# Supporting Information

## Hou et al. 10.1073/pnas.0912087107

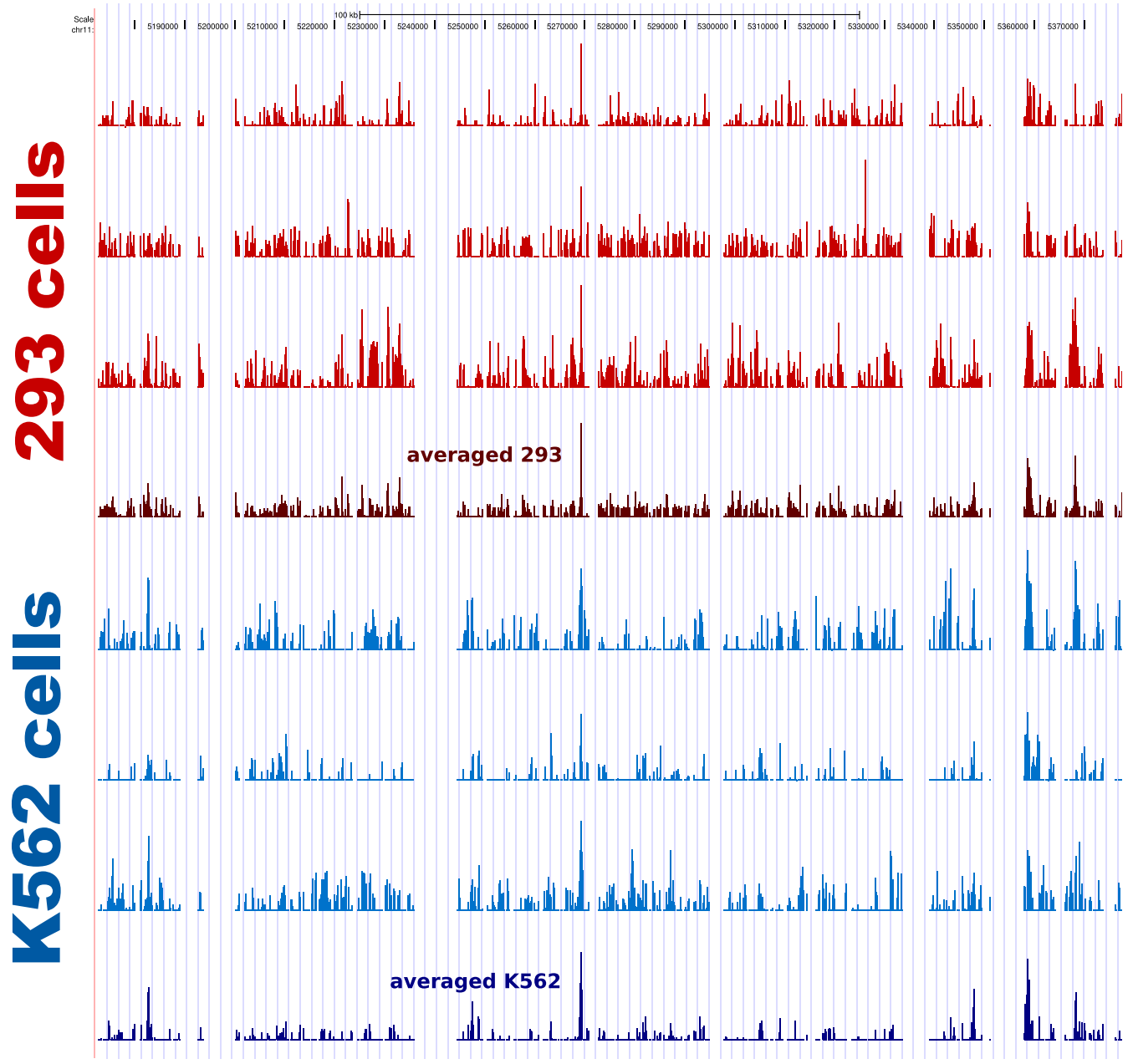


**Fig. S1.** CTCF localization on chr11:5,180,000–5,380,000 in K562 and 293T cells was derived from three independent biological replicates of ChIP-chip experiments. Raw data are shown for 293T cells in red and K562 cells in blue. Averaged data are shown below for each cell type. The averaged data were used to predict CTCF peaks with the ACME algorithm optimized as described in the main text to a training set of sites validated by ChIP.

