Supplementary Figure S1

A

R²=0.99

log2(shSMARCA4 Rep2) gene counts

log(shSMARCA4 Rep1) gene counts

1x10⁶

1x10³

1x10¹

1x10¹   1x10³   1x10⁶

B

R²=0.99

log(shSCRAM Rep2) gene counts

log(shSCRAM Rep1) gene counts

1x10⁶

1x10³

1x10¹

1x10¹   1x10³   1x10⁶

C

Mean count = 922

# of genes

2000

1500

1000

500

0

0   2500   5000   7500   10,000

shSMARCA4 gene expression count

D

Mean count = 1572

# of genes

2000

1500

1000

500

0

0   2500   5000   7500   10,000

shSCRAM gene expression count

E

shSMARCA4 / shSCRAM log2 fold change

2

1

0

-1

-2

10⁻¹   10⁻²   10⁻³   10⁻⁴   10⁻⁵

mean expression
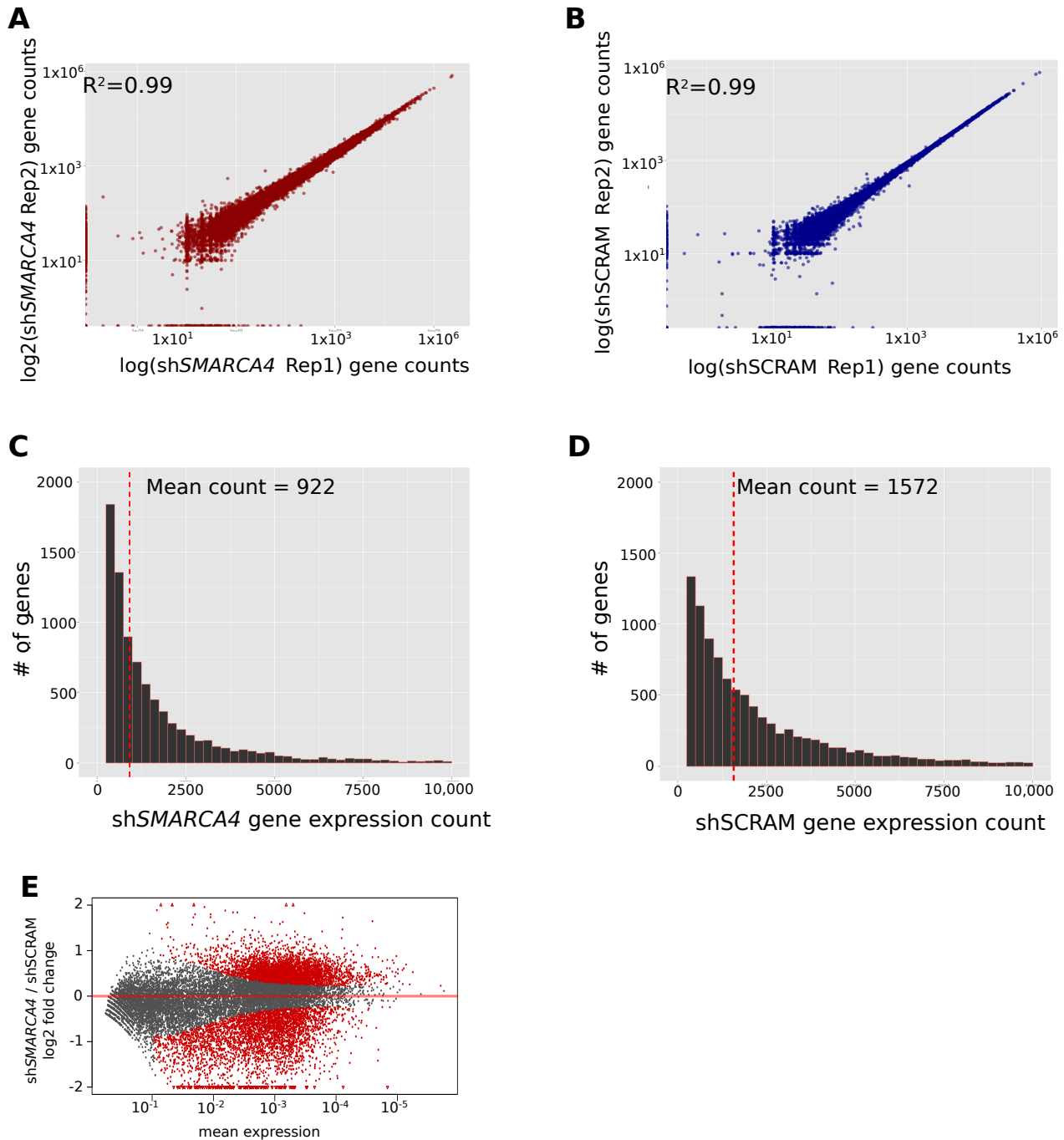
Supplementary Figure S1 - RNAseq Analysis in *SMARCA4* knockdown cells:
Scatterplots showing the gene expression Pearson Correlation coefficients
between the RNA-seq biological replicates for **A)** sh*SMARCA4* and **B)** shSCRAM samples.
Histograms showing the number of genes and the expression levels for **C)** sh*SMARCA4* and **D)** shSCRAM
combined datasets. *SMARCA4* depletion results in a lower mean gene expression.
E) MA-plot showing the log2 fold change and the mean expression of significantly altered (red) and unaltered
genes (black).

