**The American Journal of Human Genetics**

**Supplemental Data**

# Invariant TAD Boundaries Constrain Cell-Type-Specific Looping Interactions between Promoters and Distal Elements around the *CFTR* Locus

**Emily M. Smith, Bryan R. Lajoie, Gaurav Jain, and Job Dekker**

**Table S1: 5C probe set.** This table contains all the probes used in the 5C experiment: their name, sequence, and genomic location.

**Table S2: Mapping Statistics.** This table contains all mapping statistics for the raw data of our 10 5C libraries.

**Table S3: 5C library read depth.** This table summarizes the 5C read depth for each replicate, indicating the number of reads in the raw data, cis-purged data, singleton- removed data, and final coverage corrected data.

| | Raw Data | | *cis*-purged Data | | Singleton-removed Data | | Coverage-corrected Data | |
|---|---|---|---|---|---|---|---|---|
| | #interactions | #reads | #interactions | #reads | #interactions | #reads | #interactions | #reads |
| Caco2-R1 | **50,986** | **12,930,158** | **24,492** | **9,240,611** | **24,449** | **7,265,532** | **20,351** | **7,354,017** |
| Caco2-R2 | **50,970** | **12,724,492** | **27,090** | **9,783,411** | **27,046** | **9,453,039** | **22,278** | **8,248,685** |
| Calu3-R1 | **66,029** | **18,308,618** | **36,812** | **12,073,332** | **36,768** | **11,970,277** | **32,886** | **8,222,536** |
| Calu3-R2 | **148,785** | **25,380,481** | **39,883** | **15,507,757** | **39,908** | **15,378,124** | **36,733** | **8,522,545** |
| Capan1-R1 | **126,462** | **18,101,240** | **38,046** | **6,041,896** | **38,002** | **5,760,920** | **35,337** | **8,853,211** |
| Capan1-R2 | **161,499** | **27,547,883** | **39,308** | **8,722,529** | **39,264** | **8,380,681** | **36,799** | **8,897,244** |
| GM12878-R1 | **88,198** | **24,408,480** | **37,880** | **9,907,568** | **37,836** | **9,687,772** | **34,869** | **10,734,618** |
| GM12878-R2 | **190,199** | **28,011,778** | **40,190** | **15,253,503** | **40,146** | **14,800,666** | **39,392** | **9,711,216** |
| HepG2-R1 | **161,365** | **31,681,378** | **39,568** | **24,630,673** | **39,524** | **24,508,802** | **35,515** | **11,181,660** |
| HepG2-R2 | **147,911** | **26,426,516** | **39,909** | **16,945,637** | **39,865** | **16,718,446** | **35,701** | **9,771,604** |

**Table S4: Pearson correlation between datasets.** This table contains Pearson correlations between all the 5C replicates used in this study.

| | Caco2 R1 | Caco2 R2 | Calu3 R1 | Calu3 R2 | Capan1 R1 | Capan1 R2 | GM12878 R1 | GM12878 R2 | HepG2 R1 | HepG2 R2 |
|---|---|---|---|---|---|---|---|---|---|---|
| Caco2 R1 | **1** | **0.738** | **0.617** | **0.645** | **0.617** | **0.594** | **0.518** | **0.503** | **0.565** | **0.447** |
| Caco2 R2 | | **1** | **0.754** | **0.816** | **0.751** | **0.752** | **0.731** | **0.677** | **0.798** | **0.589** |
| Calu3 R1 | | | **1** | **0.911** | **0.792** | **0.802** | **0.651** | **0.642** | **0.736** | **0.551** |
| Calu3 R2 | | | | **1** | **0.792** | **0.796** | **0.669** | **0.656** | **0.747** | **0.564** |
| Capan1 R1 | | | | | **1** | **0.896** | **0.707** | **0.647** | **0.779** | **0.569** |
| Capan1 R2 | | | | | | **1** | **0.726** | **0.632** | **0.755** | **0.539** |
| GM12878 R1 | | | | | | | **1** | **0.799** | **0.825** | **0.705** |
| GM12878 R2 | | | | | | | | **1** | **0.816** | **0.897** |
| HepG2 R1 | | | | | | | | | **1** | **0.762** |
| HepG2 R2 | | | | | | | | | | **1** |

**Table S5: Probes Removed in Probe Filtering Step.** This table shows the 34 probes removed from all datasets after our probe filtering step.

**Trim Amount: 0.85**

| Flagged Probe | #datasets in which this probe is flagged |
|---|---|
| 5C_2410_EMS03_FOR_102 | 10 |
| 5C_2410_EMS03_FOR_140 | 10 |
| 5C_2410_EMS03_FOR_349 | 10 |
| 5C_2410_EMS03_FOR_429 | 10 |
| 5C_2410_EMS03_FOR_773 | 10 |
| 5C_2410_EMS03_FOR_165 | 10 |
| 5C_2410_EMS03_FOR_2 | 10 |
| 5C_2410_EMS03_REV_10 | 9 |
| 5C_2410_EMS03_FOR_75 | 9 |
| 5C_2410_EMS03_FOR_74 | 9 |
| 5C_2410_EMS03_REV_13 | 8 |
| 5C_2410_EMS03_FOR_8 | 8 |
| 5C_2410_EMS03_FOR_51 | 8 |
| 5C_2410_EMS03_FOR_54 | 7 |
| 5C_2410_EMS03_FOR_197 | 7 |
| 5C_2410_EMS03_FOR_762 | 7 |
| 5C_2410_EMS03_REV_111 | 7 |
| 5C_2410_EMS03_FOR_228 | 7 |
| 5C_2410_EMS03_FOR_389 | 7 |
| 5C_2410_EMS03_FOR_350 | 7 |
| 5C_2410_EMS03_REV_523 | 6 |
| 5C_2410_EMS03_FOR_407 | 6 |
| 5C_2410_EMS03_FOR_1 | 6 |
| 5C_2410_EMS03_FOR_57 | 5 |
| 5C_2410_EMS03_FOR_129 | 5 |
| 5C_2410_EMS03_FOR_677 | 5 |
| 5C_2410_EMS03_FOR_117 | 5 |
| 5C_2410_EMS03_FOR_248 | 4 |
| 5C_2410_EMS03_FOR_246 | 4 |
| 5C_2410_EMS03_FOR_864 | 4 |
| 5C_2410_EMS03_FOR_283 | 4 |
| 5C_2410_EMS03_FOR_607 | 4 |
| 5C_2410_EMS03_FOR_658 | 4 |
| 5C_2410_EMS03_FOR_298 | 4 |

**Table S6: Individual Interactions Removed in Singleton Removal Step.** This table lists the 44 individual interactions removed in the singleton removal step.

| z-score 12<br><br>Probe-Probe interaction | #datasets in which this interaction is flagged |
|---|---|
| 5C_2410_EMS03_FOR_124_5C_2410_EMS03_REV_17 | 10 |
| 5C_2410_EMS03_FOR_128_5C_2410_EMS03_REV_17 | 10 |
| 5C_2410_EMS03_FOR_464_5C_2410_EMS03_REV_405 | 10 |
| 5C_2410_EMS03_FOR_126_5C_2410_EMS03_REV_17 | 9 |
| 5C_2410_EMS03_FOR_128_5C_2410_EMS03_REV_203 | 9 |
| 5C_2410_EMS03_FOR_21_5C_2410_EMS03_REV_203 | 9 |
| 5C_2410_EMS03_FOR_339_5C_2410_EMS03_REV_898 | 9 |
| 5C_2410_EMS03_FOR_605_5C_2410_EMS03_REV_777 | 9 |
| 5C_2410_EMS03_FOR_7_5C_2410_EMS03_REV_405 | 9 |
| 5C_2410_EMS03_FOR_817_5C_2410_EMS03_REV_405 | 9 |
| 5C_2410_EMS03_FOR_90_5C_2410_EMS03_REV_613 | 9 |
| 5C_2410_EMS03_FOR_93_5C_2410_EMS03_REV_405 | 9 |
| 5C_2410_EMS03_FOR_467_5C_2410_EMS03_REV_719 | 8 |
| 5C_2410_EMS03_FOR_548_5C_2410_EMS03_REV_405 | 8 |
| 5C_2410_EMS03_FOR_579_5C_2410_EMS03_REV_72 | 8 |
| 5C_2410_EMS03_FOR_125_5C_2410_EMS03_REV_17 | 7 |
| 5C_2410_EMS03_FOR_19_5C_2410_EMS03_REV_203 | 7 |
| 5C_2410_EMS03_FOR_214_5C_2410_EMS03_REV_761 | 7 |
| 5C_2410_EMS03_FOR_71_5C_2410_EMS03_REV_761 | 7 |
| 5C_2410_EMS03_FOR_87_5C_2410_EMS03_REV_782 | 7 |
| 5C_2410_EMS03_FOR_105_5C_2410_EMS03_REV_405 | 6 |
| 5C_2410_EMS03_FOR_208_5C_2410_EMS03_REV_711 | 6 |
| 5C_2410_EMS03_FOR_503_5C_2410_EMS03_REV_420 | 6 |
| 5C_2410_EMS03_FOR_125_5C_2410_EMS03_REV_665 | 5 |
| 5C_2410_EMS03_FOR_132_5C_2410_EMS03_REV_17 | 5 |
| 5C_2410_EMS03_FOR_139_5C_2410_EMS03_REV_405 | 5 |
| 5C_2410_EMS03_FOR_154_5C_2410_EMS03_REV_405 | 5 |
| 5C_2410_EMS03_FOR_641_5C_2410_EMS03_REV_420 | 5 |
| 5C_2410_EMS03_FOR_68_5C_2410_EMS03_REV_594 | 5 |
| 5C_2410_EMS03_FOR_98_5C_2410_EMS03_REV_405 | 5 |
| 5C_2410_EMS03_FOR_12_5C_2410_EMS03_REV_203 | 4 |
| 5C_2410_EMS03_FOR_136_5C_2410_EMS03_REV_405 | 4 |
| 5C_2410_EMS03_FOR_237_5C_2410_EMS03_REV_665 | 4 |
| 5C_2410_EMS03_FOR_394_5C_2410_EMS03_REV_405 | 4 |
| 5C_2410_EMS03_FOR_5_5C_2410_EMS03_REV_808 | 4 |
| 5C_2410_EMS03_FOR_502_5C_2410_EMS03_REV_420 | 4 |
| 5C_2410_EMS03_FOR_503_5C_2410_EMS03_REV_421 | 4 |
| 5C_2410_EMS03_FOR_597_5C_2410_EMS03_REV_666 | 4 |

| | |
|---|---|
| 5C_2410_EMS03_FOR_626_5C_2410_EMS03_REV_420 | 4 |
| 5C_2410_EMS03_FOR_63_5C_2410_EMS03_REV_405 | 4 |
| 5C_2410_EMS03_FOR_641_5C_2410_EMS03_REV_421 | 4 |
| 5C_2410_EMS03_FOR_672_5C_2410_EMS03_REV_203 | 4 |
| 5C_2410_EMS03_FOR_751_5C_2410_EMS03_REV_665 | 4 |
| 5C_2410_EMS03_FOR_883_5C_2410_EMS03_REV_405 | 4 |

**Table S7: Significant Interactions.** This table lists all significant intra-TAD and inter - TAD interactions in the 5 cell types.

