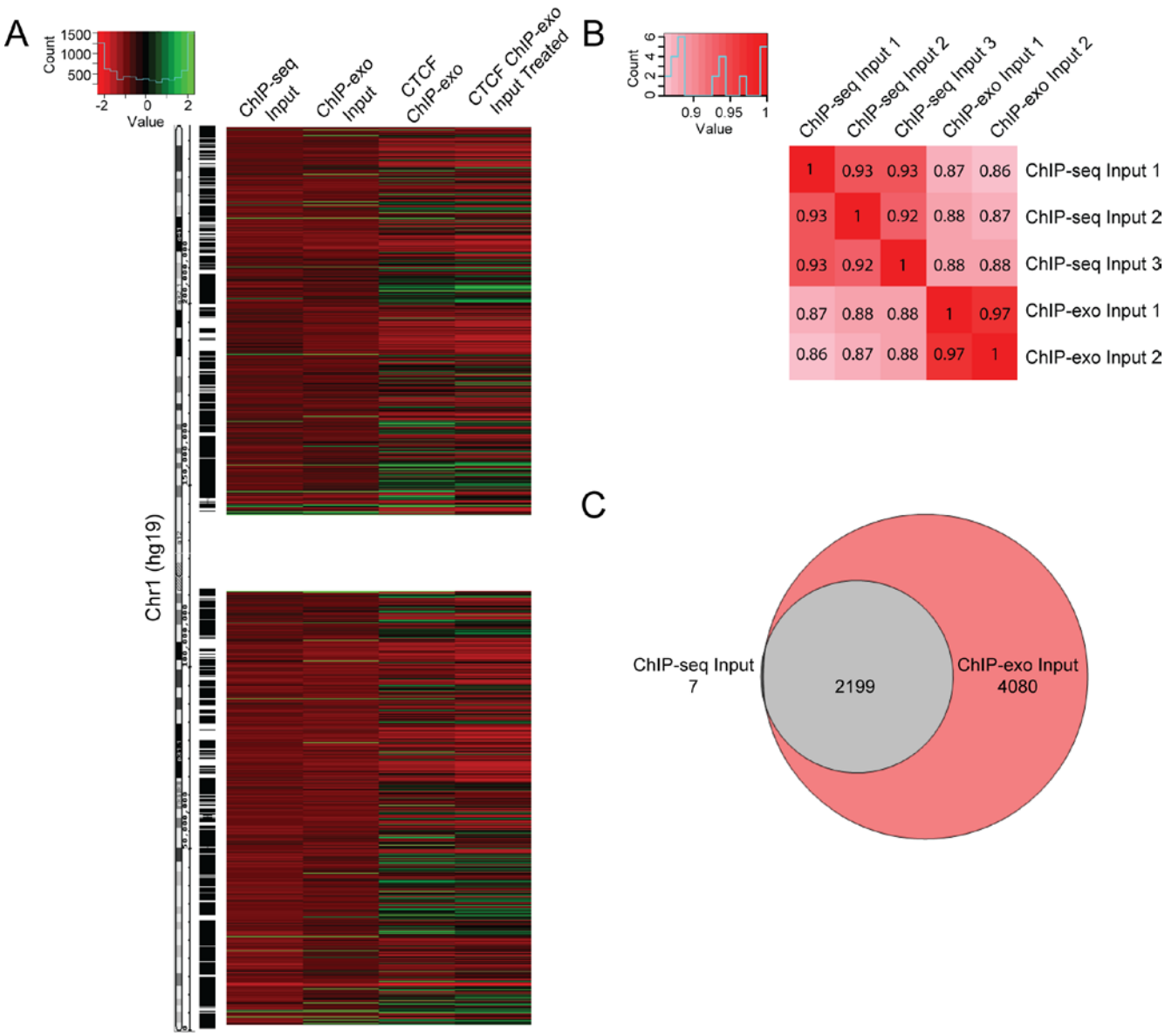
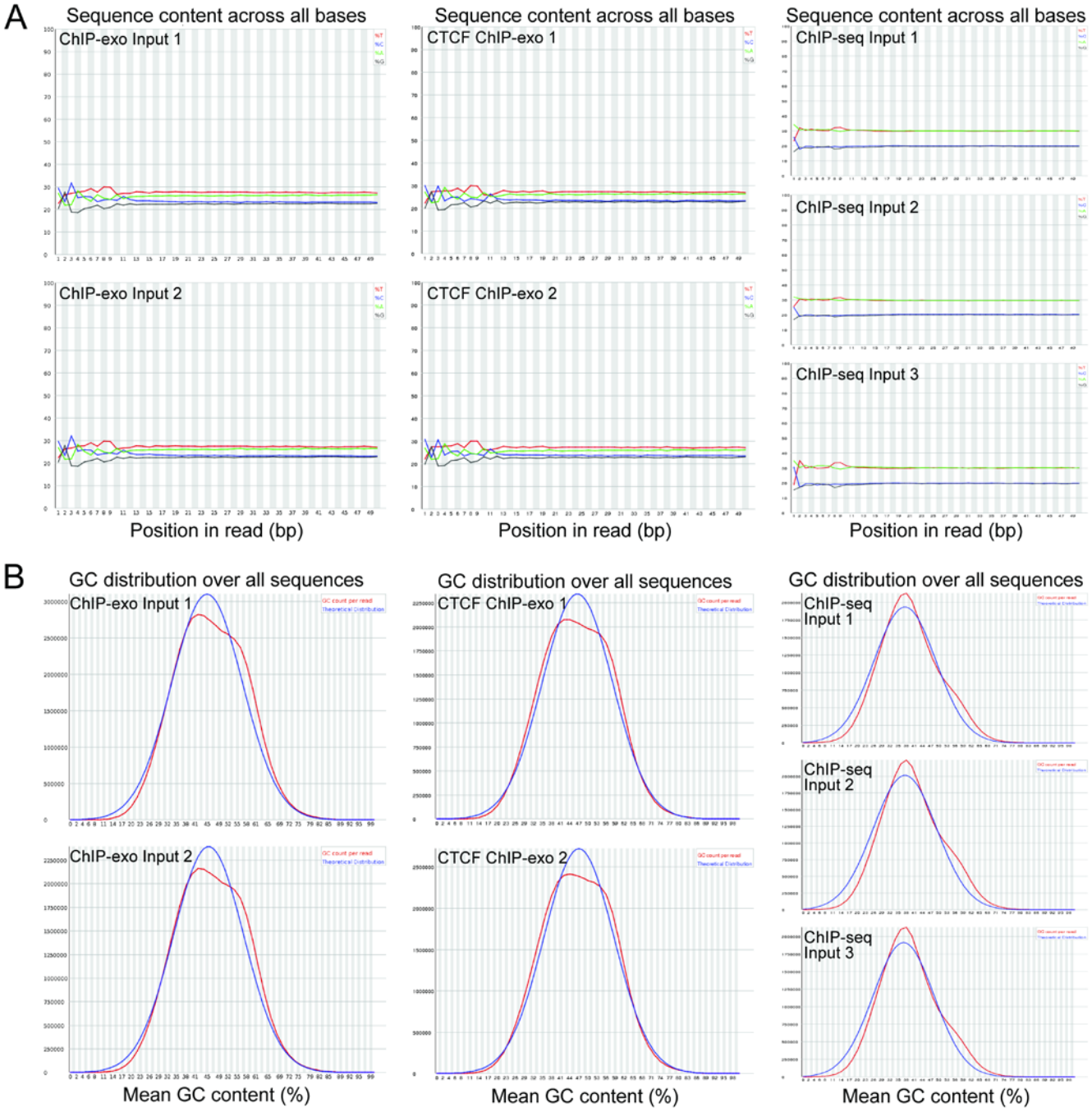
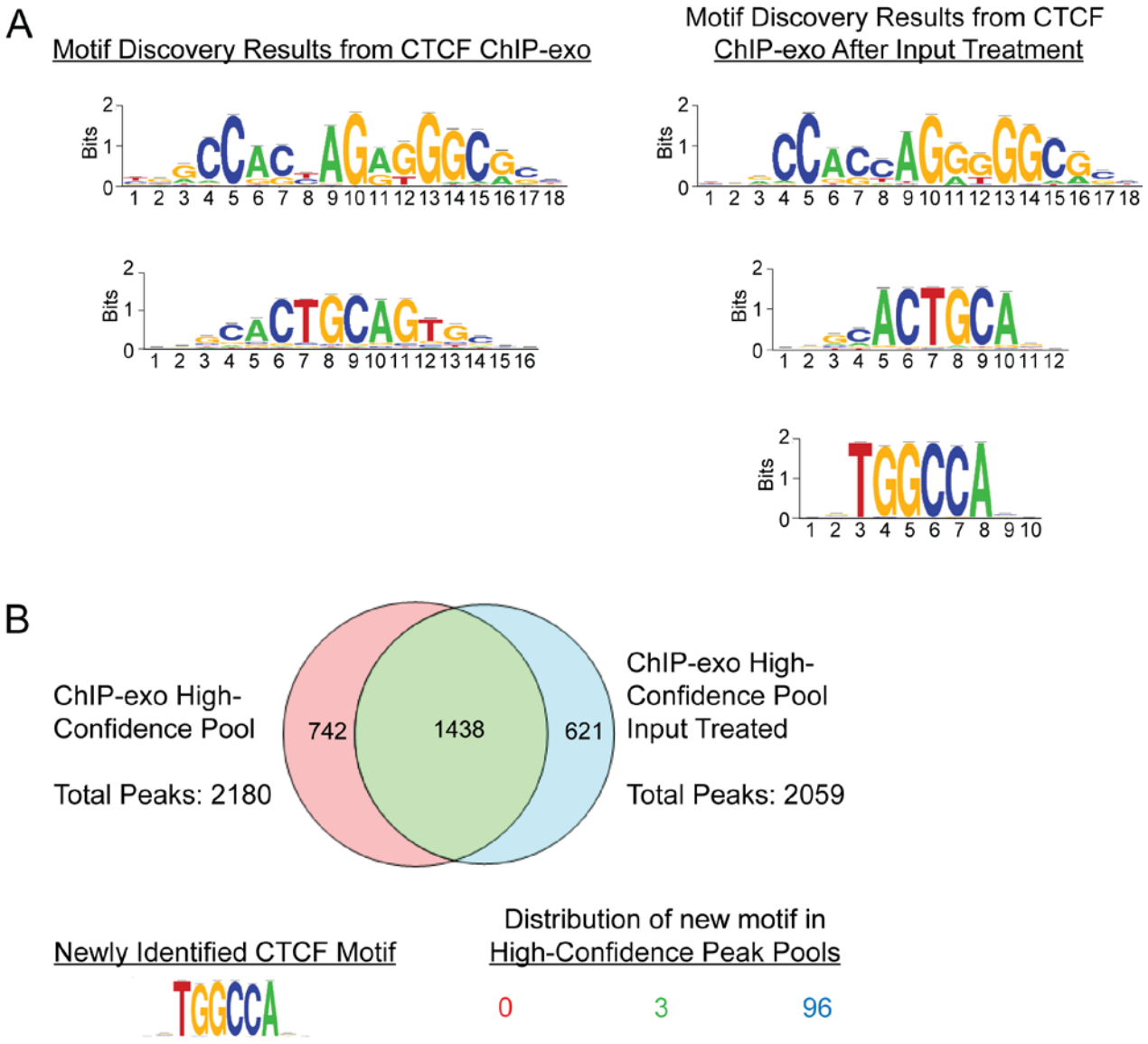


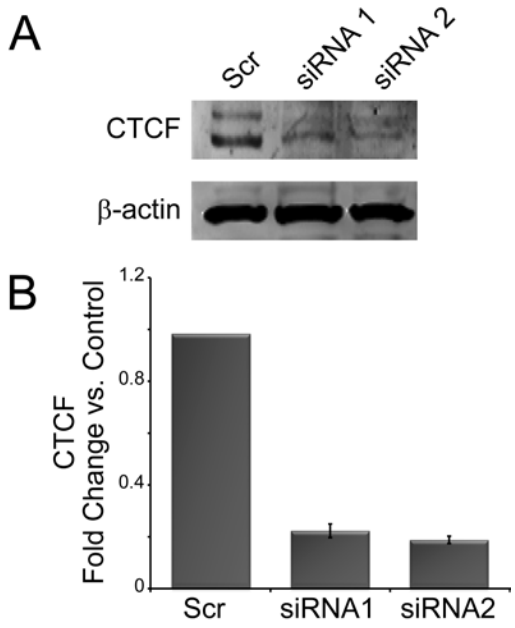
Supplementary Figure S1. Technique specific removal of false positive peaks in ChIP-exo data. (**A**) Color-coded heat map enrichment profiles visualized within the Integrated Genome Browser (IGB) for the standard ChIP-seq input, the bead-bound ChIP-exo input and the ChIP-exo profiles without and with input treatment mapped onto chromosome 1. (**B**) Genome-wide correlations (1,000 bp window) for standard ChIP-seq inputs compared to the bead-bound ChIP-exo inputs in HeLa cells. (**C**) Venn diagram showing the intersect of peaks (gray) within the standard ChIP-seq input (not