Supplementary Figure S2

A MCF-10A sh*SMARCA4*/shSCRAM Differentially Expressed Gene Positions

B

C

D
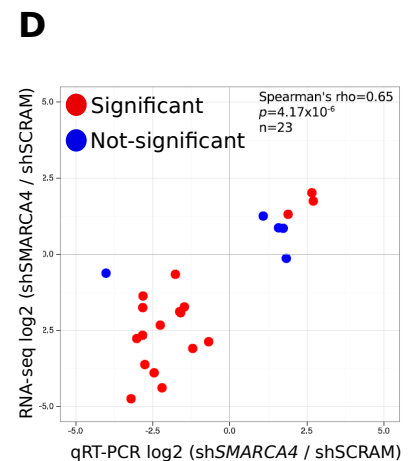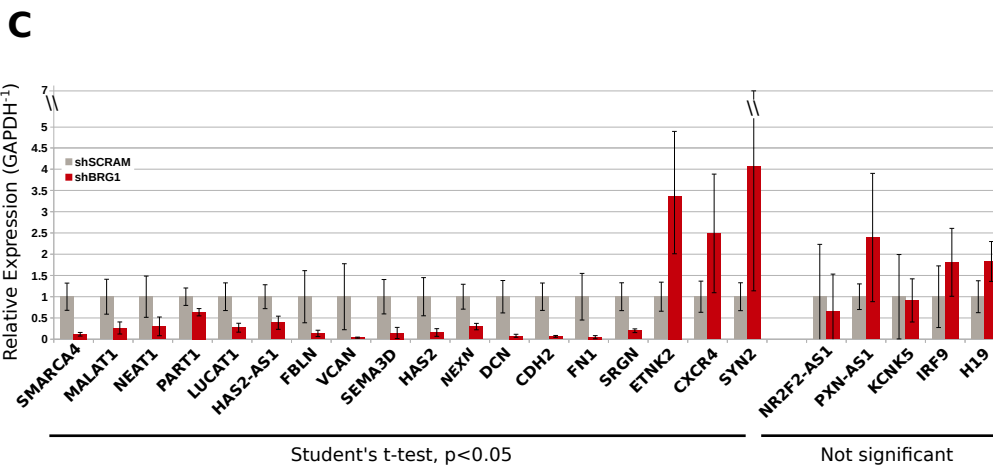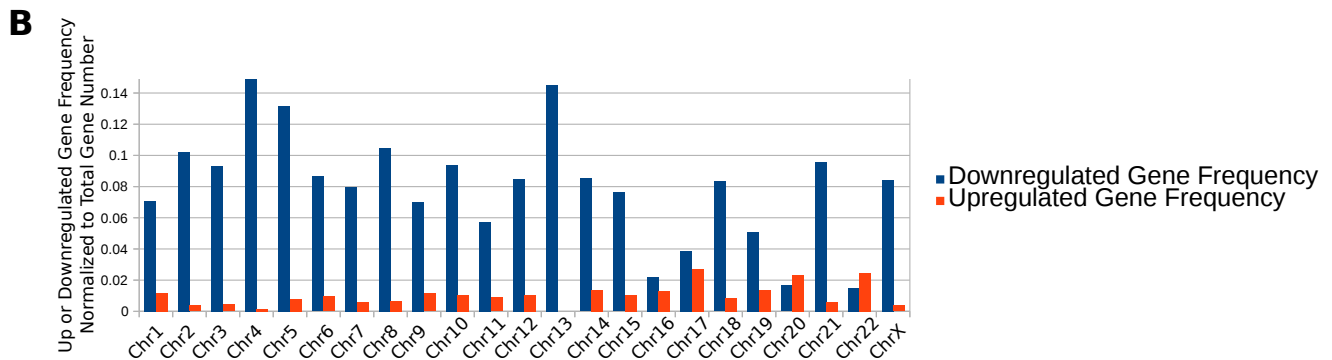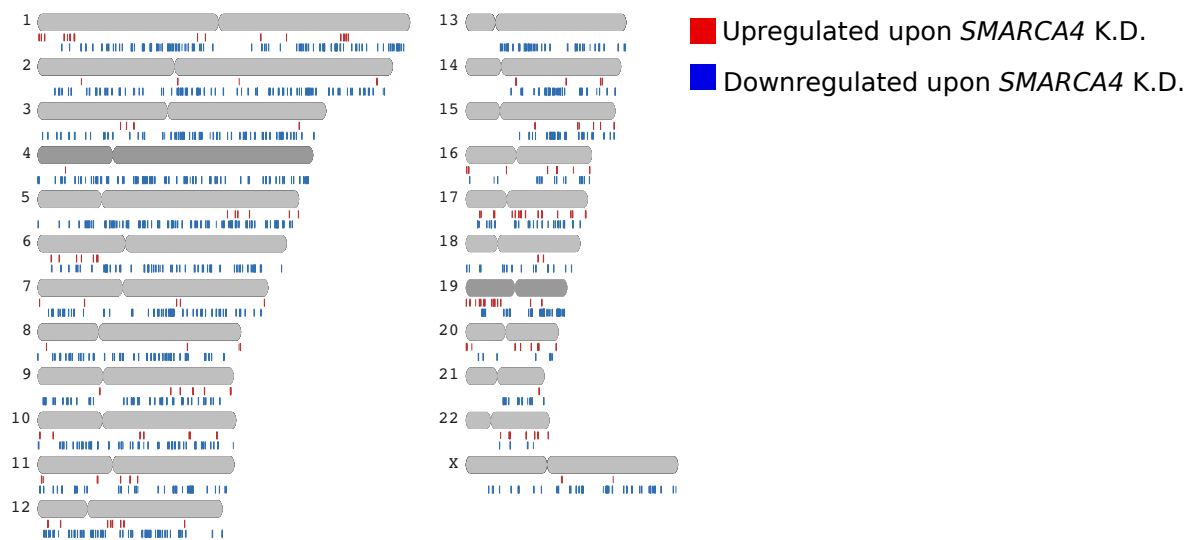
Supplementary Figure S2 - Validation of RNA-seq Analysis
A) Chromosome ideograms showing the locations of up (red) and down (blue) regulated genes upon *SMARCA4* knockdown.
B) Bar graph showing the gene frequencies of up (red) and down (blue) regulated genes for each chromosome.