**Table S8: RT-PCR Primers.** Primers used in our qPCR experiment testing gene expression in the 5C region.

Genes in 5C Region - qPCR primers

| Primer Name | Primer Sequence | Exon # | Genome Coordinates (hg18) |
|---|---|---|---|
| **TES:** *testin isoform 1* | | | |
| TES_F_A | GCCCCTTGTTTAAAATGCAA | 2 | 115661854-115661873 |
| TES_R_A | TGCTCAAGAGGACATCATGC | 3 | 115676346-115676365 |
| **CAV2:** *caveolin 2 isoform a and b* | | | |
| CAV2_F_A | GGCTCAACTCGCATCTCAAG | 1 | 115927202-115927221 |
| CAV2_R_A | CAGGAACACCGTCAGGAACT | 2 | 115927656-115927675 |
| **CAV1:** *caveolin 1* | | | |
| CAV1_F_A | GAGCTGAGCGAGAAGCAAGT | 2 | 115953887-115953906 |
| CAV1_R_A | CAAATGCCGTCAAAACTGTG | 3 | 115986275-115986294 |
| **MET:** *met proto-oncogene isoform a precursor* | | | |
| MET_F_B | CCAATGACCTGCTGAAATTG | 11 | 116197041-116197060 |
| MET_R_B | CTTTTCCAAGGACGGTTGAA | 12 | 116198805-116198824 |
| **CAPZA2:** *capping protein (actin filament) muscle Z-line* | | | |
| CAPZA2_F_A | GAAGGAGGCAACTGATCCAA | 5 | 116331553-116331572 |
| CAPZA2_R_A | GCTTGGAACTGATGGCTTTC | 6 | 116333601-116333620 |
| **ST7:** *suppression of tumorigenicity 7 isoform b* | | | |
| ST7_F_A | TTCCAGTAACGGGGACTCAG | 3 | 116546967-116546986 |
| ST7_R_A | TGGATTTCGCCATACTTTGC | 4 | 116557086-116557105 |
| **WNT2:** *wingless-type MMTV integration site family* | | | |
| WNT2_F_B | GTGGATGCAAAGGAAAGGAA | 3 | 116742416-116742435 |
| WNT2_R_B | AGCCAGCATGTCCTGAGAGT | 4 | 116725090-116725109 |
| **ASZ1:** *ankyrin repeat, SAM and basic leucine zipper* | | | |
| ASZ1_F_B | CACGTCAGGGTCATAAA | 6 | 116812115-116812137 |
| ASZ1_R_B | GCTGTTGAAGTTTTCCTTCCA | 7 | 116810346-116810366 |
| **CFTR:** *cystic fibrosis transmembrane conductance regulator* | | | |
| CFTR2-3F | CCCTTCTGTTGATTCTGCTG | 2 | 116931609-116931628 |
| CFTR2-3R | AAGGGCATTAATGAGTTTAGGA | 3 | 116936357-116936378 |
| **CTTNBP2:** *cortactin binding protein 2* | | | |
| CTTNBP2_F_C | AAAATGGCTTCACACCCTTG | 6 | 117210181-117210200 |
| CTTNBP2_R_C | TGTCTGTCCTCCATCAGCAG | 7 | 117207818-117207837 |
| **LSM8:** *U6 snRNA-associated Sm-like protein LSm8* | | | |
| LSM8_F_A | CAGCTCTTCACAGGGGGTAG | 3 | 117615633-117615652 |
| LSM8_R_A | CTGCTCGAATATTCCCCAAA | 4 | 117619247-117619000 |
| **ANKRD7:** *ankyrin repeat domain 7 isoform b* | | | |
| ANKRD7_F_A | ACCTTTGCACCTAGCCTGTG | 2 | 117661724-117661743 |

| | | | |
|---|---|---|---|
| ANKRD7_R_A | ATCTGGGTCTGCACCAAAGT | 3 | 117662034-117662053 |

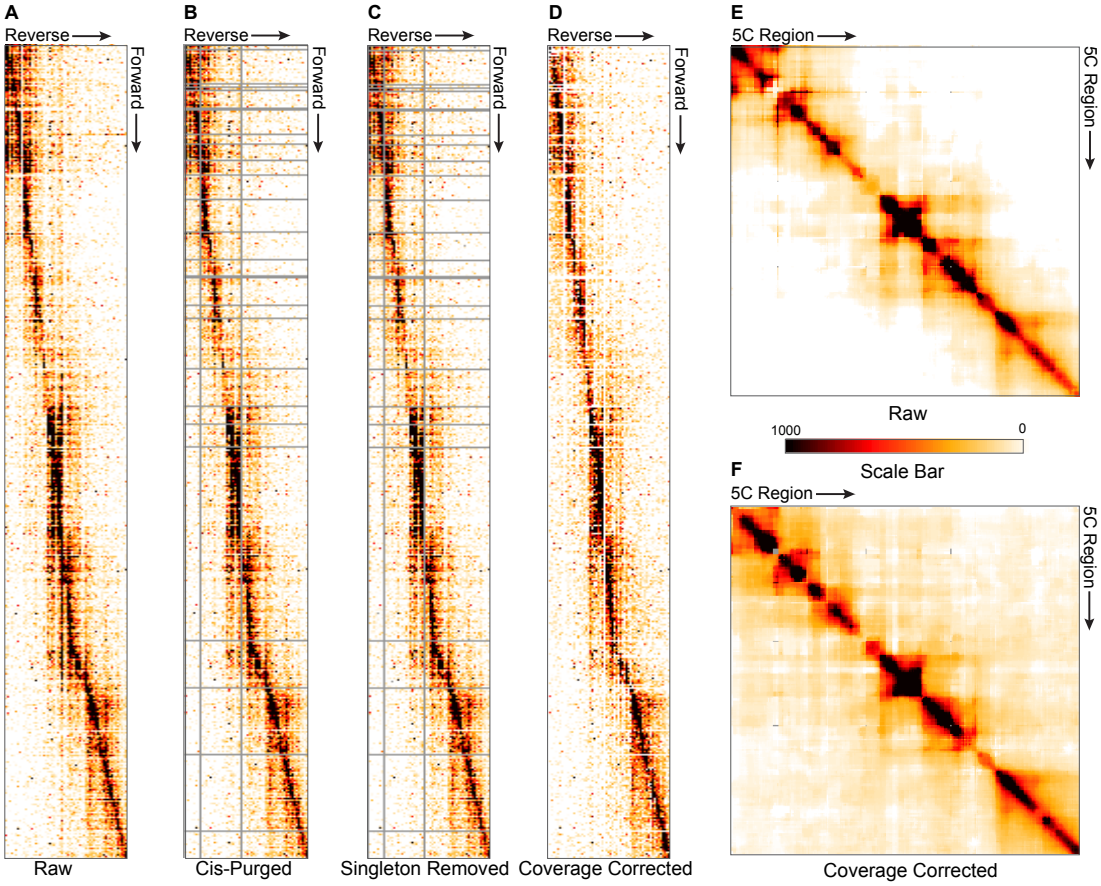
**HPRT1:** *Chromosome X: hypoxanthine phosphoribosyltransferase 1*

| | | | |
|---|---|---|---|
| HPRT2-3F | TGAGGATTTGGAAAGGGTGT | 2 | 133435114-133435133 |
| HPRT2-3R | TAATCCAGCAGGTCAGCAAA | 3 | 133436965-133436984 |

**Figures S1-S10: Correction pipeline for each replicate.** A) Raw data. B) Data after cis-purge. Grey stripes represent primers that were removed. C) Data after singleton removal. Grey stripes are the primers removed in the previous step, grey pixels are the individual reactions that were removed in this step. D) Final coverage corrected data. Grey lines have been removed and are now represented in white. E) Binned raw data (100kb, 10kb step). F) Binned coverage corrected data (100kb, 10kb step).

Caco2 R1

Figure S1



Caco2 R2

Figure S2

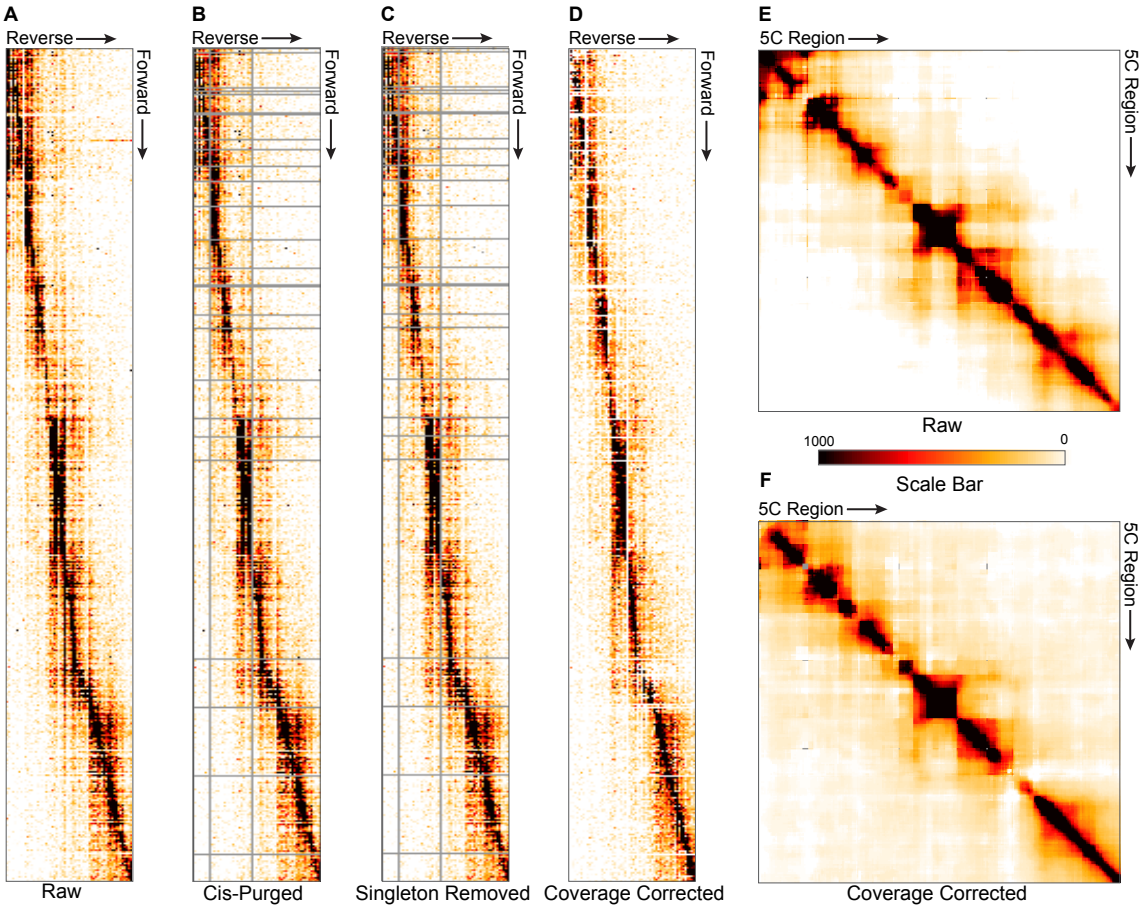
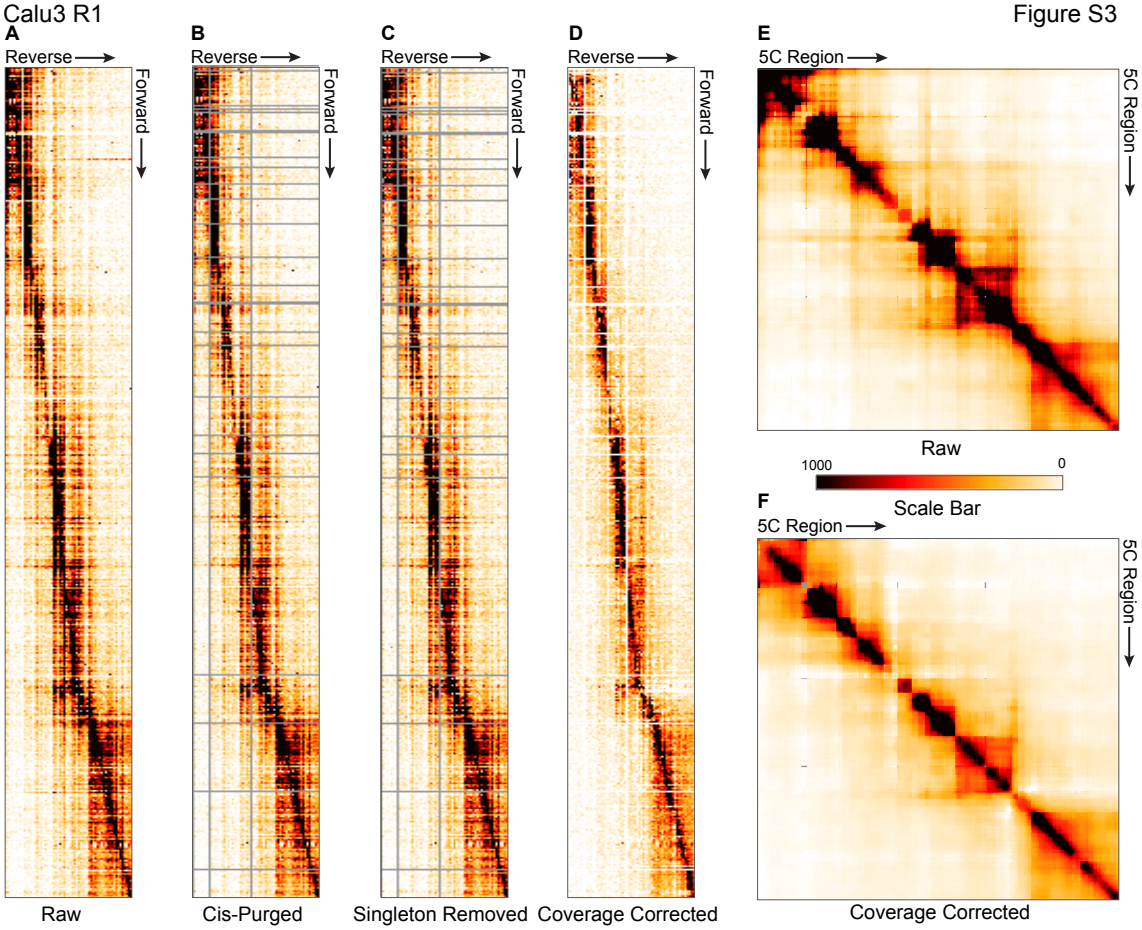
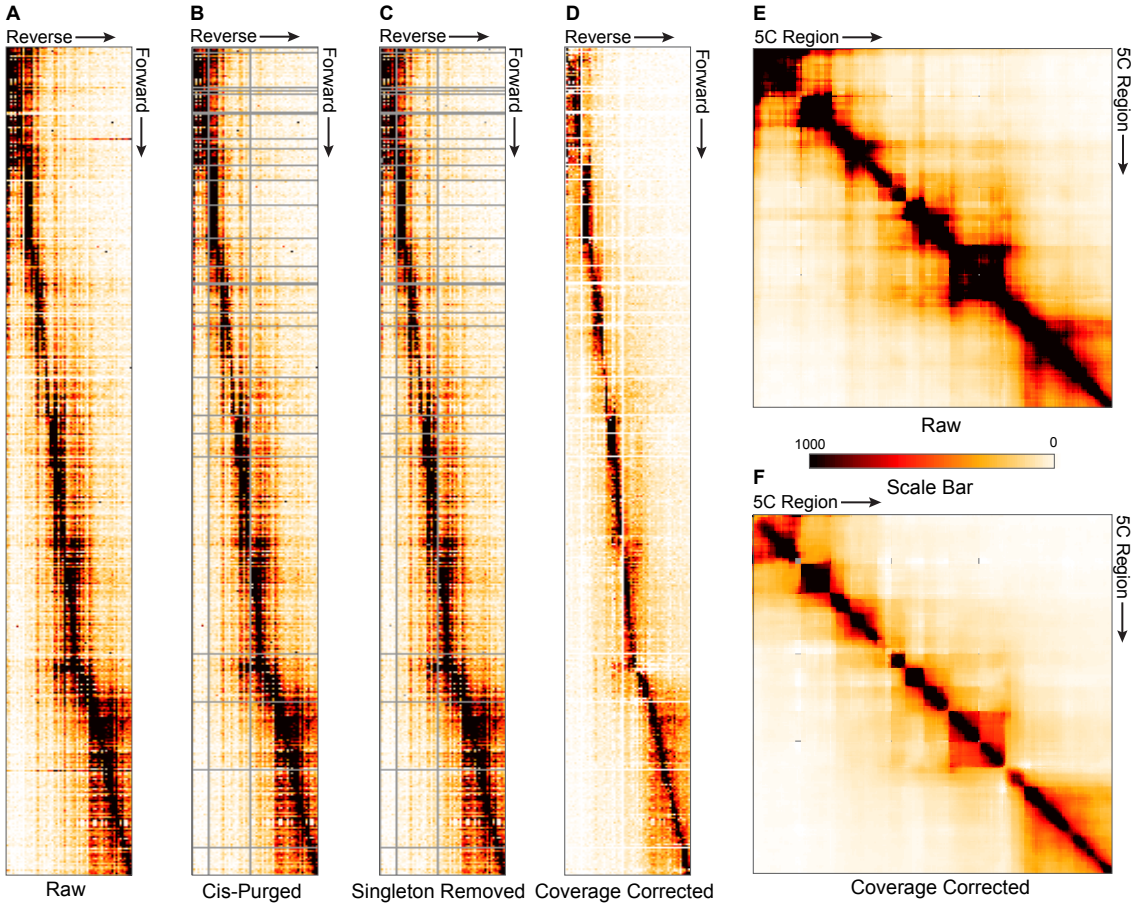