observable) and the ChIP-exo input (red) generated by the PAtCh-Cap method. Together, these data indicate that while a significant number of false positive peaks are shared between the standard ChIP-seq input and the bead-bound ChIP-exo input, there a significant number of additional false positive peaks that are unique to the ChIP-exo process. Use of the ChIP-exo input generated by PAtCh-Cap significantly removes these artifacts in the CTCF ChIP-exo data as demonstrated by the substantial reduction in green loci, particularly for pericentromeric regions in (A).



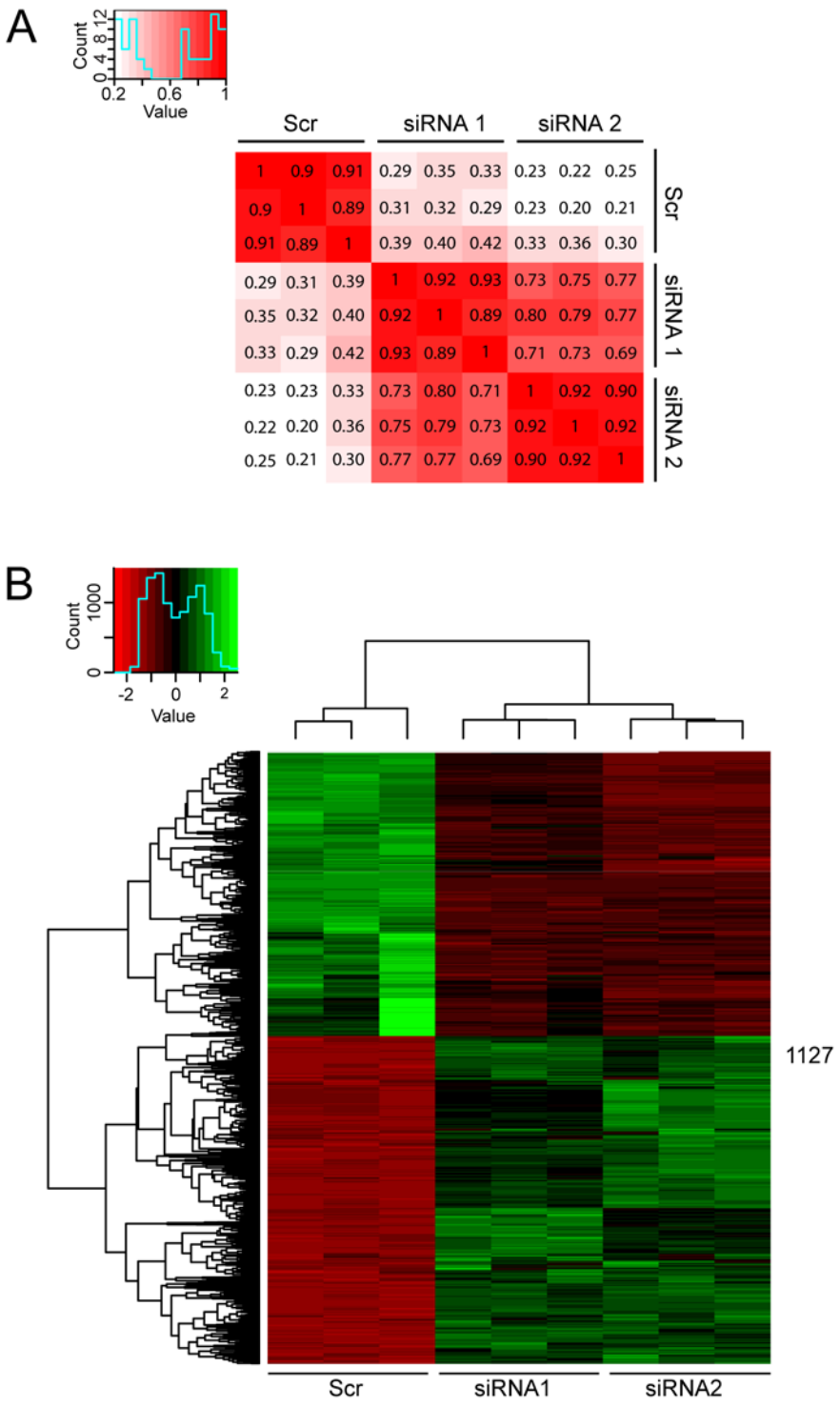
Supplementary Figure S2. (**A**) FastQC plots of nucleotide frequency as a function of base position for the ChIP-exo input, CTCF