C) qRT-PCR validation of the RNA-seq data for 23 genes. The y-axis shows the relative expression level of each gene compared to GAPDH. 18 of 23 genes showed significant differential expression.Error bars:S.D.
**D)** Scatterplot showing the correlation of log2 fold change values for the 23 genes from the RNA-seq and qRT-PCR analyses. There is a significant positive correlation between the RNA-seq and qRT-PCR data (Spearman's rho=0.65).
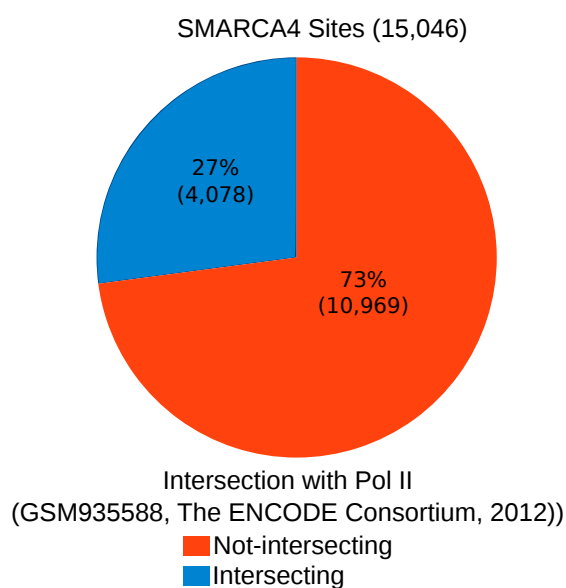
Supplementary Figure S3

