Cell Reports, Volume *41*

Supplemental information

Systematic discovery and functional dissection of enhancers needed for cancer cell fitness and proliferation

Poshen B. Chen, Patrick C. Fiaux, Kai Zhang, Bin Li, Naoki Kubo, Shan Jiang, Rong Hu, Emma Rooholfada, Sihan Wu, Mengchi Wang, Wei Wang, Graham McVicker, Paul S. Mischel, and Bing Ren

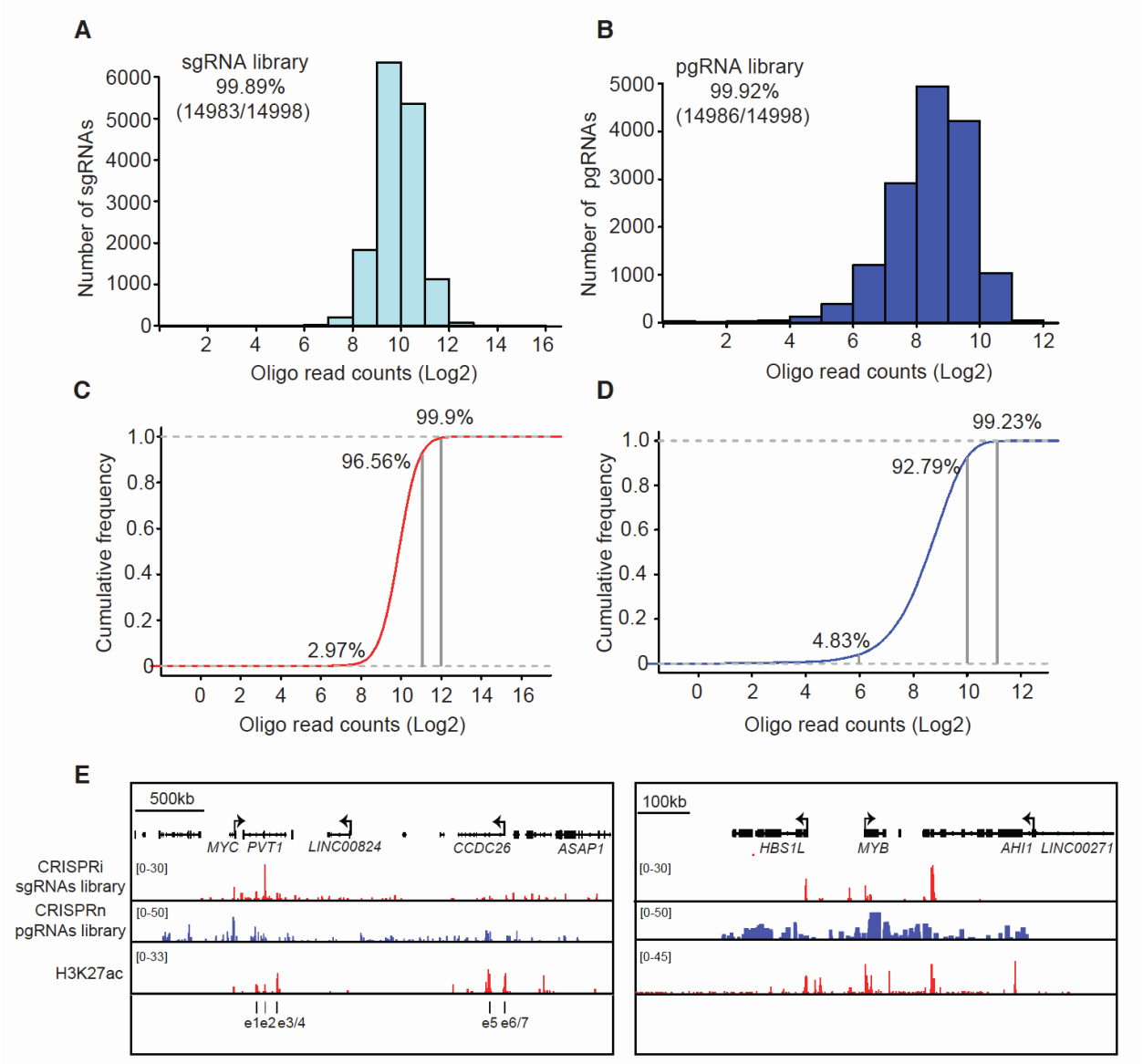


Figure S1. Comparison of different CRISPR-based enhancer screening strategies, Related to Figure 1
(A-D) Library construction and quality control. The plasmid library recovered > 99.5% of oligos in both single guide RNA, sgRNA (A), and paired guide RNA, pgRNA (B), library