ChIP-exo, and ChIP-seq input replicates. The four colors represent the bases A (green), T (red), G (black) or C (blue). (**B**) FastQC plots showing that the GC count per read (red) for each ChIP-exo input, CTCF ChIP-exo, and ChIP-seq input data set relative to the theoretical distribution (blue).



Supplementary Figure S3. (**A**) Results for the *de novo* motif search analysis performed by RSAT on the high-confidence CTCF peaks identified from the ChIP-exo data without and with input treatment (refer to Figure 2A). The newly identified motif is only observed from the *de novo* search conducted on the input treated high-confidence pool. (**B**) Venn diagram from Figure 2B showing the overlap of high-confidence peaks identified from data sets with and without input treatment (top). Distribution analysis of the newly identified CTCF motif within each of these pools (color coded numbers; bottom) clearly shows that this sequence is found nearly exclusively within the population of high-confidence peaks that were only identified after input treatment (blue).



Supplementary Figure S4. (**A**) Representative immunoblot indicating CTCF protein depletion after HeLa cells were treated with two different siRNAs relative to a scrambled siRNA (Scr) control. (**B**) Sufficient CTCF depletion at the mRNA level in HeLa cells was further confirmed by qRT-PCR. mRNA levels are represented as the fold C_T change relative to *HPRT1* and a scrambled siRNA (normalized to 1).



Supplementary Figure S5. (**A**) Genome-wide correlations (1,000 bp window) of the RNA-seq data acquired for each of the scrambled and CTCF siRNA replicates. (**B**) Comprehensive heat map from the RNA-seq data depicting gene transcripts exhibiting a two-fold up- (green) or down-regulation (red) after CTCF depletion relative to the scrambled siRNA control (p-value cut-off of 0.05).

Supplementary Table S1. Incidence of CTCF motif within ±1000 bps of regulated gene promoter TSS

	CTCF ChIP-exo	CTCF ChIP-exo Input Treated
Total genes altered after CTCF depletion (from RNA-seq)	1127	
Number of ChIP-exo peaks intersected with above altered genes	771	448
Number of intersected peaks containing core CTCF motif within promoter	398	396
% CTCF motif	52%	88%