A

Sample Name	Biological Replicate	Raw Reads	Mapped Reads
SMARCA4 ChIP 1	1	14,642,947	9,057,272
SMARCA4 ChIP 2	2	25,641,108	20,088,204
INPUT 1	1	25,624,664	22,301,264
INPUT 2	2	37,688,445	32,712,701

B
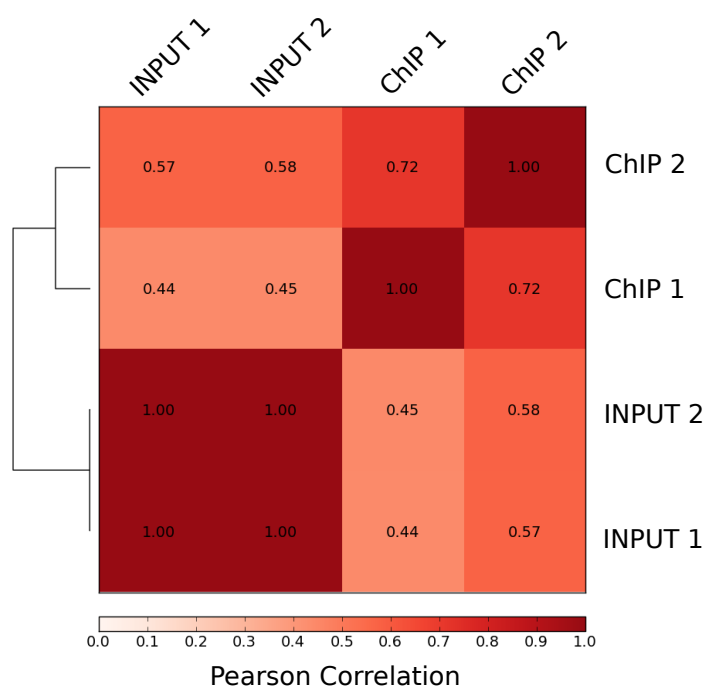
C

**Supplementary Figure S3 - SMARCA4 ChIP-seq Analysis**
A) Table showing the sequenced and mapped reads for each ChIP-seq biological replicate.
B) Matrix showing the Pearson correlation of the signal intensity between the pull down (ChIP) and input samples.
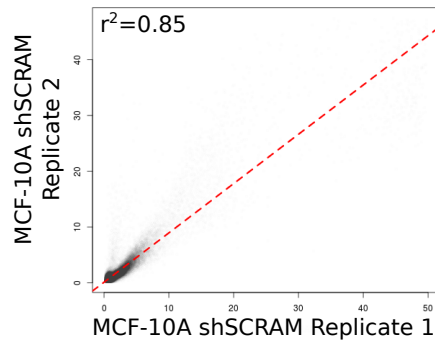**C)** Pie chart showing intersections of SMARCA4 peaks with publicly available Pol2 peaks from MCF-10A cells.
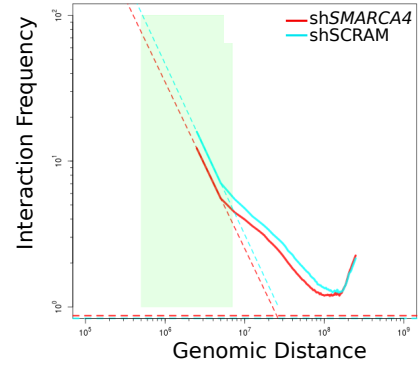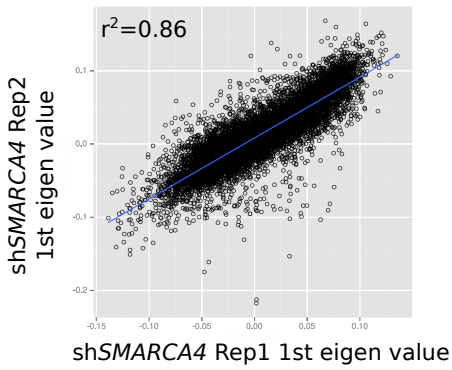
Supplementary Figure S4
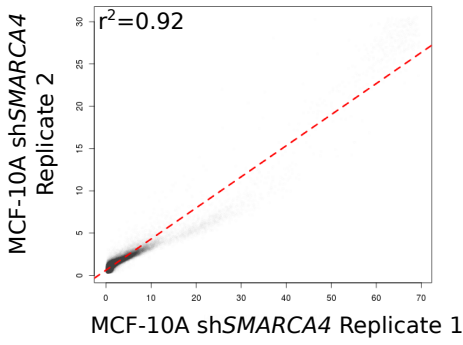
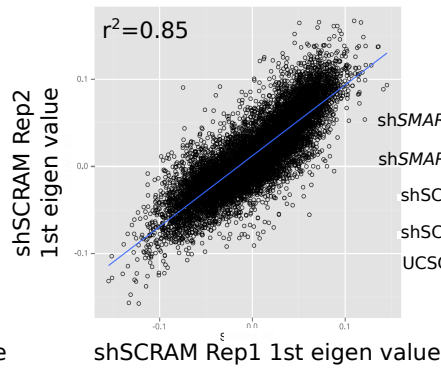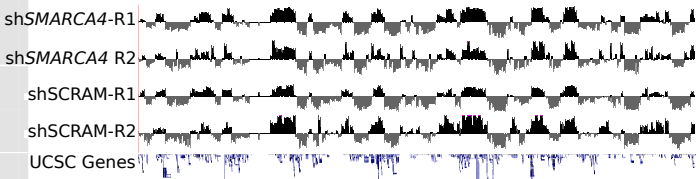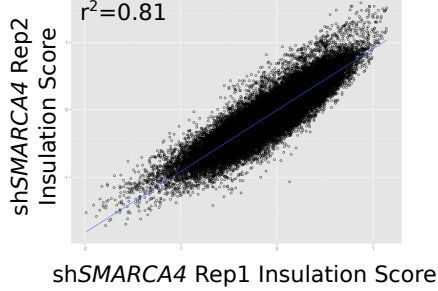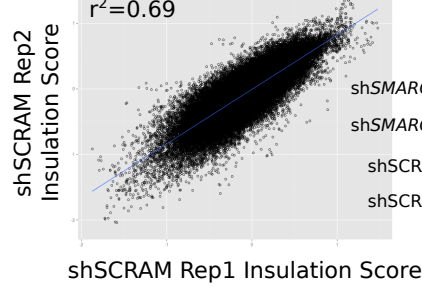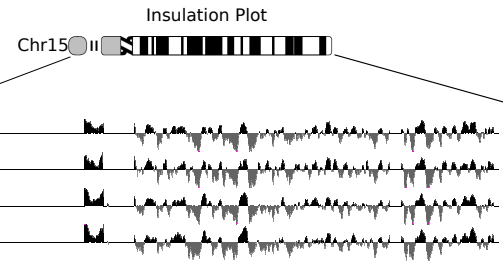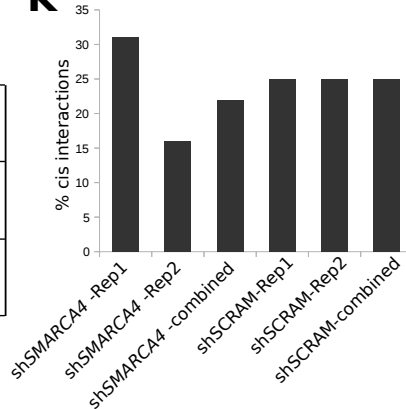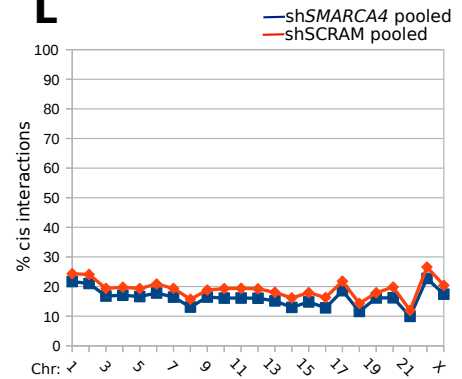**Supplementary Figure S4 - Reproducibility of Hi-C Experiments**
A-B) Scatter plots comparing normalized interactions between pairs of 2.5Mb bins in the two biological replicates from **A)** sh*SMARCA4* and **B)** shSCRAM datasets. Pearson correlation (r^2 value) is shown on the graphs.
C) Scaling plot at 2.5Mb resolution showing that in both sh*SMARCA4* and shSCRAM Hi-C samples, interaction frequency decreases similarly as a function of genomic distance. The difference between the sh*SMARCA4* and shSCRAM curves is due to the differences in % cis interactions.
D-E) Scatter plots showing the correlation of the 1st eigenvector values for each 250kb bin from the compartment analysis for each biological replicate in **D)** sh*SMARCA4* and **E)** shSCRAM datasets.
F) Example of the first eigen values for each biological replicate across Chr18. The replicates show high correlation in genomic compartmentalization.
G-H) Insulation score correlation among the biological replicates for each 40kb bin for
G) sh*SMARCA4* and **H)** shSCRAM datasets.
I) Example of the insulation plot across Chr15 for each biological replicate.
J) Pearson correlations of the pooled Hi-C replicates between shSCRAM, sh*SMARCA4,* and previously published parental (wildtype) MCF-10A cells (Barutcu et al. 2015).
K) Bargraph showing the cis interaction percentage of individual and pooled sh*SMARCA4* and shSCRAM Hi-C replicates.
L) The percentage of cis/trans interaction frequencies of pooled sh*SMARCA4* and shSCRAM datasets for each chromosome.