designs. Distribution of oligo read counts (A-B) and cumulative frequency in the plasmid DNA library (C-D). (E) The genome browser snapshot shows the CRISPRi score from sgRNA library (top) coupling with CRISPR interference (CRISPRi) or pgRNA library (middle) coupling with CRISPR nuclease (CRISPRn) screens at the *MYC* (left) and *MYB* (right) loci. H3K27ac ChIP-seq signal, H3K27ac (bottom), and e1-e7 are the functional enhancers identified from the previous study (Fulco *et al.*, 2016).

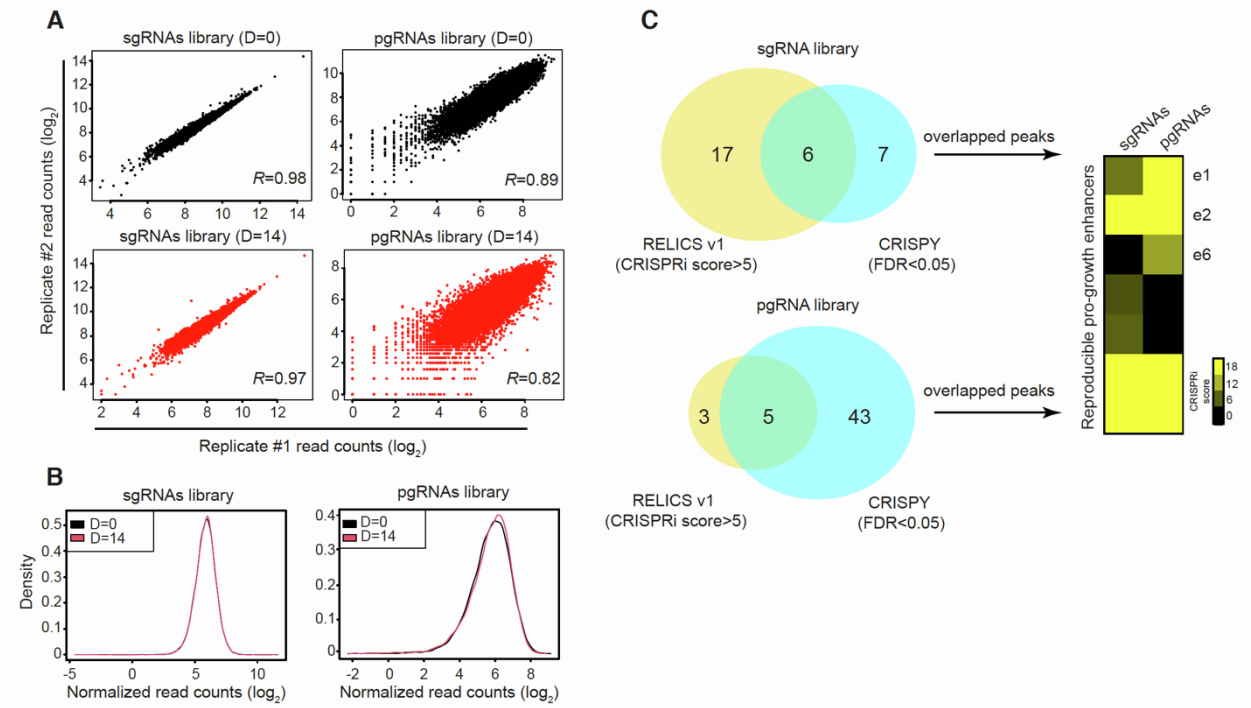


Figure S2. Comparison of essential enhancers identified from single and paired gRNA CRISPRi screens, Related to Figure 1