Figure S3


Calu3 R1
A
Reverse
Forward
Raw
B
Reverse
Forward
Cis-Purged
C
Reverse
Forward
Singleton Removed
D
Reverse
Forward
Coverage Corrected
E
5C Region
5C Region
Raw
1000
0
Scale Bar
F
5C Region
5C Region
Coverage Corrected

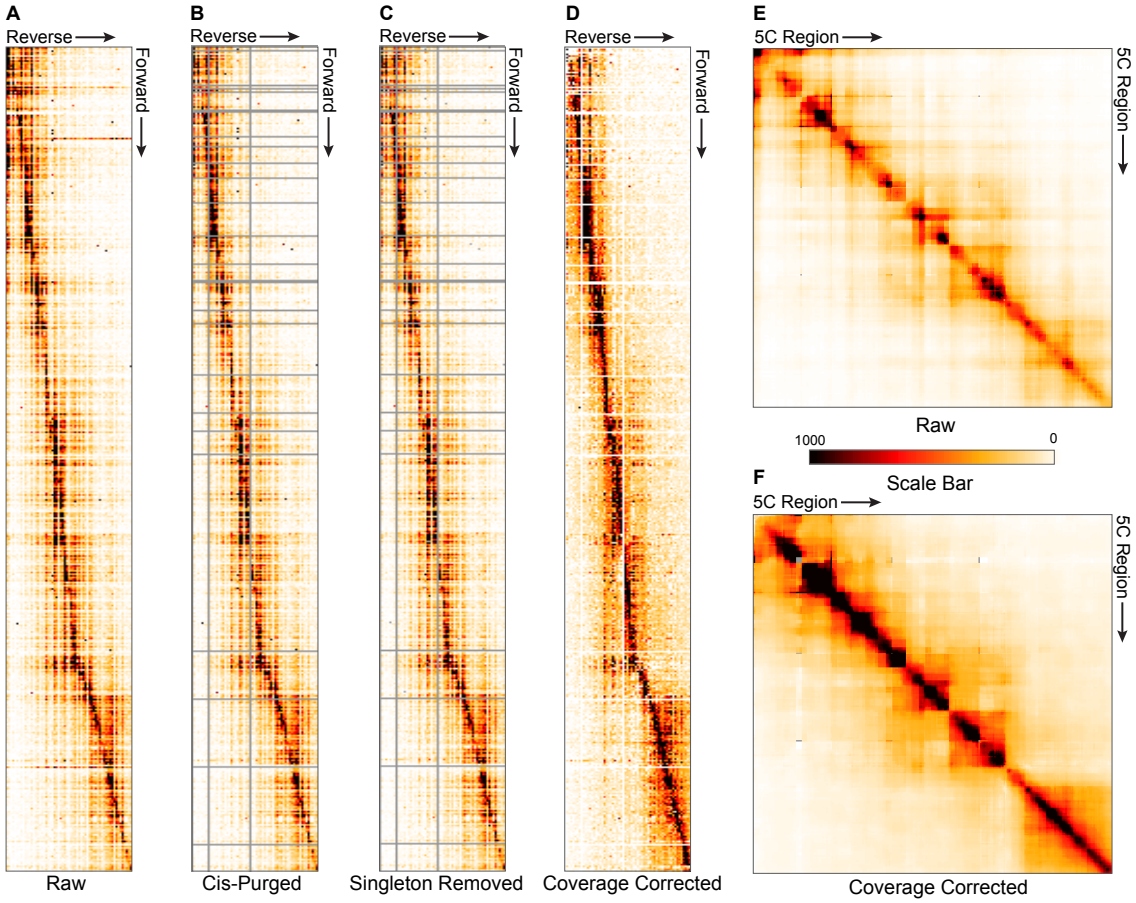
Calu3 R2

Figure S4



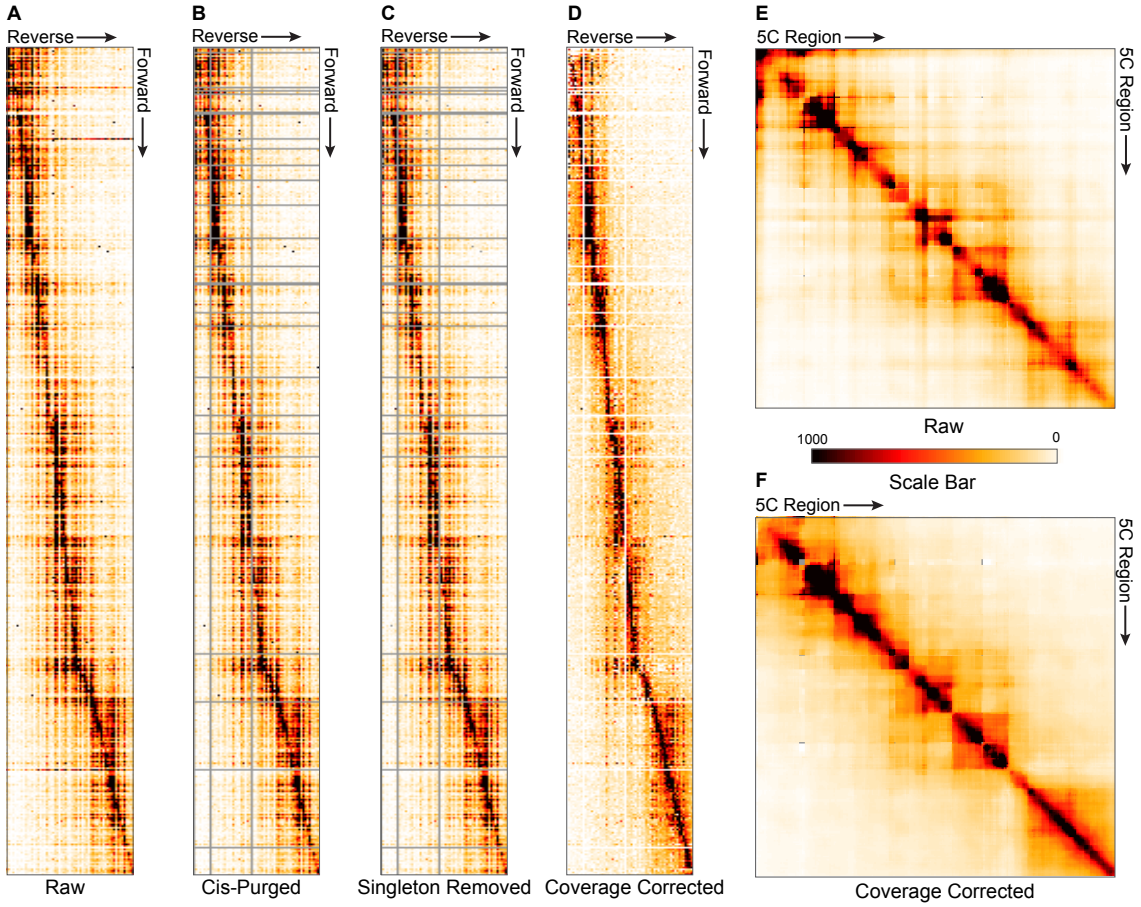
Capan1 R1

