985753-28988018 | RP24-479I12 |
| chrX:161037281-161038515   | RP23-239H22  | chr11:17430005-17450052 | RP23-299L18 |
| chr16:52100,518-52,400,16B(control) | RP24-540H15  |                                |             |

### BAC clones for 3C validation

| 3C          | BAC clone | Location          |
|-------------|-----------|-------------------|
| control BAC (Eroe3 locus) | RP23-148C24 | chr18:32,375,680-32,420,011 |
| EcoRI3C_Chr11 | RP23-358E19 | chr11:102,804,615-103,037,611 |
| HindIII 3C_Chr10 | RP24-488O12 | chr10:79515657-79702617 |
Handoko L*, Han X*, Li G* et al., Chromatin interactome in ES cells

| Locus   | Primer (5’-3’)                  | Left primer (5’-3’)                  |
|---------|---------------------------------|--------------------------------------|
| Cyp2    | Inverse: CCAAGCTGAAATCCAGCTCCACG | GTAATTCGACTAAAGCACTTTATGGA           |
|         | Nested: CACACTTATCTTGACCAGTCAAGA | GAAATCAGCTCTTCTTTGAGCCGAGCAGAGCAGA  |
| Podgamma12 | Inverse: CTTTCATCCGCCTAATAACA  | CAATACAGAGAGAGAGCAGAGAGAGAGAGAGA   |
|         | Nested: CACACTTATCTTGACCAGTCAAGA | GAAATCAGCTCTTCTTTGAGCCGAGCAGAGA   |

The underlined sequences are the GSFLX 454 adapter.

**Primer for 4C validation**

| Interaction | Interaction site (5’-3’) | Anchor site (5’-3’) | Note                      |
|-------------|--------------------------|---------------------|---------------------------|
| A-1         | TGGTTAGGACACACACTATGG    | TGCACCTGGAGACAGAACG | chr10, Hind III, control primers |
| A-2         | GCCACAGCTAGGAGACACTGAG   | TGCACCTGGAGACAGAACG | chr10, Hind III, control primers |
| A-3         | AGATCACGTGGCTAGACACAGAC | TGCACCTGGAGACAGAACG | chr10, Hind III, control primers |
| A-4         | CACATGCGATAGTCAGTGAAA    | TGCACCTGGAGACAGAACG | chr10, Hind III, control primers |
| A-5         | AGACGCTCTAGCAGACACTGTC   | CTGAGACCTGGAGACAGAAC | chr10, Hind III, the interacting fragment* |
| A-6         | ATCGAGCAGATTTCCTTCTTGG   | TCTTGGCTTGACAGCTTGTCT | chr10, Hind III, control primers |
| A-7         | AAGACGAAGAATCTTGGGCTTCT  | CTTTGGCTCTGCTGATG   | chr10, Hind III, control primers |
| Ercc3 control interaction1 | TCTGCTCTGCTGCTTGTAGTT | MGGCCCTACTCCAGACGAGT | chr10, Hind III, Ercc3 fragments from Ercc3 |
| Ercc3 control interaction2 | AGATGACCTCTGCTGTAGT   | CCAAGGTGTACTGTAAGAGAGCAGA | chr18, Hind III, Ercc3 fragments from Ercc3 |
| Loading control | TCTGCGCTCTACCTCCTGCT   | CTTTGGCTCTGCTGATG   | chr18, Ercc3 gene locus, HindIII |
| A-1         | AGACGACAGGAGAGAGACAGCTA | CCGAGGATTAAATAGTGGAGC | chr11, Ercc3, control primers |
| A-2         | ACCAGCGTCTGCTATACCC     | CCGAGGATTAAATAGTGGAGC | chr11, Ercc3, control primers |
| A-3         | AACACATACCTGAGAAGGCTTTT | CCGAGGATTAAATAGTGGAGC | chr11, Ercc3, control primers |
| A-4         | GTTATGCTTGTGATTAGCTGA   | CCGAGGATTAAATAGTGGAGC | chr11, Ercc3, control primers |
| A-5         | CCTGTTAGACAGCTTCTGGG    | CCGAGGATTAAATAGTGGAGC | chr11, Ercc3, the interacting fragment* |
| A-6         | TCACTTGCCTGCTCGACTCAT   | CCGAGGATTAAATAGTGGAGC | chr11, Ercc3, control primers |
| Ercc3 control interaction1 | CTCTTGCACCTGCTTGGTGGAGC | GTAGTCACTCCGTGGTGGTGGAGC | chr18, Ercc3, Ercc3 fragments from Ercc3 |
| Ercc3 control interaction2 | TCAAGGAGAGAGAGTGGATTTGA | TCTTGGCTCTGCTGATGCT | chr18, Ercc3, Ercc3 fragments from Ercc3 |