(A-B) Quality control of CRISPRi based proliferation screens in K562. Scatter plots (A) show the correlations between two biological replicates from sgRNAs or pgRNAs library at the indicated time-point. Doubling time 0, D=0; 14 doubling time, D=14. (B) Density plots show guide RNA abundance between different time-points from the screens using sgRNA (left) or pgRNA (right) library. (C) The criteria to identify reproducible essential enhancers identified from sgRNA or pgRNA library. Left: Venn diagrams show the overlapping and nonoverlapping of essential enhancers identified using RELICS or CRISPY analysis pipeline in two screen libraries. Candidate essential enhancers are identified with CRISPRi score > 5 in RELICS v1 and with FDR < 0.05 from negative selection in CRISPY, respectively. Reproducible essential enhancers are selected as the regions identified from two independent pipelines. Right: heatmap representation of CRISPRi signal at reproducible essential enhancers identified from sgRNA or pgRNA screen. Each row represents the CRISPRi signal of an individual essential enhancer and each column represents a library type used in the CRISPRi screen.

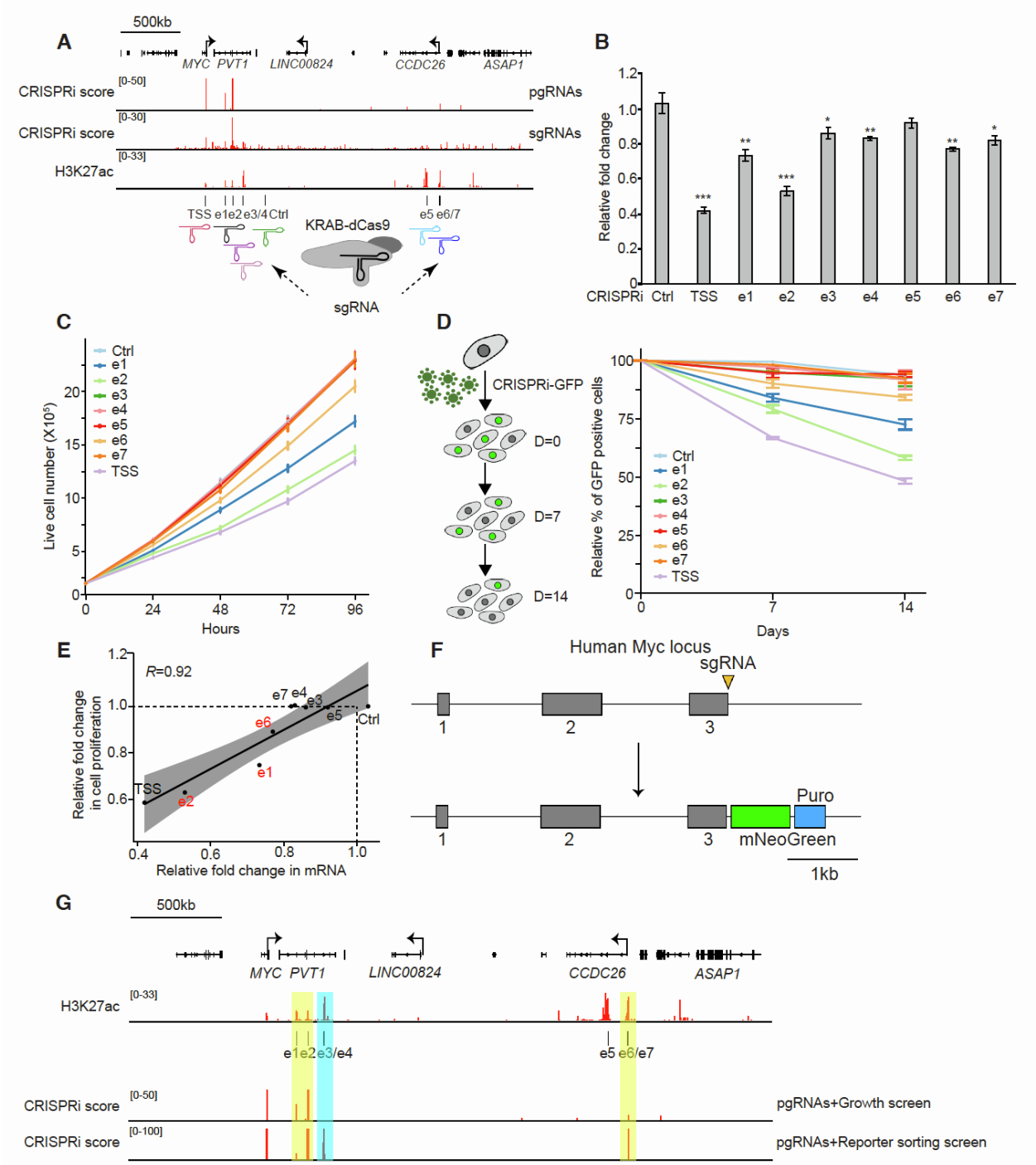
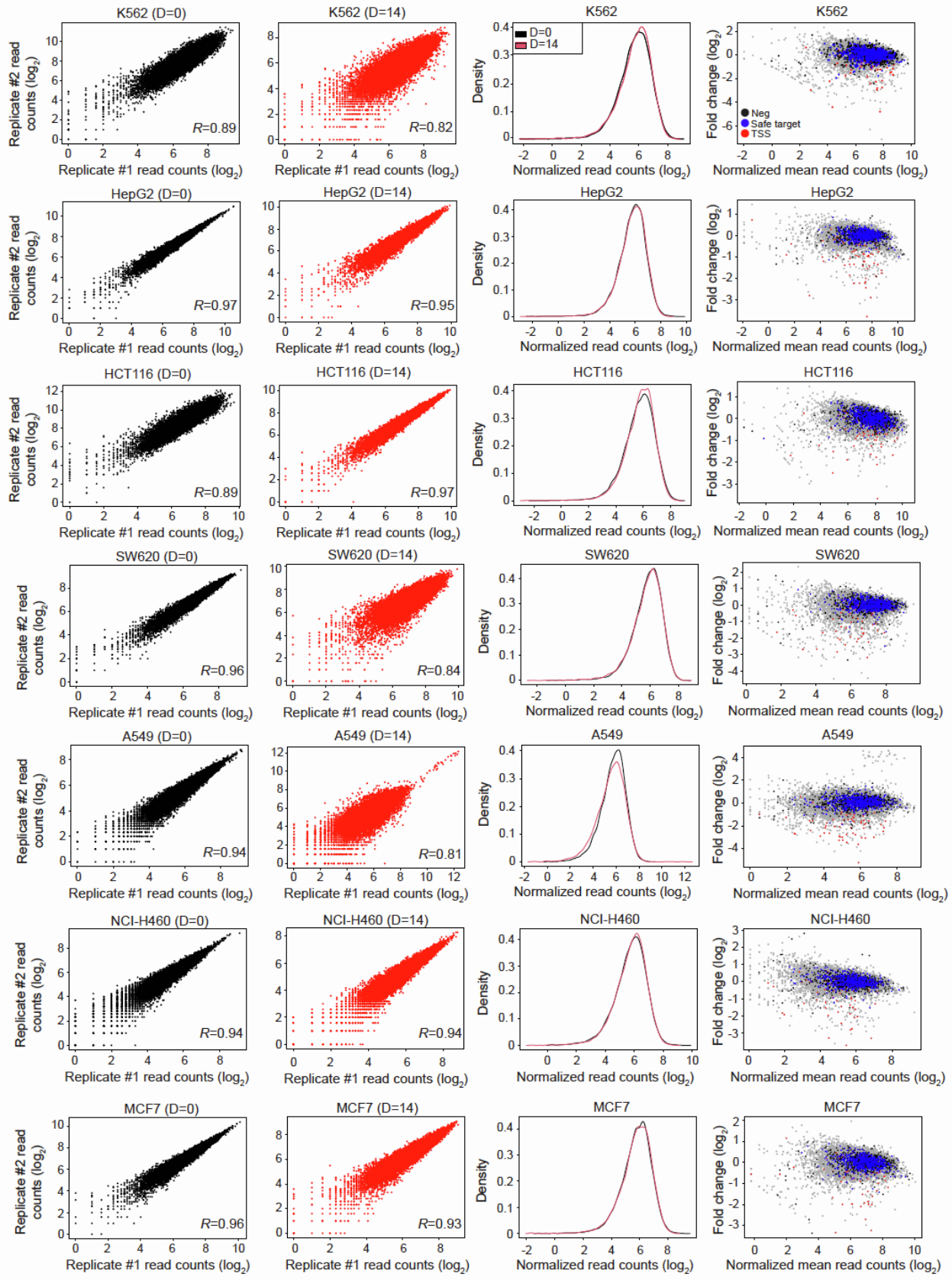