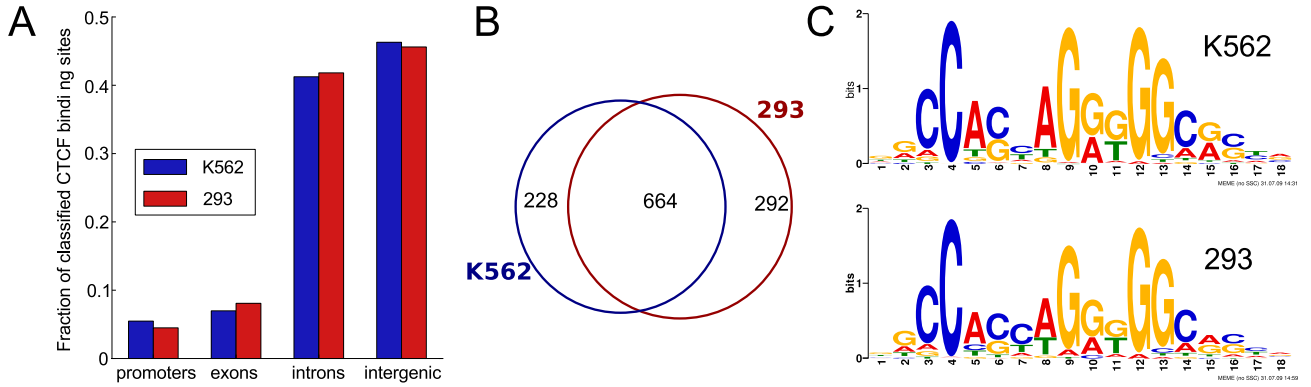


**Fig. S2.** (*A*) Bars indicate the percentages of predicted peaks from ACME analysis of CTCF ChIP-chip data that are located in promoter, exon, intron, or intergenic regions in each cell type. In agreement with earlier work, 46% of CTCF peaks on chromosomes 10 and 11 were in intergenic regions whereas a relatively higher number of interactions (about 50% compared to 35% previously reported) were in exons and introns (1–3). This result may reflect the gene-rich nature of chromosome 11 because CTCF sites closely follow the distribution of genes in other work (1, 3, 4). (*B*) A Venn diagram representing the shared and unique predicted peaks for K562 and 293T cells. Substantial (70–75%), but not complete, overlap of CTCF peaks was observed in the two cell types. (*C*) The MEME-derived (5) statistically overrepresented CTCF motif shared by the called sites in K562 cells and 293T cells agrees substantially with data published for other cell lines (1, 3), indicating that as little as 2% of the genome can provide a representative sample of a protein-binding motif.


1. Kim TH, et al. (2007) Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. *Cell* 128:1231–1245.
2. Barski A, et al. (2007) High-resolution profiling of histone methylations in the human genome. *Cell* 129:823–837.
3. Cuddapah S, et al. (2009) Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. *Genome Res* 19: 24–32.
4. Xie X, et al. (2007) Systematic discovery of regulatory motifs in conserved regions of the human genome, including thousands of CTCF insulator sites. *Proc Natl Acad Sci USA* 104: 7145–7150.
5. Bailey TL, Williams N, Misleh C, Li WW (2006) MEME: Discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Res* 34 (Web Server issue):W369–W373.