* the interacting fragment (A-5) is the interaction detected by ChIA-PET

**Primer for gene expression analysis**

| Locus  | Forward sequences (5’-3’)       | Reverse sequences (5’-3’)          |
|--------|---------------------------------|-----------------------------------|
| GAPDH  | tggcctgctctggaggagactgac     | ccttctggcctgccctgctcttg          |
| beta-Actn | ggtgagctgcagcagctgctcggc    | ccacagacagcgacagcgacaggg          |
| Tmex170 | ttgctgccctgcctctgctctg      | ccttctggcctgccctgctcttg          |
| Cntct  | ctgcggctggagagctctctctct   | gcacagctgcagcagctgctcttcctggta   |
| Dks8   | agcgcgctggagagctctctctct   | gcacagctgcagcagctgctcttcctggta   |
| Lsm2   | tggcctgctctggaggagactgac     | ccttctggcctgccctgctcttg          |
| RAB27A | ctgcggctggagagctctctctct   | gcacagctgcagcagctgctcttcctggta   |
### Primers for CTCF binding sites validation

| Primer ID | peak location | forward sequence | reverse sequence |
|-----------|---------------|------------------|-----------------|
| 1         | chr13:12760080-12761880 | agacagaagcaacgctgtf | aggagccccctgtacccag | |
| 2         | chr12:88168970-88169390 | gcggccctgagctgtgctg | ccacagtctgccccgtgctg | |
| 3         | chr12:20618030-20619240 | aaccatgacccctagagcagctgtg | acctccgaggaacaccttcgctg | |
| 4         | chr11:115801370-115801800 | cttacatctccacccacacc | ctacatagggcaggtgctlctt | |
| 5         | chr12:42005020-42005440 | gcgcccaagactagtctctacttc | ttccttgagacagcagccag | |
| 6         | chr15:36324390-36324610 | aggcaccagggaggtgcgggc | gcccttcctgtctcccttca | |
| 7         | chr15:100060850-100070270 | tccacacctctctctccctca | agcttttaagccagtttgctgta | |
| 8         | chr19:53476400-53477060 | gcaccagctgagccgttgcag | tgcagggacaggttctgctct | |
| 9         | chr11:94301240-94301660 | ctgctccacccacagccactc | gcgtcttcacagctgcttga | |
| 10        | chr13:59202020-59202060 | lgtggtcgctcctgggtctcag | cccgtaagagctgtccacctc | |
| 11        | chr11:115780500-115784720 | gtgtccgctctgggtgctg | ctgccacagttgctgctgtg | |
| 12        | chr16:176174770-176175200 | gtcctcagccagtctggaaggg | cccctctctaggggtgcctg | |
| 13        | chr16:85788450-85788870 | gcagcttgctggagctgtcag | cagctctggtgagccgctgta | |
| 14        | chr10:78149980-78150420 | ggcctcgctgctggtttgt | agtcgttcagctgcgtgcctg | |
| 15        | chr16:67202000-6720300 | lgtactgtgtgccccagtgtg | ggtgcagccaggtcaggtgcag | |