Figure S5



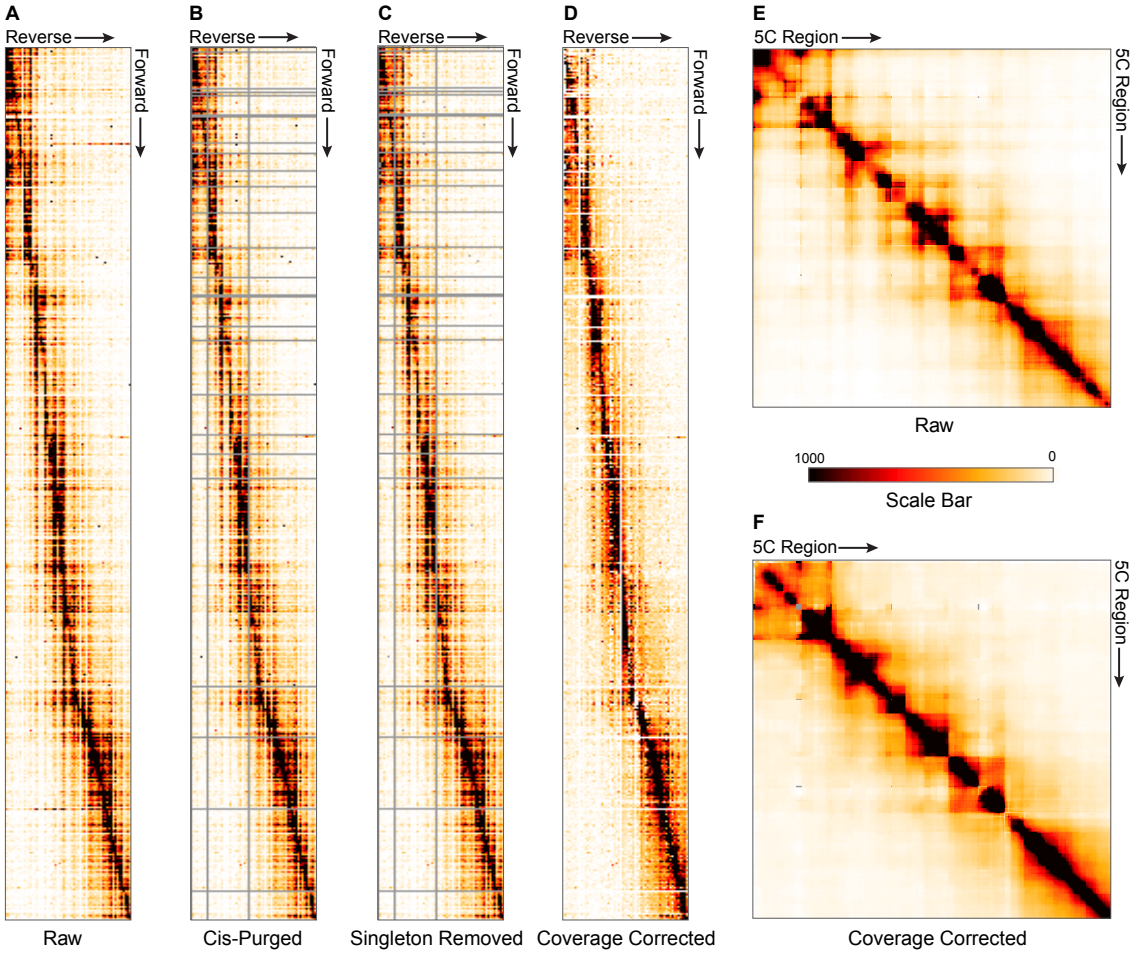
Capan1 R2

Figure S6



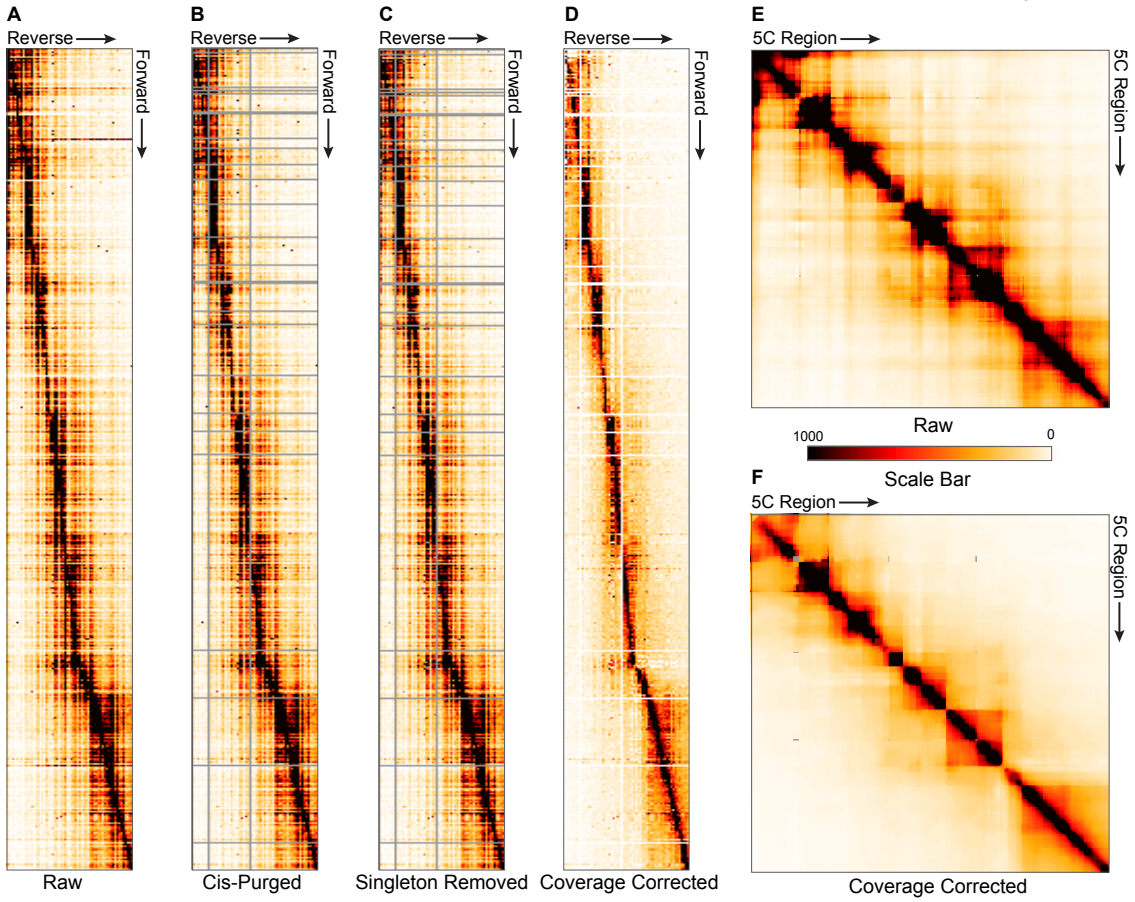
GM12878 R1

Figure S7



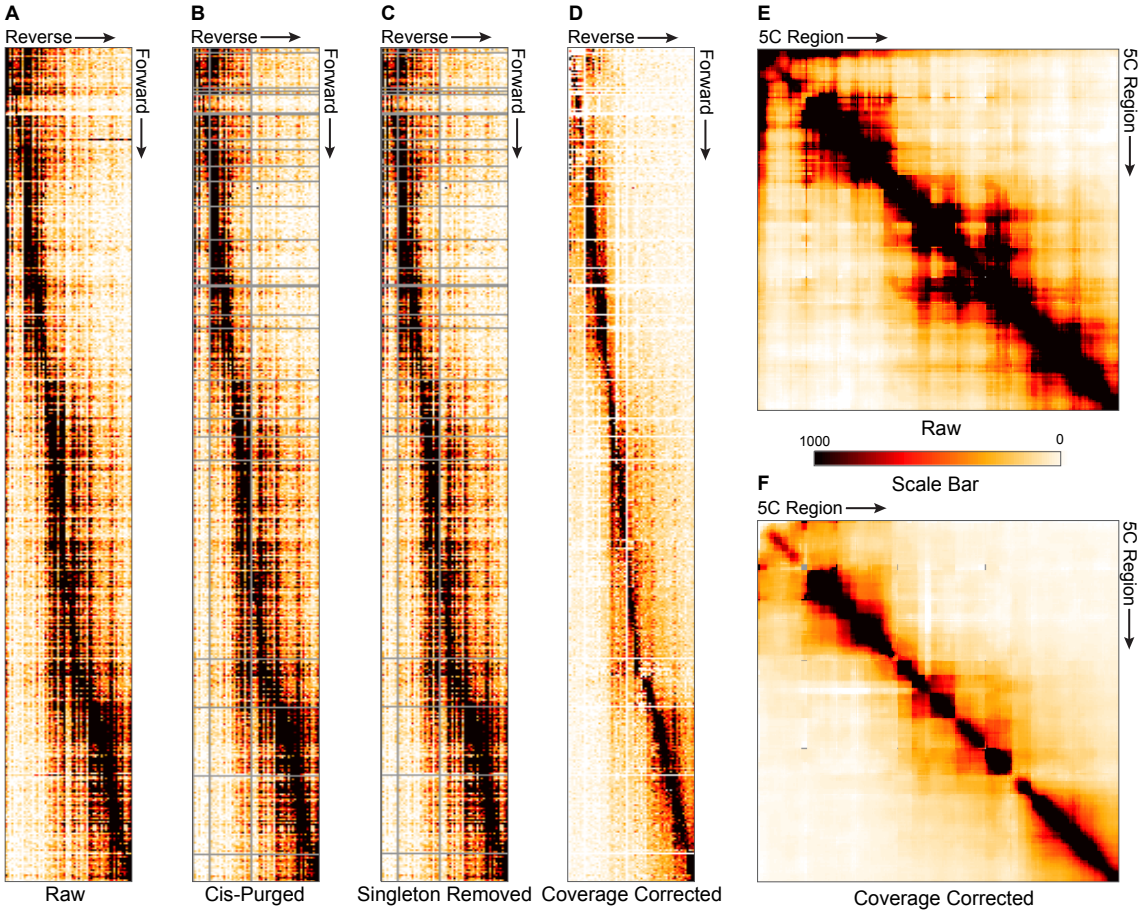
GM12878 R2

Figure S8



HepG2 R1

Figure S9



HepG2 R2

Figure S10

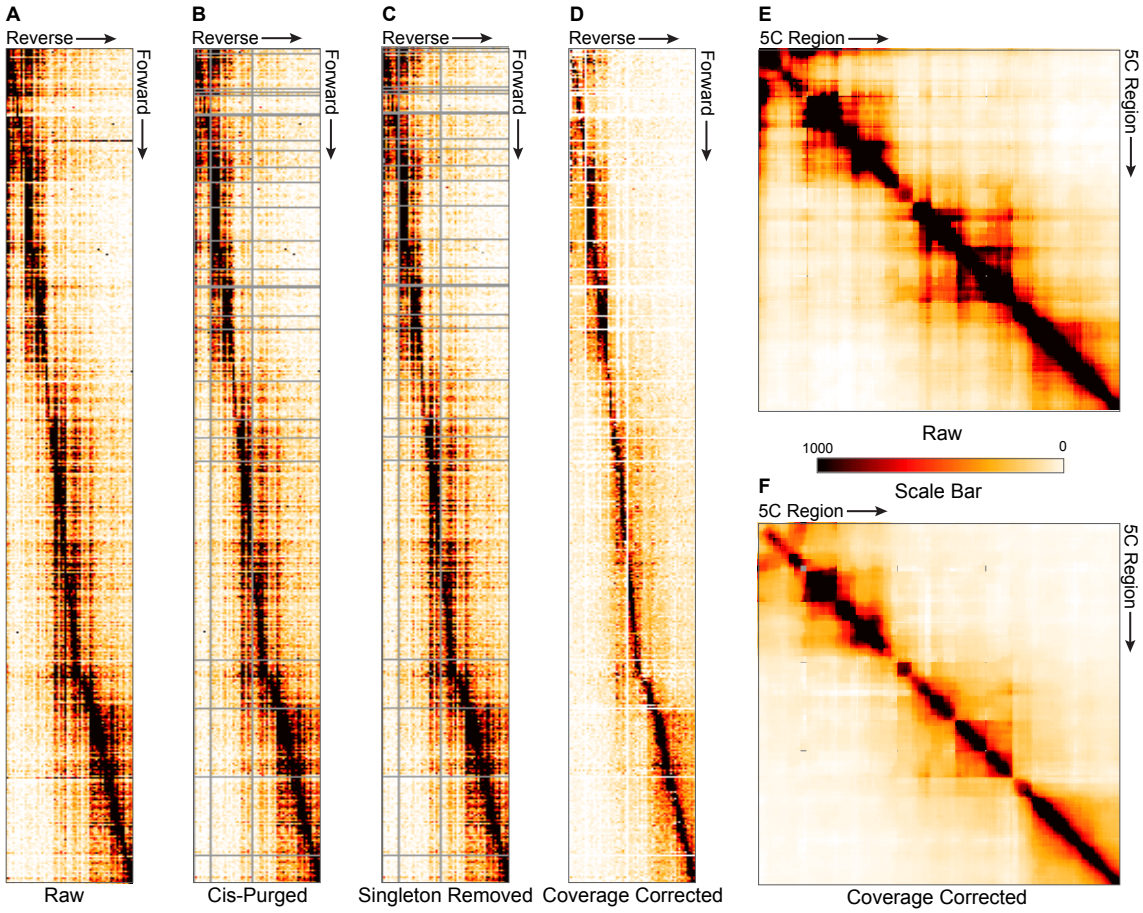
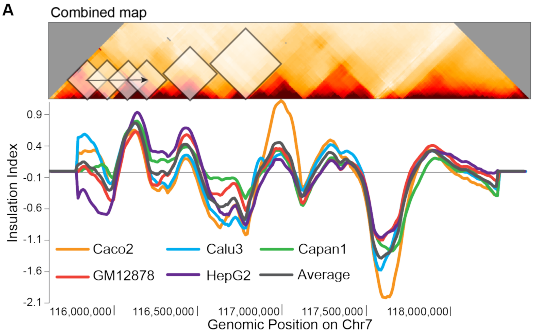


