



Supplemental Fig. S9. Independent Associations of CCR Signal Strength with lincRNA Length and Transcript Abundance.

(A) Distribution of transcript abundance density for all lincRNA loci compared to distribution of lincRNAs with CCRs. Two RNA-seq datasets for untreated K562 cells (GEO Accession number GSE46718) (Bansal et al. 2014) were used to plot transcript abundance. Note bimodal distribution of transcript abundance for all lincRNA loci (black line) and increased probability of having moderately abundant transcripts for lincRNA loci with CCRs (colored lines). p -value $< 2.2 \times 10^{-16}$; Kruskal-Wallis test.

(B, C, D, E, F) Correlation between length and lincRNA locus CCR signal strength (HP10M) in (B) DMSO-, (C) VP16-, (D) mitoxantrone-, (E) p BQ-, and (F) genistein-treated samples.

LincRNA categories by length are shown on X-axis. Means of all CCR signal strengths within each lincRNA locus are plotted to control for tendency of longer lincRNAs to have more CCRs by virtue of increased sequence space. Boxes represent 25th to 75th percentiles; whiskers, 5th to 95th percentiles; horizontal lines, median for each length interval; chi-squared test p -values are at top right in panels. Amplified samples; same treatments merged where applicable (Supplemental Table S1). The data show that CCR signal strength correlates significantly with lincRNA length for each condition tested

(G) Scatter plot of lincRNA transcript abundance vs. length based on two RNA-seq datasets for untreated K562 cells (GEO Accession number GSE46718) (Bansal et al. 2014) and based on lincRNA length from GRCh38/hg38. Smooth line was predicted by gam method. Shading indicates confidence interval around smoothed trend line. lincRNAs with RPKM > 0.1 are plotted. r value (top right in panel) indicates that transcript abundance shows a slight overall negative correlation with length.

(H) Box and whisker plots of CCR signal strength vs. transcript abundance within indicated lincRNA length categories sub-divided based on lower or higher than average CCR signal strength. * p -value < 0.05 ; ** p -value < 0.01 ; < 10 kb ($p = 0.00595$), 10-49 kb ($p = 0.02924$), 50-99 kb ($p = 0.03446$), ≥ 100 kb ($p = 0.03446$); chi-squared test. Union set of all amplified samples merged (Supplemental Table S1). Note that genic CCR signal strength correlates significantly with transcript abundance across all lincRNA length categories.