| 16        | chr11:100073330-100073750 | cctggtgcctctactctactac | gctagagactggaaccacctta | |
| 17        | chr10:27765450-27765890 | cagcgctgtggaagactccacca | accggaagagctgaaccccgtg | |
| 18        | chr10:120790420-120794450 | gttgcctggcgggtttgga | fttgccttcagctgctgta | |
| 19        | chr11:34375170-34376520 | gagctgcaactggtttggt | ltcctgccctctagtgcacca | |
| 20        | chr12:4212400-4212600 | ggccacactgttctgctg | gcagattctccctctcacat | |
| 21        | chr12:16741700-16742210 | ctcacaggctcctctatatgt | gtcgttcagogaacctacca | |
| 22        | chr12:23690780-23691420 | tccagatctccgtctctcct | gtcctccacactccacag | |
| 23        | chr15:87378500-87378960 | gggcaaggggtctgaaaaa | tcccctctacagggccag | |
| 24        | chr15:35344080-35344550 | cgcgtgagagaccttccttg | ggtcttggagacagacacttcctg | |
| 25        | chr11:67335360-67345050 | agacagagacagactcagapaa | cagctcctccctgctcacta | |
| 26        | chr10:82591720-82592190 | tagtccctgcttcctcttgt | acacagcagagactacgct | |
| 27        | chr11:49906200-49906670 | agacgccctggctagccctctg | tgcgagagccagagacaggt | |
| 28        | chr12:84800050-84801390 | lgtgtggctcatctctctgt | acagctgcagagacagcgct | |
| 29        | chr11:77302970-77303450 | cttctgtcgacttccacttcagc | tgggtttgtgttgttctccaa | |
| 30        | chr16:63680070-63681350 | gcacagctgtgagacagccatct | tatccctctaggggacctg | |
| 31        | chr12:93181330-93181630 | ccagccagacagctctctct | tgtgttagctcctcagccag | |
| 32        | chr11:167550150-167550660 | ggcagggactgagctcgct | tgtcgtgcctgcctgctcgft | |
| 33        | chr11:119640100-119640660 | gcggggccgctgctgtgctgc | ccagccagagcctgctgtg | |
| 34        | chr11:111870630-111871420 | ggtttgatttgagcagccgctg | tctgtcctttctcctctctctctc | |
| 35        | chr16:71560030-715610100 | ggcagccactctcctactgcgtcatctctcagcagctgc | ctcagccagactgcagctgcagctgc | |
| 36        | chr13:15393010-15394300 | caggagagagagagagctg | gggtgtgtctcctcttcactaca | |
| 37        | chr16:91027880-91028320 | gcagagctggctgctgtgctgct | cagctgcagagctgtccttcctt | |
| 38        | chr16:60680120-60680150 | gcggccagccctcgctgctg | ccgctgctcctccctccctct | |
| 39        | chr10:89924060-90025120 | gaaacagagagagtgggaggtgta | gcagctgtctcctgcctgccctc | |
| 40        | chr16:89962250-90026760 | gaaatgtgctggagccagcagc | aacatttgaacgcttcagccac | |
| 41        | chr11:74948370-74948840 | gcggctcgtggtggfctagcag | gcacacagacaccctctccag |
### Primers for p300 ChIP-Seq validation