Supplementary Figure S5

A

shSMARCA4

Chr5

B

shSCRAM

Chr5

Chr9

Chr9

Chr19

Chr19
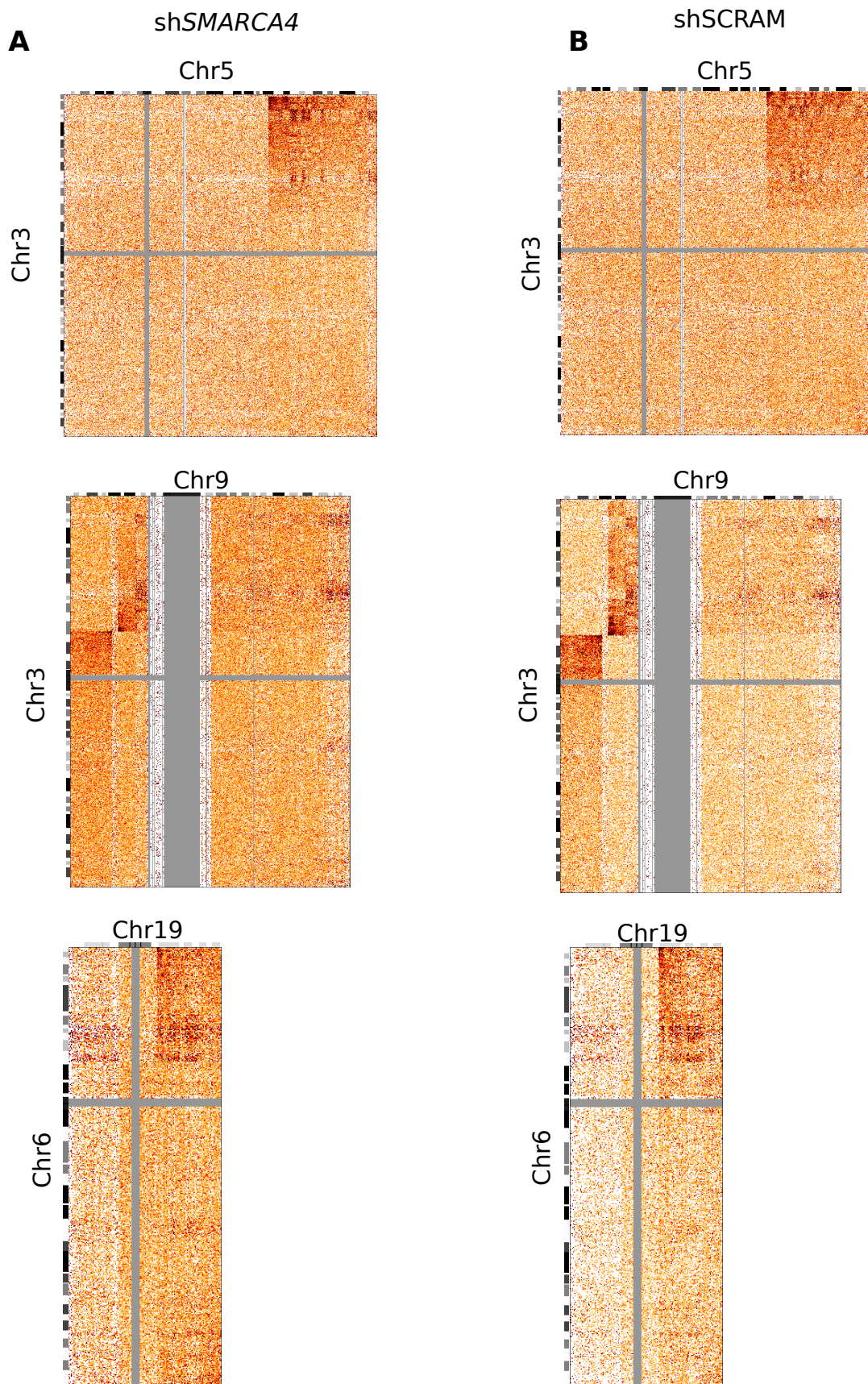
Supplementary Figure S5 - Translocations in shSCRAM and shSMARCA4 MCF-10A cells
Inter-chromosomal interaction heatmaps showing the translocated regions in the MCF-10A genome for **A)** shSMARCA4 and **B)** shSCRAM Hi-C datasets.