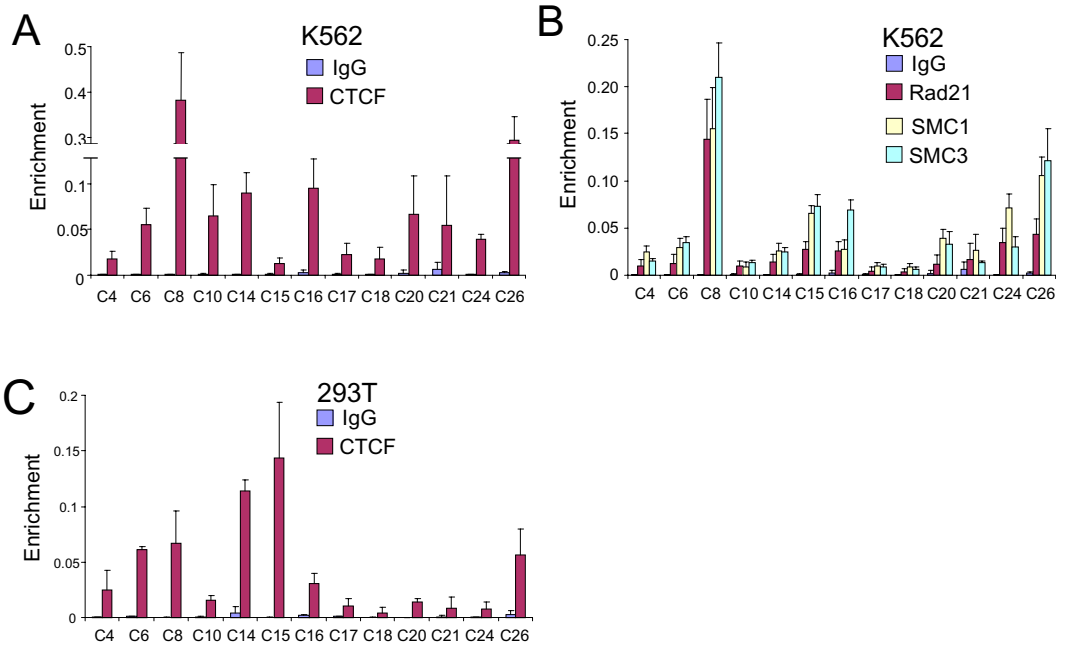


**Fig. S3.** (*A*) ChIP and real-time qPCR analysis confirm in vivo CTCF binding in K562 cells at the 13 sites most closely surrounding the β-globin locus. The enrichment varied over a wide range, but all sites had an at least an 8- to 10-fold signal over the IgG control. (*B*) ChIP experiments using K562 cells with antibodies against Rad21, SMC1, and SMC3, three members of the cohesin complex, indicate that these components are enriched at all 13 CTCF-binding sites in K562 cells. (*C*) Real-time qPCR confirmation of occupancy of predicted CTCF sites in 293T cells determined using a CTCF antibody. Three independent chromatin preparations were studied for each experiment. The error bars represent SEM.

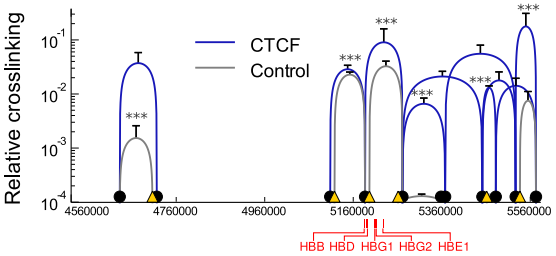


**Fig. S4.** Specificity of CTCF site long-range interactions. 3C was carried out with K562 cell chromatin. Specific interactions between and among CTCF sites (black circles) are depicted by blue curves with the height of the curve corresponding to the cross-linking frequency. A primer for each interacting site was next paired with one from the non-CTCF fragment adjacent to its partner CTCF site (yellow triangles); reduced interaction frequencies (20% or less) between CTCF sites and non-CTCF sites are depicted as gray curves. Cross-linking is plotted relative to the signal for two fragments in the α-tubulin gene. Note the log scale of the *y* axis. Error bars represent SD. ***$P < 0.001$.