| Primer ID | peak location | forward sequence (5'-3') | reverse sequence (5'-3') |
|-----------|---------------|---------------------------|--------------------------|
| 1         | chr8:44619470-44820010 | caaagagcgctggggtgaag | ggcctatgtcctgctcaaa |
| 2         | chr7:29974180-29976050 | cagggccagctgctacgttc | gcttgcagaaagtctgcaggt |
| 3         | chr7:132953210-132953750 | cagctgccattcaagccaaatgg | tttgtgagagagggagataa |
| 4         | chr5:90317280-90318130 | cagggagagagacaccccaaaaggggcctctgattccctcaaa |
| 5         | chr5:135221820-135221850 | caggagtcgctgtcagagag | cgggtacagagatgtgctagag |
| 6         | chr10:39694680-103965290 | aagggggtgtctctgctgag | ccggaaagcagcgcacaacc |
| 7         | chr4:57784000-57785270 | gctggtggctctgtcgctgtg | ggcacatagagctccaccaaa |
| 8         | chr3:9228810-9229240 | ccagagggctcctagatctgcc | cagggccacacgtcttctccttc |
| 9         | chr3:6652910-66853530 | cccacttcgctgtctgcgc | ggctcgatgcagaatgtcctcctc |
| 10        | chr3:30915900-30915960 | tgtattggtgttgagagtaacaa | ccgttccacgagaggtgggtaa |
| 11        | chr2:162743500-162743880 | gctggtggctctgtcgctgtg | ccccaagtgtcacaagggaagtg |
| 12        | chr10:30264610-30265090 | gttggagagagagcgagag | aavcctctgctggtgggtgaag |
| 13        | chr17:4505410-45055880 | cgccgctccctgctgcc | gttgggaattcgccagcactc |
| 14        | chr16:56875920-56876450 | ttcctgctctgtcctgctgct | ggtggacagctgacgctggag |
| 15        | chr13:66995560-66995980 | agctgccctccttcacac | ctttgccagcggagcttcctct |
| 16        | chr12:87390730-87391280 | ctcgagccttccttccttc | gtggctgtcagacgagctggag |
| 17        | chr12:74004680-74005270 | gcacagacagatatacgtccg | cagacagctctgcgtcagtccg |
| 18        | chr11:8469290-8469810 | ggaacccagcctgcctcctcaaa | cctcgagatgctccaaccaaa |
| 19        | chr18:7261330-182761800 | cccaggggtgttgctcccag | tcttaagctcttggtggag |
| 20        | chr18:62752300-182752880 | gcacacgctccctgtaaaag | tgtggccagcgcggccaaag |
| 21        | chr15:45777900-154678420 | cctgcgtctctctctctctct | gaaagcctgagatgtgctcc |
| 22        | chr5:19780300-19784600 | ggaagagacccctgacccctcc | ttgagctgtgcggcggtgtg |
| 23        | chr11:50734550-50735600 | cgcctgccctgctgttcatag | ggcacacacacgcccagccc |
| 24        | chr15:15163230-151635020 | cgtggctgaagggggagagt | cttgcaaccaaggagcagtg |
| 25        | chr8:10893790-10893810 | gcctgccgtccgtcaggtgga | aaaggggccccagctgtgtaa |
| 26        | chr7:12511126-12512140 | cccggtctgtcagacgagggt | cgacctggtgtggttaaatgg |
| 27        | chr11:16491843-16491836 | ccctgctgccatataccttcg | gccggagtgtctgtgcttctg |
| 28        | chr11:161345.234-84146382 | agccacacagacgggcttcacaa | aggcgagagcagctgccaa |
| 29        | chr9:79731,656-97,322,185 | tctccggcaggccacattcc | caaggagtgggtttggtgct |
| 30        | chr10:36,940,400-36,940,632 | aagctgctgtctgctctgtt | gttggagacgcgccaggccg |
| 31        | chr12:78405020-78465510 | cagctgccctgctgttcatag | ggcctagcgcagcgcaagcgc |
| 32        | chr2:18152080-181522580 | atctcaagtggctgcctgga | agattcgtctgtggtacagtg |
| 33        | chr10:7,295,365-7,295,795 | cccggtctgtcagagcagcag | tgtccagagagctgagccacaa |
| 34        | chr19:85,812,887-85,822,216 | ccctgagctgcctgcagccattc | gttgcagcaggtgctctcctc |
| 35        | chr9:12979068-12979367 | cggaggtgtctgcgctgtcctg | neccgtcgtcagatgttggc |
| 36        | chr10:852,572,876-852,573,203 | cggaggtgtctgcgctgtcctg | neccgtcgtcagatgttggc |
| 37        | chr10:32975971-32976320 | acgtgtggtctgcgctgtgctt | neccgtcgtcagatgttggc |
| 38        | chr16:54,143,025-130,724 | caaggggtgctgcctgcagcc | cagacggagacggtgtgcctc |
| 39        | chr10:6,64,143,033-6,64,146,688 | ggggggcctggtgcgtgttt | gtggccagcaggtgctgcctc |
| 40        | chr13:7110510-71105570 | ggcctgagcagcagctgagcctg | cagacagagcttgtgtaag |
| 41        | chr5:53,373,035-55,374,217 | gggggggtgggagagagag | cggggcaaaagtcccttccttct |
| 42        | chr4:47,495,036-47,495,840 | gcagctgcctgccaaatggtc | cagccctgtggcttttaa |
| 43        | chr15:62732630-62733200 | aatggagagagctggtggagcctgcgctg | cagacagagct tgtgtaag |
| 44        | chr12:48,15,124,46-48,15,64,478 | gcctgagctgctgcagctgag | cggggcctggtgctgcctcct |
| 45        | chr18:35080260-35080710 | gctggtggctgcctgcagctg | cggggcctggtgctgcctcct |
| 46        | chr3:132621800-132621540 | cggcagctgcctgcagctgcct | cggggcctggtgctgcctcct |