Supplementary Figure S6
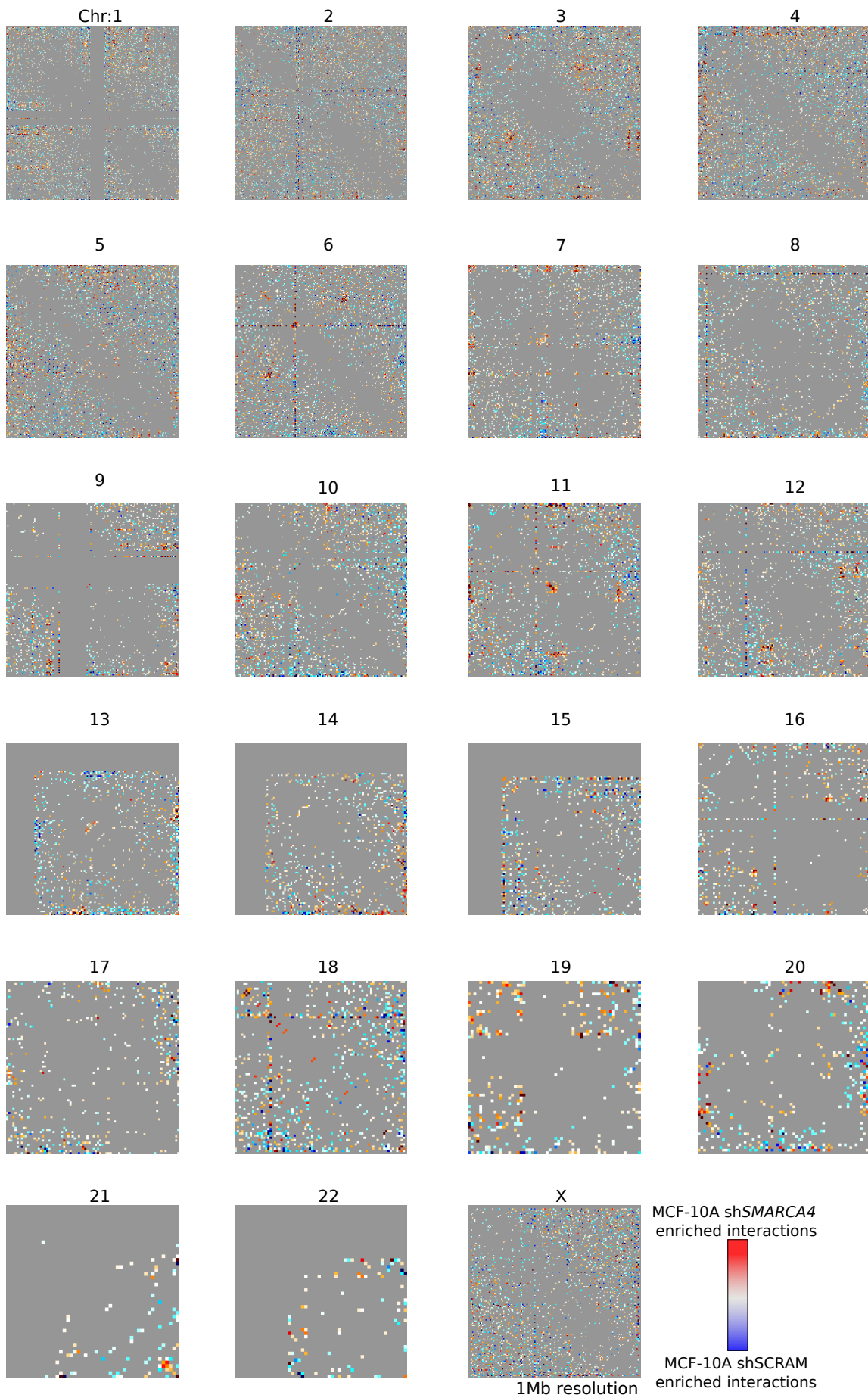
Supplementary Figure S6 - SMARCA4-dependent significant interactions
sh*SMARCA4* / shSCRAM interactions that are significantly different (see Methods) at 1Mb for each chromosome (versus itself).