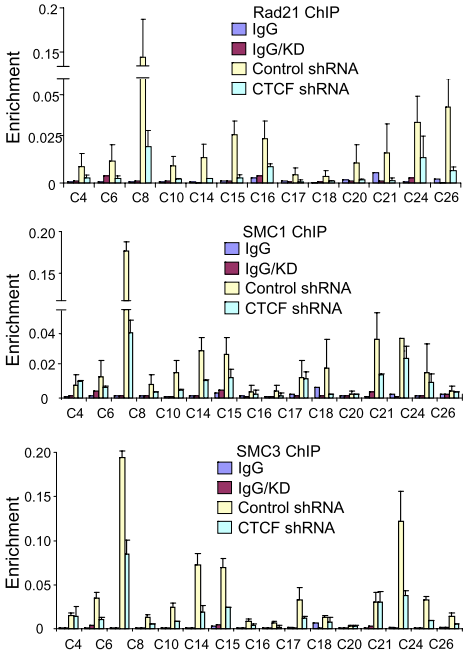


**Fig. S5.** ChIP analysis of cohesin binding after knock-down of CTCF by shRNA. K562 cells were transduced with a control shRNA or an shRNA directed to CTCF. ChIP was performed with antibodies to Rad21, SMC1, or SMC3. The results of three chromatin preparations are shown ±SEM.

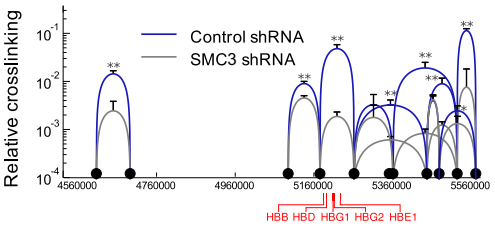


**Fig. S6.** SMC3 knock-down results in loss of CTCF site interactions. The 3C assay was carried out using chromatin from K562 cells transduced with a control shRNA or an shRNA directed to SMC3. Interactions between and among CTCF sites determined by 3C are indicated by blue curves, and reduction of these interactions after knock-down of SMC3 by shRNA is indicated by gray curves. Cross-linking is relative to the signal for two fragments in the α-tubulin gene. The results of three chromatin preparations are shown ±SD. **$P < 0.01$, *$P < 0.05$.