Handoko L*, Han X*, Li G* et. al., Chromatin interactome in ES cells
### Primers for Lamin B ChIP-Seq validation

| Primer ID         | Lamin region | Forward sequence (5’-3’) | Reverse sequence (5’-3’) |
|-------------------|--------------|--------------------------|--------------------------|
| ESC-LAD-Primer1   | chr6:133157900-133885600 | cacacaggccagggctgactgag | ccggctgtgaatttttgtctgctg |
| ESC-LAD-Primer2   | chrX:165031880-165334680 | ccctctctctctctccctgag | acatccagagctgctgtgctgct |
| ESC-LAD-Primer3   | chr1:105384250-105329960 | ccctctctctctctcgctgctg | acatccagagctgctgtgctgct |
| ESC-LAD-Primer4   | chr2:140119310-140360550 | ccctctctctctctcgctgctg | acatccagagctgctgtgctgct |
| ESC-LAD-Primer5   | chr17:90808780-939092800 | ccctctctctctctcgctgctg | acatccagagctgctgtgctgct |
| ESC-LAD-Primer6   | chr14:109788790-110546270 | ccctctctctctctcgctgctg | acatccagagctgctgtgctgct |
| ESC-LAD-Primer7   | chr16:72167190-72472270 | ccctctctctctctcgctgctg | acatccagagctgctgtgctgct |
| ESC-LAD-Primer10  | chr8:52808620-53033740 | ccctctctctctctcgctgctg | acatccagagctgctgtgctgct |
| ESC-LAD-Primer11  | chr12:101749700-102015260 | ccctctctctctctcgctgctg | acatccagagctgctgtgctgct |
| ESC-LAD-Primer12  | chr11:110306970-111345490 | ccctctctctctctcgctgctg | acatccagagctgctgtgctgct |
| ESC-LAD-Primer13  | chr13:86146570-86383580 | ccctctctctctctcgctgctg | acatccagagctgctgtgctgct |
| ESC-LAD-Primer14  | chr2:110978130-111520876 | ccctctctctctctcgctgctg | acatccagagctgctgtgctgct |
| ESC-LAD-Primer15  | chr5:85707090-86717180 | ccctctctctctctcgctgctg | acatccagagctgctgtgctgct |
| ESC-LAD-Primer16  | chr12:124499710-124748270 | ccctctctctctctcgctgctg | acatccagagctgctgtgctgct |
| ESC-LAD-Primer17  | chr7:76357690-76651450 | ccctctctctctctcgctgctg | acatccagagctgctgtgctgct |
| ESC-LAD-Primer19  | chr9:88619710-88702820 | ccctctctctctctcgctgctg | acatccagagctgctgtgctgct |
| ESC-LAD-Primer20  | chr4:89265500-89598550 | ccctctctctctctcgctgctg | acatccagagctgctgtgctgct |
| Lamin-negative1   | chr19:47382500-47387500 | ccctctctctctctcgctgctg | acatccagagctgctgtgctgct |
| Lamin-negative2   | chr1:53045000-53050000 | ccctctctctctctcgctgctg | acatccagagctgctgtgctgct |
| Lamin-negative3   | chr1:53045000-53050000 | ccctctctctctctcgctgctg | acatccagagctgctgtgctgct |
| Lamin-negative4   | chr1:43780000-43785000 | ccctctctctctctcgctgctg | acatccagagctgctgtgctgct |
| Lamin-negative5   | chr1:53045000-53050000 | ccctctctctctctcgctgctg | acatccagagctgctgtgctgct |
| Lamin-negative6   | chr1:43780000-43785000 | ccctctctctctctcgctgctg | acatccagagctgctgtgctgct |
| Lamin-negative7   | chr1:43780000-43785000 | ccctctctctctctcgctgctg | acatccagagctgctgtgctgct |
| Lamin-negative8   | chr1:43780000-43785000 | ccctctctctctctcgctgctg | acatccagagctgctgtgctgct |
| Lamin-negative9   | chr1:43780000-43785000 | ccctctctctctctcgctgctg | acatccagagctgctgtgctgct |
| Lamin-negative10  | chr1:43780000-43785000 | ccctctctctctctcgctgctg | acatccagagctgctgtgctgct |