Supplementary Figure S7

A   shSCRAM

B   shSMARCA4

0.38

0

C

*p<1x 10⁻¹⁸⁰

*p=1.37 x 10⁻⁵

shSMARCA4 / shSCRAM Interaction Score Difference (Normalized z-score difference)

shSCRAM   shSMARCA4   shSCRAM   shSMARCA4

in-cis   in-trans

D

p=1.19x10⁻¹⁰⁶   p=3.08x10⁻¹⁷⁷

Interaction Frequency (z-score)

shSCRAM   shSMARCA4   shSCRAM   shSMARCA4

Intra chromosomal arm cis-interactions   Inter chromosomal arm cis-interactions

E

N.S. (p=0.17)

shSMARCA4 /shSCRAM 1000X Randomized Region Interaction Frequency (z-Score)

shSCRAM   shSMARCA4

Supplementary Figure S7 - Large scale alterations in *SMARCA4* knockdown cells
Chromosome by chromosome interaction heatmaps for **A)** shSCRAM and **B)** sh*SMARCA4* Hi-C datasets.
C) Boxplot showing the sh*SMARCA4* / shSCRAM interaction difference scores for cis and trans-interactions. sh*SMARCA4* MCF-10A cells show higher frequency of trans-interactions (Replicate Pearson correlation r^2=0.94 for sh*SMARCA4* and r^2=0.85 for shSCRAM biologial replicates.)
D) Boxplot showing the intra-arm and inter-arm chromosomal interaction frequencies for sh*SMARCA4* and shSCRAM datasets. p-values: Wilcoxon rank-sum test.
E) 1000X randomized set of regions to serve as a control for telomeric interactions. p-values: Wilcoxon rank-sum test.

Supplementary Figure S8

A

shSCRAM TAD border
interaction pileup mean

B

sh*SMARCA4* TAD border
interaction pileup mean

C

D

E

F

Supplementary Figure S8 - Topologically Associated Domain (TAD) Analysis
Mean of pile up interaction frequencies 1Mb around the TAD boundaries for **A)** shSCRAM and **B)** sh*SMARCA4* cells
Frequency of **C)** UCSC genes and **D)** Pol2 binding at and +/- 1Mb of TAD boundaries. **E)** Venn diagram showing
the overlap of SMARCA4 peaks from this study with previously published CTCF ChIP-seq peaks from MCF-10A
cells (Ross-Innes et al. 2011). **F)** Venn diagram showing the SMARCA4 and CTCF overlap for peaks that are
located at TAD boundaries.

Supplementary Figure S9

A

B

C

D

SMARCA4 knockdown affects the nucleosome occupancy and chromatin accessibility near CTCF sites.
A) Plot showing the nucleosome occupancy around the CTCF sites in wildtype and *SMARCA4* knockdown cells. MNase-seq data from wildtype and *SMARCA4* knockdown MEFs (Tolstorukov et al, 2013) were used. The MEF CTCF ChIP-seq data was downloaded from the ENCODE project (ENCFF001YAQ). Nucleosome occupancy was computed for each CTCF site by adding the nucleosome tag counts in 75 bp windows with 1bp smoothing step from the peak center of the CTCF sites, and normalizing the resulting score by the average tag count (positions from -1kb to -0.7 kb from the CTCF site).
B) Results of repeated random subsampling of 10% of the CTCF sites (light colored traces) superimposed on the results for the full dataset (dark lines), showing that the SMARCA4-dependent decrease in nucleosome occupancy around CTCF sites is not due to a few outlier sites harboring an unusual decrease.
C) Randomization of the MNase-seq signal by randomly shifting the CTCF site positions 2 to 100kb up or downstream. There is no significant pattern or difference between WT and mutant MNase-seq signals.
D) Bar graph showing SMARCA4-dependent altered DNase I hypersensitivity near CTCF sites. Changes in chromatin accessibility (Morris et al. 2014) were determined by calculating the Tet+/Tet- tag density values of each DNase I hotspot in the control and *SMARCA4* dominant negative expressing 3134 murine epithelial cells. The hotspots with a Tet+/Tet- ratio more than +/- 1 standard deviation of the average ratio from all the hotspots were considered as "affected". The affected sites were intersected with the CTCF ChIP-seq dataset from the same cell lines (Stavreva et al. 2015). The *SMARCA4*-DN affected DNase I sites overlap significantly more with CTCF sites (Chi-square Test with Yates' correction) compared to control cells.