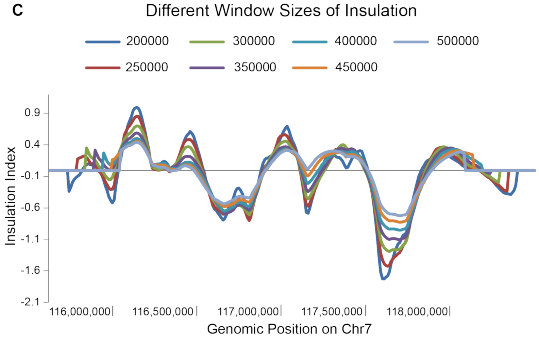
**Figure S11: Insulation Index of all cell lines and different window search spaces, and correlation with CTCF binding** A) The insulation index method is illustrated by sliding a diamond along the bottom of the triangular heatmap. Interactions within the diamond are summed and divided by the average sum for the region to calculate the insulation score of the bin at the bottom corner of the diamond. Below the heatmap is plotted the insulation index for all cell lines and the average, as in **Figure 2**. The size of the window over which can be changed, but does not affect the insulation profile (panel C). B) A) Insulation profiles run on the combined heatmap (all 10 datasets) using different diamond (window) sizes. As the size increases the index smoothens out but maintains the same peaks and minima. B). Browser shot (hg18) of the 2.8 Mb region studied here with the positions of the 7 TAD indicated (Top track, in black bars). Below are CTCF ChIP seq tracks (CTCF P = peaks, CTCF s= raw signal) for GM12878 and HepG2 cells. CTCF sites are within 10-20 Kb of TAD boundaries but are also found inside TADs. Also indicated are positions of all genes (UCSC gene track). CTCF ChIP- seq data were generated at the Broad Institute and in the lab of Dr. Bradley Bernstein at Massachusetts General Hospital/Harvard Medical School. This data is publicly available as part of the ENCODE dataset. B) Pearson correlations between the insulation profiles for the different cell types studies here. C) Pearson correlation of insulation profiles for the combined dataset calculated with the different diamond sizes.

Figure S11

A




C




B

| Pearson: Insulation | Caco2 | Calu3 | Capan1 | GM12878 | HepG2 | Average |
|---|---|---|---|---|---|---|
| Caco2 | 1 | 0.912 | 0.798 | 0.844 | 0.568 | 0.933 |
| Calu3 | | 1 | 0.819 | 0.866 | 0.745 | 0.944 |
| Capan1 | | | 1 | 0.876 | 0.811 | 0.926 |
| GM12878 | | | | 1 | 0.890 | 0.958 |
| HepG2 | | | | | 1 | 0.876 |
| Average | | | | | | 1 |

D

| Pearson: Windows | 200000 | 250000 | 300000 | 350000 | 400000 | 450000 | 500000 |
|---|---|---|---|---|---|---|---|
| 200000 | 1 | 0.982 | 0.953 | 0.925 | 0.896 | 0.877 | 0.865 |
| 250000 | | 1 | 0.989 | 0.968 | 0.942 | 0.922 | 0.903 |
| 300000 | | | 1 | 0.991 | 0.972 | 0.952 | 0.930 |
| 350000 | | | | 1 | 0.992 | 0.976 | 0.955 |
| 400000 | | | | | 1 | 0.993 | 0.978 |
| 450000 | | | | | | 1 | 0.995 |
| 500000 | | | | | | | 1 |

**Figure S12: Scaling plots for all cell lines, interactions and peaks.** For all plots: red represents inter-TAD interactions, blue intra-TAD interactions, and green all interactions. A) Scaling plot of all interactions for Caco2 cells. B) Scaling plot of all interactions for Calu3 cells C) Scaling plot of all interactions for Capan1 cells. D) C) Scaling plot of all interactions for GM12878 cells. E) D) Scaling plot of all interactions for HepG2 cells. F) E) Scaling plot of peaks significant interactions in Caco2 cells. G) F) Scaling plot of significant interactions in Capan1 cells. H) G) Scaling plot of significant interactions in Calu3 cells. H). I) Scaling plots of significant interactions in GM12878 cells. J) H) Scaling plots of significant interactions in HepG2 cells.

Figure S12

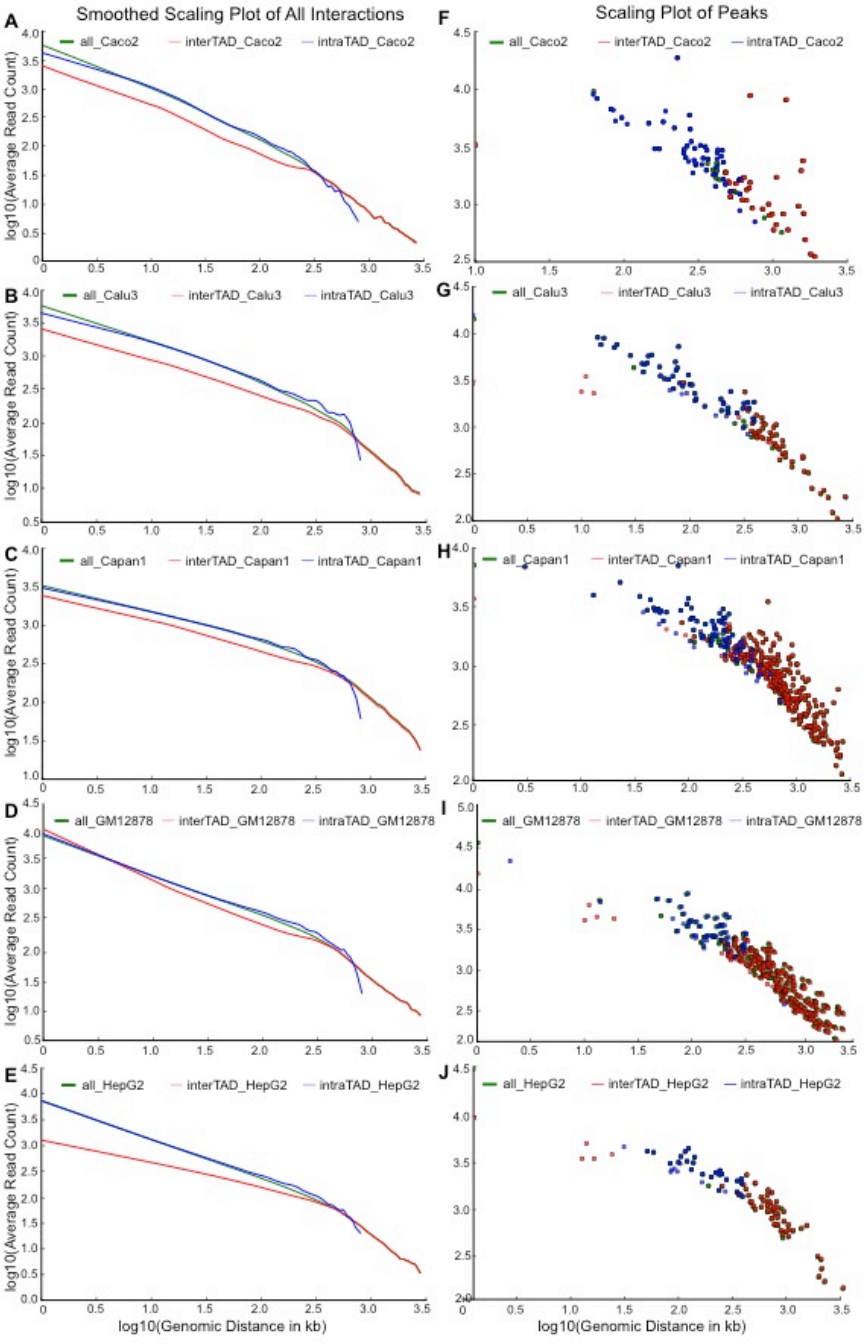

A
Smoothed Scaling Plot of All Interactions
all_Caco2
interTAD_Caco2
intraTAD_Caco2
log10(Average Read Count)
B
all_Calu3
interTAD_Calu3
intraTAD_Calu3
C
all_Capan1
interTAD_Capan1
intraTAD_Capan1
D
all_GM12878
interTAD_GM12878
intraTAD_GM12878
E
all_HepG2
interTAD_HepG2
intraTAD_HepG2
log10(Genomic Distance in kb)
F
Scaling Plot of Peaks
all_Caco2
interTAD_Caco2
intraTAD_Caco2
G
all_Calu3
interTAD_Calu3
intraTAD_Calu3
H
all_Capan1
interTAD_Capan1
intraTAD_Capan1
I
all_GM12878
interTAD_GM12878
intraTAD_GM12878
J
all_HepG2
interTAD_HepG2
intraTAD_HepG2
log10(Genomic Distance in kb)

**Figure S13: Interactions deemed significant in only one replicate have higher signal in the other replicate as compared to interactions deemed not significant in both replicates.** For all plots: The box plot on the far left represents z-scores for interactions called significant in both replicates. The blue box plot of peaks in replicate 1 represents those interactions in replicate 1 that were deemed significant, but were not significant in replicate 2. The green box shows the z-scores of those same interactions in the second replicate, where they were not counted significant. This coloring is reversed when we look at replicate 2. The blue box now indicates the z-score of interactions deemed significant in replicate 2 but not in replicate 1. The green box shows the z-scores of those same interactions in replicate 1, where they are not significant. The box plot on the far right shows the z-scores of all interactions in both replicates that were not significant. This is the same for all graphs.