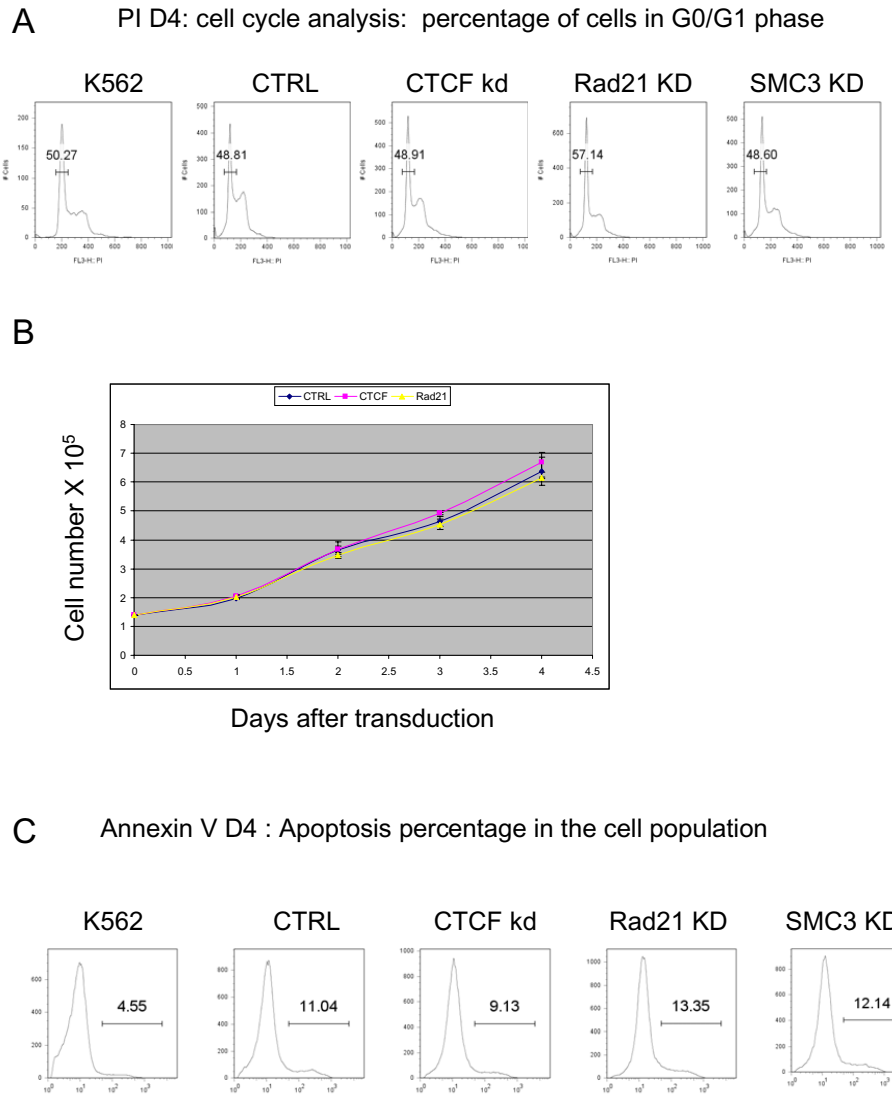


**Fig. S7.** CTCF or cohesin knock-down does not affect K562 cell cycle parameters. (*A*) Cells were stained with propidium iodide and sorted by FACS to investigate whether loss of interaction among CTCF sites after reduction of cohesins or CTCF was influenced by a change in cell cycle distribution. There was no obvious change in the distribution of cells within the cell cycle at day 4 after CTCF or cohesin reduction, the time point at which the 3C experiments were carried out. (*B*) CTCF or cohesin knock-down does not slow K562 cell proliferation over the 4-day time course. (*C*) CTCF or cohesin knock-down does not cause apoptosis in K562 cells. Note that apoptosis begins to dramatically increase after day 4 of shRNA treatment in cells with cohesin subunit knock-downs.

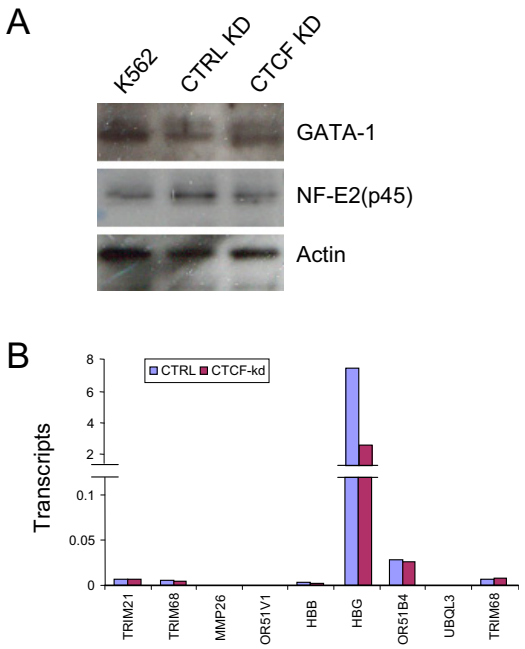


**Fig. S8.** (*A*) Western blot analysis was carried out with lysates of K562 cells or with cells transduced with a control shRNA or an shRNA directed to CTCF for 4 days. Antibodies to GATA-1 and NF-E2(p45), two important erythroid transactivators, were used. Actin served as the loading control. No obvious change in protein levels was observed. (*B*) An RT–PCR analysis of selected genes in regions flanking the globin locus was performed after K562 cells were transduced with a control shRNA or an shRNA directed against CTCF. No change in expression of these silent genes was observed after CTCF reduction.

## Other Supporting Information Files

Table S1 (DOC)