Figure S13

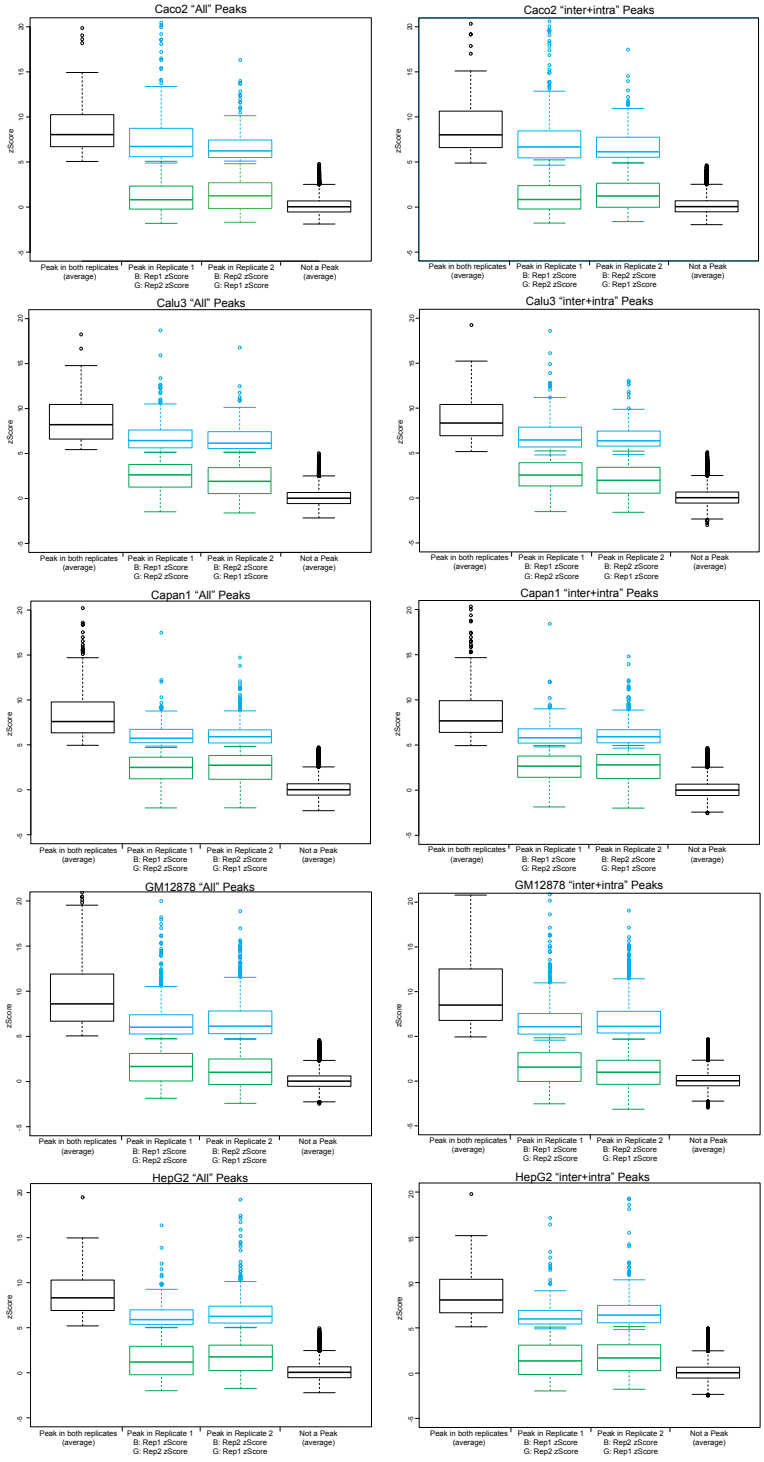

Caco2 "All" Peaks
Caco2 "inter+intra" Peaks
Calu3 "All" Peaks
Calu3 "inter+intra" Peaks
Capan1 "All" Peaks
Capan1 "inter+intra" Peaks
GM12878 "All" Peaks
GM12878 "inter+intra" Peaks
HepG2 "All" Peaks
HepG2 "inter+intra" Peaks
zScore
Peak in both replicates (average)
Peak in Replicate 1 B: Rep1 zScore G: Rep2 zScore
Peak in Replicate 2 B: Rep2 zScore G: Rep1 zScore
Not a Peak (average)

**Figure S14: Overlap of peak calling between replicates on the whole data set.**

Overlap of significant interactions in Caco2 replicates (A), Calu3 replicates (B), Capan1 replicates (C), GM12878 replicates (D) and HepG2 replicates (E).

Figure S14

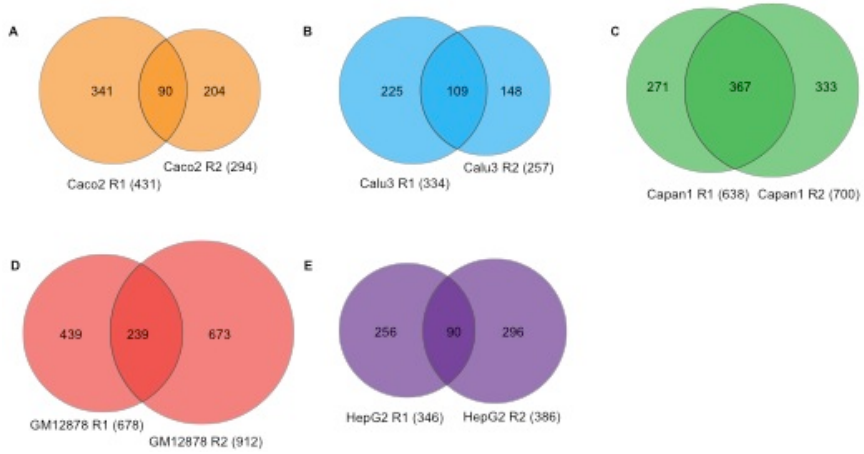

A
341
90
204
Caco2 R2 (294)
Caco2 R1 (431)
B
225
109
148
Calu3 R2 (257)
Calu3 R1 (334)
C
271
367
333
Capan1 R1 (638)
Capan1 R2 (700)
D
439
239
673
GM12878 R1 (678)
GM12878 R2 (912)
E
256
90
296
HepG2 R1 (346)
HepG2 R2 (386)

**Figure S15: Overlap of peak calling between replicates on the intraTAD dataset.**

Overlap of significant interactions in Caco2 replicates (A), Calu3 replicates (B), Capan1 replicates (C), GM12878 replicates (D) and HepG2 replicates (E).

Figure S15

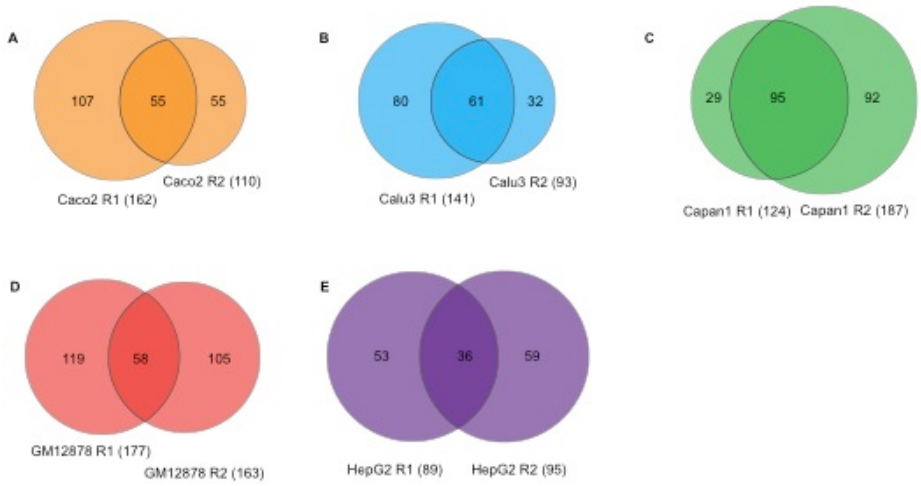

A
107
55
55
Caco2 R2 (110)
Caco2 R1 (162)
B
80
61
32
Calu3 R2 (93)
Calu3 R1 (141)
C
29
95
92
Capan1 R1 (124)
Capan1 R2 (187)
D
119
58
105
GM12878 R1 (177)
GM12878 R2 (163)
E
53
36
59
HepG2 R1 (89)
HepG2 R2 (95)

**Figure S16: Overlap of peak calling between replicates on the interTAD dataset.**

Overlap of significant interactions in Caco2 replicates (A), Calu3 replicates (B), Capan1 replicates (C), GM12878 replicates (D) and HepG2 replicates (E).

Figure S16

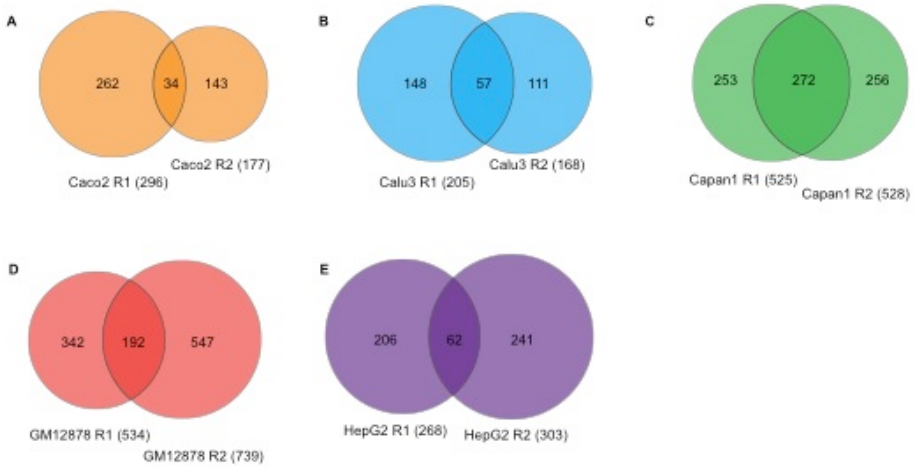

A
262
34
143
Caco2 R2 (177)
Caco2 R1 (296)
B
148
57
111
Calu3 R2 (168)
Calu3 R1 (205)
C
253
272
256
Capan1 R1 (525)
Capan1 R2 (528)
D
342
192
547
GM12878 R1 (534)
GM12878 R2 (739)
E
206
62
241
HepG2 R1 (268)
HepG2 R2 (303)

**Figure S17: Overlap of peaks called using the different spaces: whole, intra-TAD and inter-TAD, for all cell lines.** A) Overlap of peak calling methods in Caco2 cells. B) Overlap of peak calling methods in Calu3 cells. C) Overlap of peak calling methods in Capan1 cells. D) Overlap of peak calling methods in GM12878 cells. E) Overlap of peak calling methods in HepG2 cells. F) Diagram of peak calling space (same as Figure 4A).

Figure S17

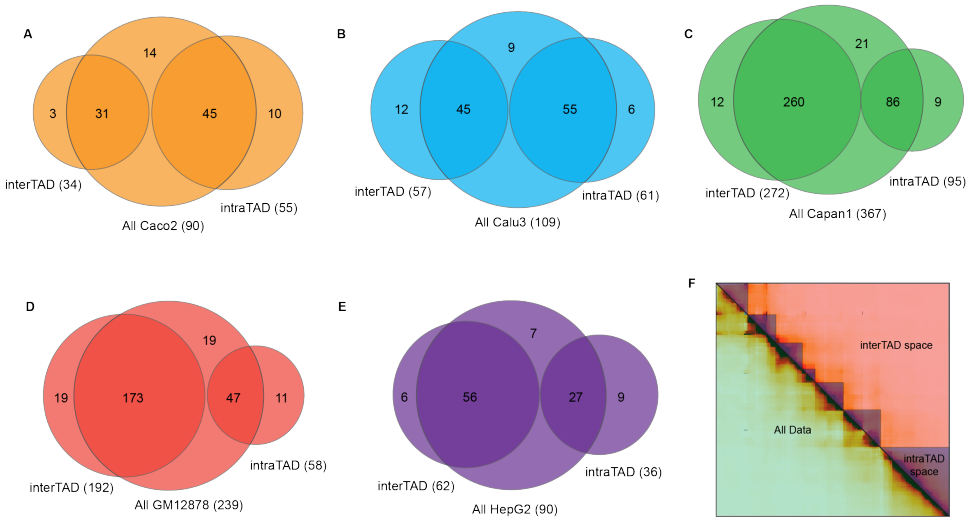

A
14
3
31
45
10
interTAD (34)
intraTAD (55)
All Caco2 (90)
B
9
12
45
55
6
interTAD (57)
intraTAD (61)
All Calu3 (109)
C
21
12
260
86
9
interTAD (272)
intraTAD (95)
All Capan1 (367)
D
19
19
173
47
11
intraTAD (58)
interTAD (192)
All GM12878 (239)
E
7
6
56
27
9
interTAD (62)
intraTAD (36)
All HepG2 (90)
F
interTAD space
All Data
intraTAD space

**Figure S18: Significant interactions of the *CAV2* promoter.** A genome browser snapshot of the region is above each profile view. In each anchor plot: Corrected read count is plotted versus genomic distance. The yellow bar represents the location of the anchoring interaction. The dotted black lines represent the loess standard deviation, and the thick black line represents the loess average. The red line represents the interactions of the genomic fragment indicated by the yellow line with all other interrogated fragments in the region. Significant interactions are indicated by large circles.

Figure S18

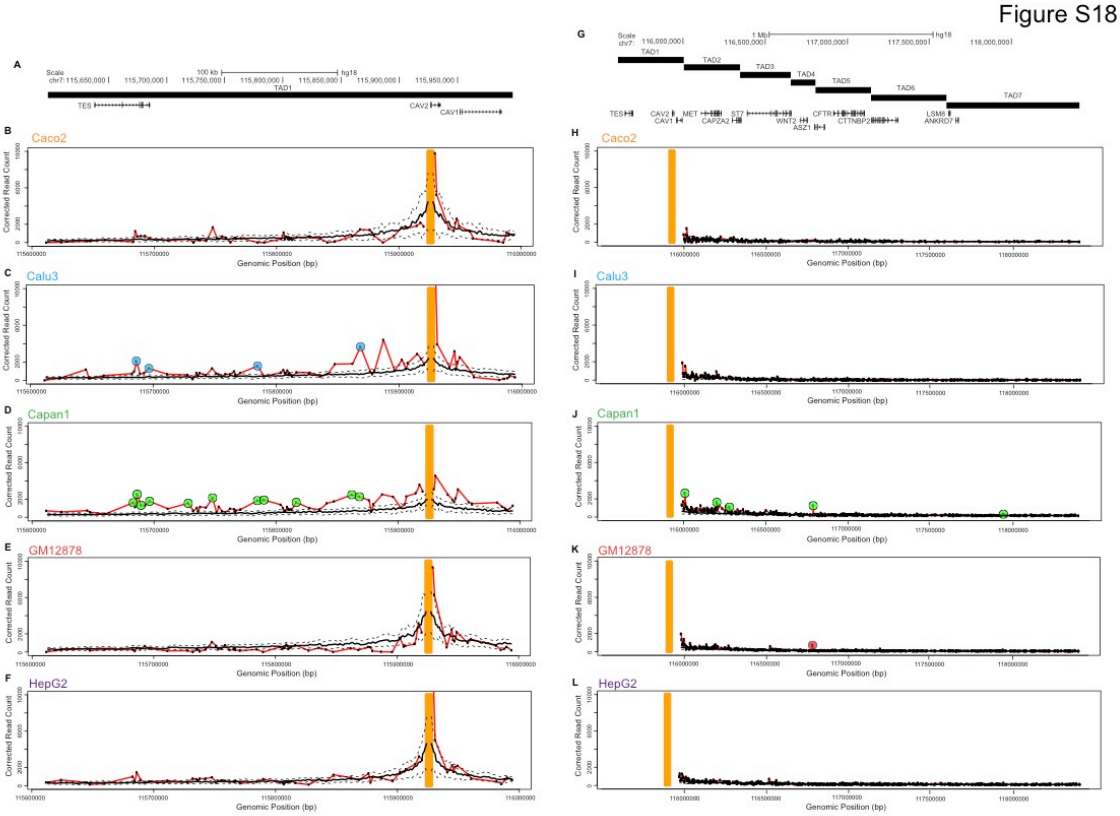

A
Scale
chr7: 115,650,000
115,700,000
100 kb
115,750,000
115,800,000
115,850,000
hg18
115,900,000
115,950,000
TAD1
TES
CAV2
CAV1
B
Caco2
C
Calu3
D
Capan1
E
GM12878
F
HepG2
Corrected Read Count
Genomic Position (bp)
G
Scale
chr7: 116,000,000
1 Mb
116,500,000
117,000,000
117,500,000
hg18
118,000,000
TAD1
TAD2
TAD3
TAD4
TAD5
TAD6
TAD7
TES
CAV2
CAV1
MET
CAPZA2
ST7
WNT2
ASZ1
CFTR
CTTNBP2
LSM8
ANKRD7
H
Caco2
I
Calu3
J
Capan1
K
GM12878
L
HepG2
Corrected Read Count
Genomic Position (bp)

**Figure S19: Significant interactions of *MET* and *CAPZA2* promoters.** A genome browser snapshot of the region is above each profile view. In each anchor plot: Corrected read count is plotted versus genomic distance. The yellow bar represents the location of the anchoring interaction. The dotted black lines represent the loess standard deviation, and the thick black line represents the loess average. The red line represents the interactions of the genomic fragment indicated by the yellow line with all other interrogated fragments in the region. Significant interactions are indicated by large circles.

Figure S19

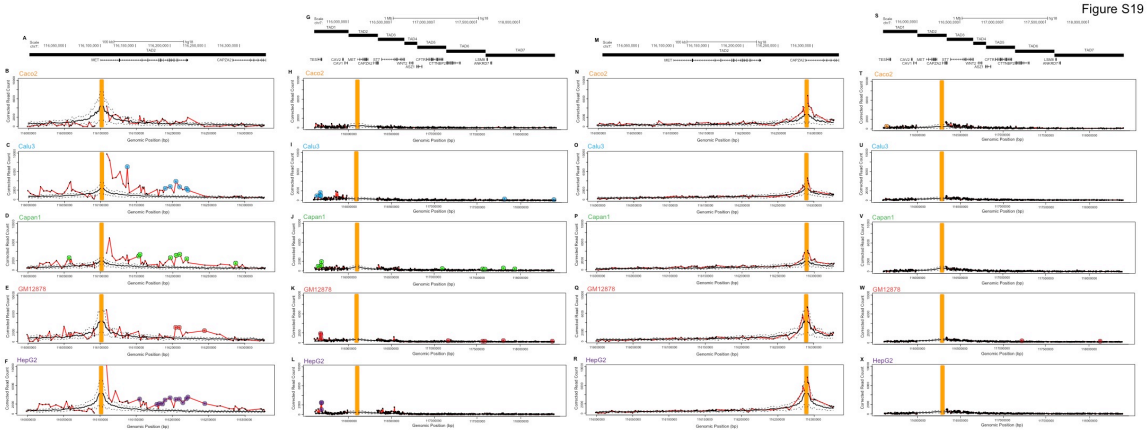

A
G
M
S
B
H
N
T
Caco2
C
I
O
U
Calu3
D
J
P
V
Capan1
E
K
Q
W
GM12878
F
L
R
X
HepG2
Corrected Read Count
Genomic Position (bp)

**Figure S20: Significant interactions of the *ST7* promoter.** A genome browser snapshot of the region is above each profile view. In each anchor plot: Corrected read count is plotted versus genomic distance. The yellow bar represents the location of the anchoring interaction. The dotted black lines represent the loess standard deviation, and the thick black line represents the loess average. The red line represents the interactions of the genomic fragment indicated by the yellow line with all other interrogated fragments in the region. Significant interactions are indicated by large circles.

Figure S20

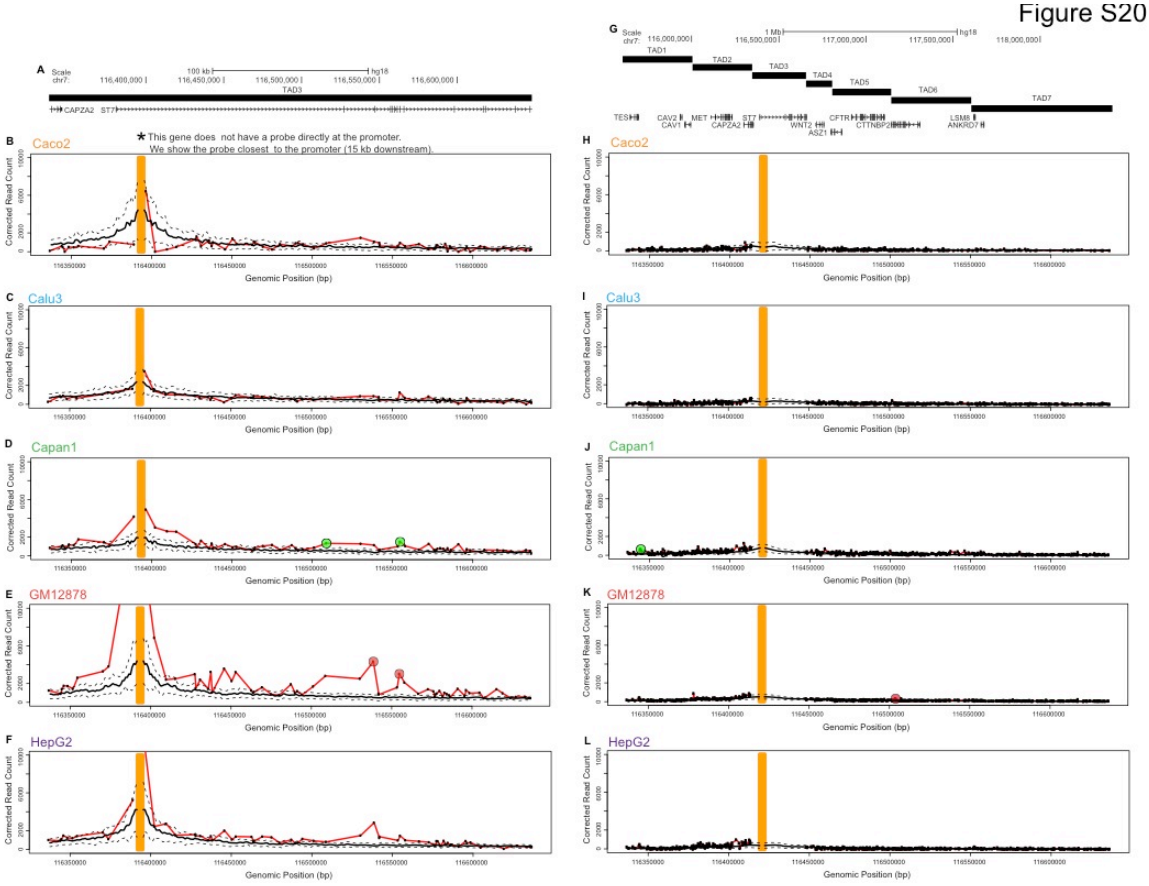

A
Scale
chr7:
116,400,000
100 kb
116,450,000
hg18
116,500,000
116,550,000
116,600,000
TAD3
CAPZA2
ST7
G
Scale
chr7:
116,000,000
1 Mb
116,500,000
hg18
117,000,000
117,500,000
118,000,000
TAD1
TAD2
TAD3
TAD4
TAD5
TAD6
TAD7
TES
CAV2
CAV1
MET
CAPZA2
ST7
WNT2
ASZ1
CFTR
CTTNBP2
LSM8
ANKRD7
B
Caco2
This gene does not have a probe directly at the promoter.
We show the probe closest to the promoter (15 kb downstream).
C
Calu3
D
Capan1
E
GM12878
F
HepG2
H
Caco2
I
Calu3
J
Capan1
K
GM12878
L
HepG2
Corrected Read Count
Genomic Position (bp)

**Figure S21: Significant interactions of the *WNT2* promoter.** A genome browser snapshot of the region is above each profile view. In each anchor plot: Corrected read count is plotted versus genomic distance. The yellow bar represents the location of the anchoring interaction. The dotted black lines represent the loess standard deviation, and the thick black line represents the loess average. The red line represents the interactions of the genomic fragment indicated by the yellow line with all other interrogated fragments in the region. Significant interactions are indicated by large circles.

Figure S21

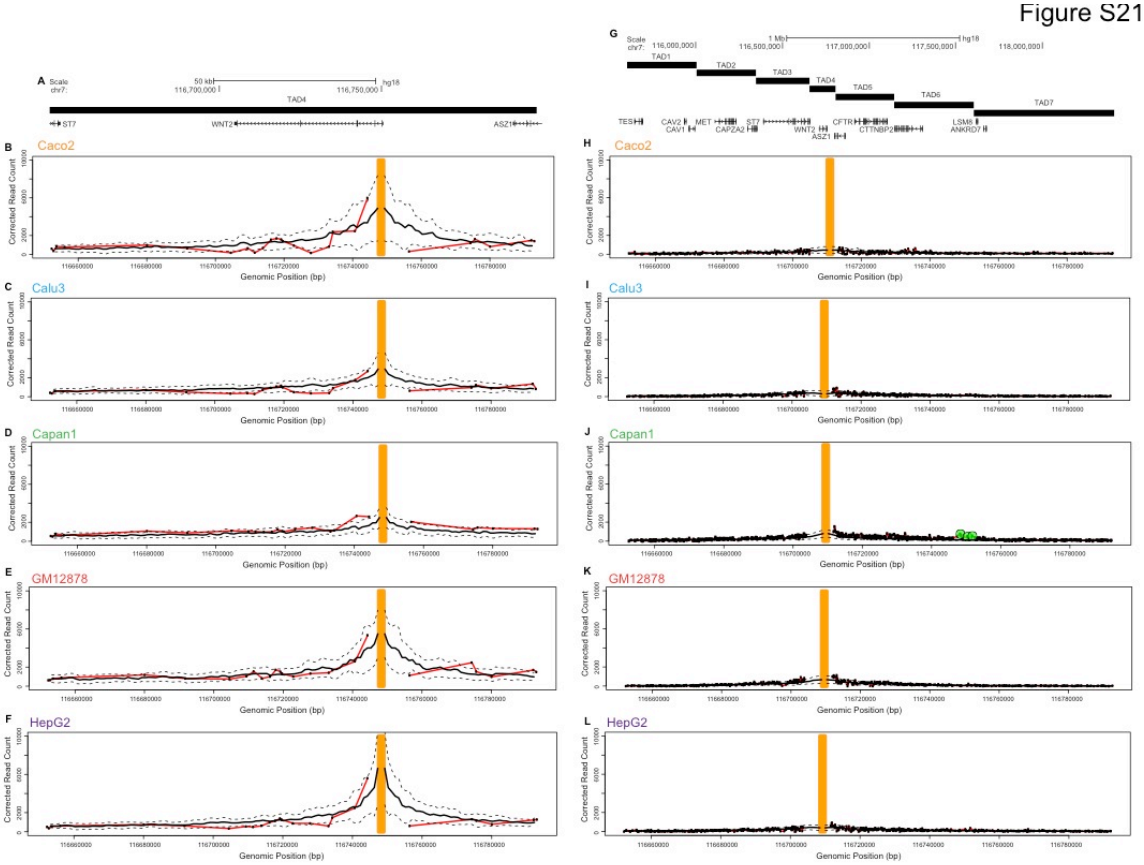

A
Scale
chr7:
50 kb
116,700,000
116,750,000
hg18
TAD4
ST7
WNT2
ASZ1
G
Scale
chr7:
116,000,000
1 Mb
116,500,000
117,000,000
117,500,000
hg18
118,000,000
TAD1
TAD2
TAD3
TAD4
TAD5
TAD6
TAD7
TESK1
CAV2
CAV1
MET
CAPZA2
ST7
WNT2
ASZ1
CFTR
CTTNBP2
LSM6
ANKRD7
B
Caco2
C
Calu3
D
Capan1
E
GM12878
F
HepG2
H
Caco2
I
Calu3
J
Capan1
K
GM12878
L
HepG2
Corrected Read Count
Genomic Position (bp)

**Figure S22: Significant interactions of *ASZ1* and *CFTR* promoters.** A genome browser snapshot of the region is above each profile view. In each anchor plot: Corrected read count is plotted versus genomic distance. The yellow bar represents the location of the anchoring interaction. The dotted black lines represent the loess standard deviation, and the thick black line represents the loess average. The red line represents the interactions of the genomic fragment indicated by the yellow line with all other interrogated fragments in the region. Significant interactions are indicated by large circles.

Figure S22

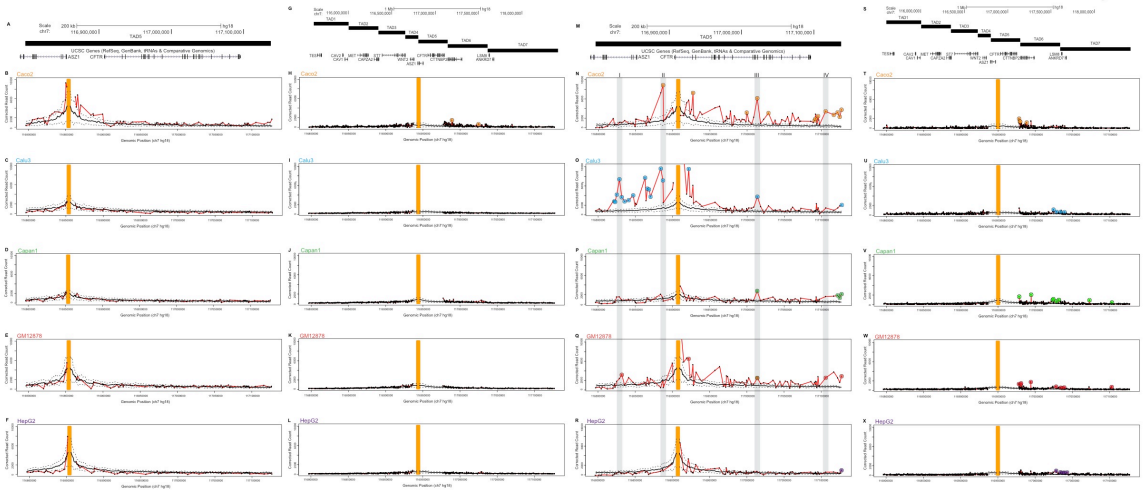


**Figure S23: Significant interactions of the *CTTNBP2* promoter within TAD6.** A genome browser snapshot of the region is above each profile view. In each anchor plot: Corrected read count is plotted versus genomic distance. The yellow bar represents the location of the anchoring interaction. The dotted black lines represent the loess standard deviation, and the thick black line represents the loess average. The red line represents the interactions of the genomic fragment indicated by the yellow line with all other interrogated fragments in the region. Significant interactions are indicated by large circles.

Figure S23

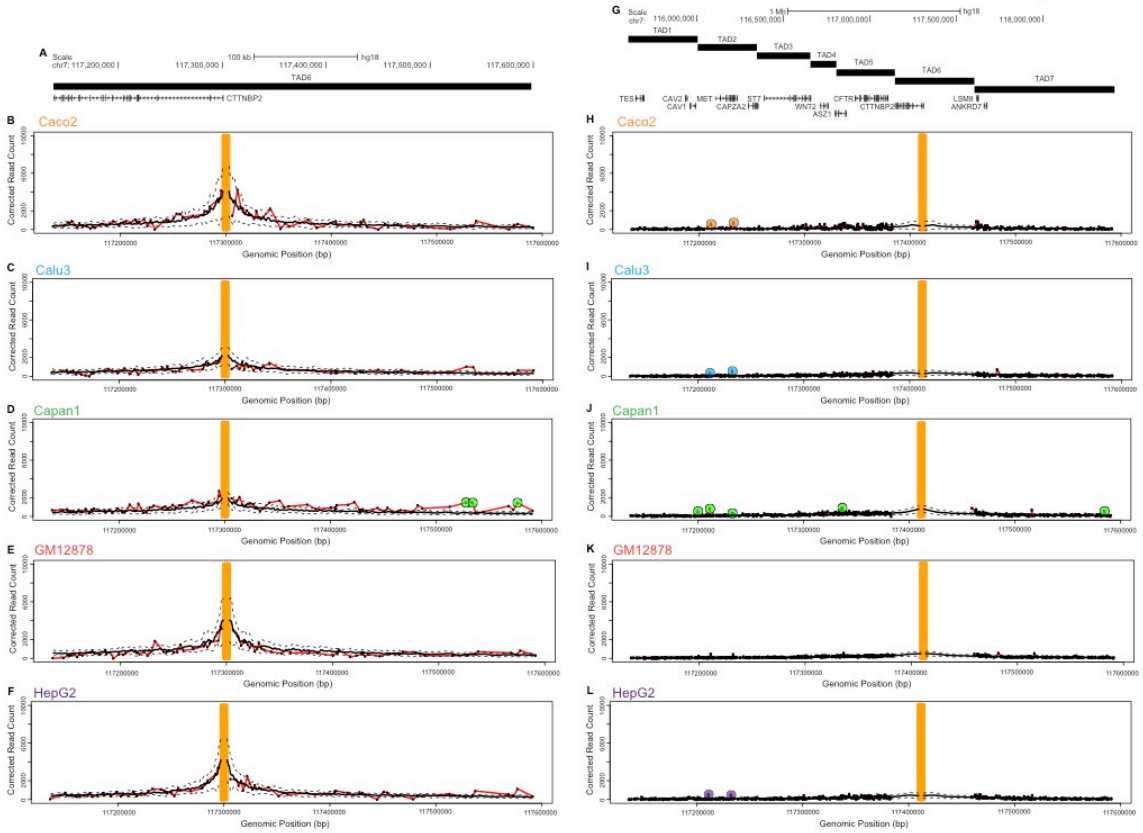
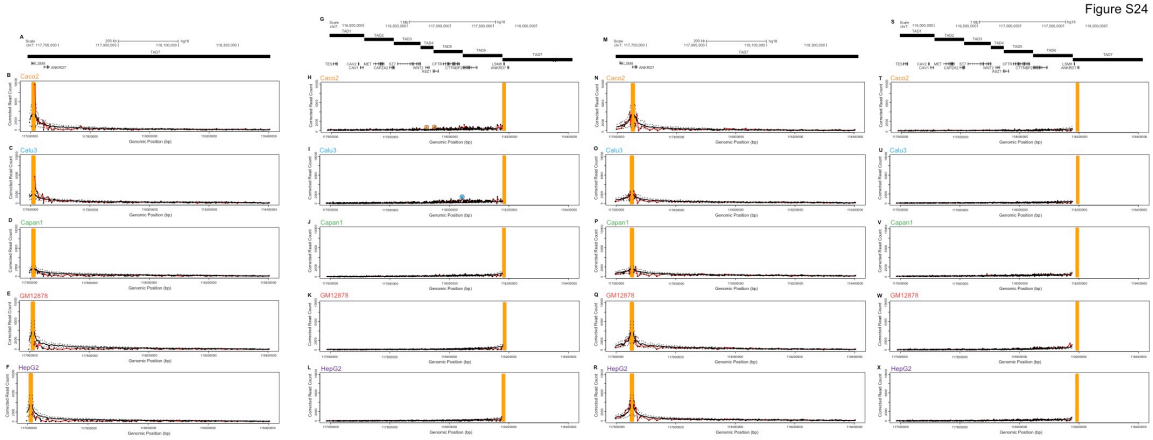

A
Scale
chr7: 117,200,000
117,300,000
100 kb
117,400,000
hg18
117,500,000
117,600,000
TAD6
CTTNBP2
G
Scale
chr7:
116,000,000
1 Mb
116,500,000
117,000,000
117,500,000
hg18
118,000,000
TAD1
TAD2
TAD3
TAD4
TAD5
TAD6
TAD7
TES
CAV2
CAV1
MET
CAPZA2
ST7
WNT2
ASZ1
CFTR
CTTNBP2
LSM8
ANKRD7
B
Caco2
C
Calu3
D
Capan1
E
GM12878
F
HepG2
H
Caco2
I
Calu3
J
Capan1
K
GM12878
L
HepG2
Corrected Read Count
Genomic Position (bp)

**Figure S24: Significant interactions of the *LSM8* and Ankyrin Repeat Domain- containing Protein 7 (*ANKRD7* [MIM 610731]) promoters.** A genome browser snapshot of the region is above each profile view. In each anchor plot: Corrected read count is plotted versus genomic distance. The yellow bar represents the location of the anchoring interaction. The dotted black lines represent the loess standard deviation, and the thick black line represents the loess average. The red line represents the interactions of the genomic fragment indicated by the yellow line with all other interrogated fragments in the region. Significant interactions are indicated by large circles.

Figure S24


A
G
M
S
B
Caco2
H
N
T
C
Calu3
I
O
U
D
Capan1
J
P
V
E
GM12878
K
Q
W
F
HepG2
L
R
X
Corrected Read Count
Genomic Position (bp)