Figure S3. Validation of pgRNA library design using different functional assays, Related to Figure 1
(A) Top: Genome browser snapshot shows CRISPRi signal from sgRNA or pgRNA CRISPRi screen at *MYC* locus in K562. Bottom: Scheme of the experimental strategy to individual validate e1-e7 functional enhancers, identified from the previous study (Fulco *et al.*, 2016), for gene regulation and cell proliferation. sgRNA targeting the TSS of *MYC* gene is served as a positive control. (B) Gene expression of *MYC* measured by RT-qPCR after silencing selected enhancers with KRAB-dCas9. Relative fold change represents gene expression ratio from the CRISPRi silenced indicated enhancer to the control samples. Data shown are mean ± SD from three biological replicates performed. *P*-values were determined by a two-tailed Student's *t*-test after silencing individual enhancer compared to control cells (*: $p < 0.1$; **: $p < 0.01$; ***: $p < 0.001$). (C) The measurement of cell proliferation after silencing selected enhancers with KRAB-dCas9 in K562. (D) The relative change in cell abundance of GFP-positive cells.

Left: Scheme of the experimental strategy to illustrate a competitive cell growth assay. Cell population was harvested at different time points and the percentage of GFP positive cells was determined by FACS. Right: The measurement of relative abundance change in GFP positive cells after silencing indicated enhancers with KRAB-dCas9. (E) Correlation between relative *MYC* gene expression and cell proliferation (Pearson's $R = 0.92$). The essential enhancers identified from pgRNA library are labeled in red. The grey shaded area represents the 95% confidence interval. (F) Scheme of the experimental strategy to illustrate CRISPR/Cas-9 mediated gene targeting for knocking in a fluorescent reporter, mNeoGreen, into the C-terminal of the *MYC* gene in K562 cells. The yellow triangle represents the sgRNA target site at the last exon of *MYC*. (G) CRISPRi signal from pgRNA library using proliferation-based screen (top) or reporter sorting-based screen (bottom). The common or assay-specific essential enhancers are highlighted in yellow or blue, respectively.


K562 (D=0)
K562 (D=14)
K562
K562
HepG2 (D=0)
HepG2 (D=14)
HepG2
HepG2
HCT116 (D=0)
HCT116 (D=14)
HCT116
HCT116
SW620 (D=0)
SW620 (D=14)
SW620
SW620
A549 (D=0)
A549 (D=14)
A549
A549
NCI-H460 (D=0)
NCI-H460 (D=14)
NCI-H460
NCI-H460
MCF7 (D=0)
MCF7 (D=14)
MCF7
MCF7
Replicate #2 read counts (log2)
Replicate #1 read counts (log2)
R=0.89
R=0.82
R=0.97
R=0.95
R=0.89
R=0.97
R=0.96
R=0.84
R=0.94
R=0.81
R=0.94
R=0.94
R=0.96
R=0.93
Density
Normalized read counts (log2)
D=0
D=14
Fold change (log2)
Normalized mean read counts (log2)
Neg
Safe target
TSS

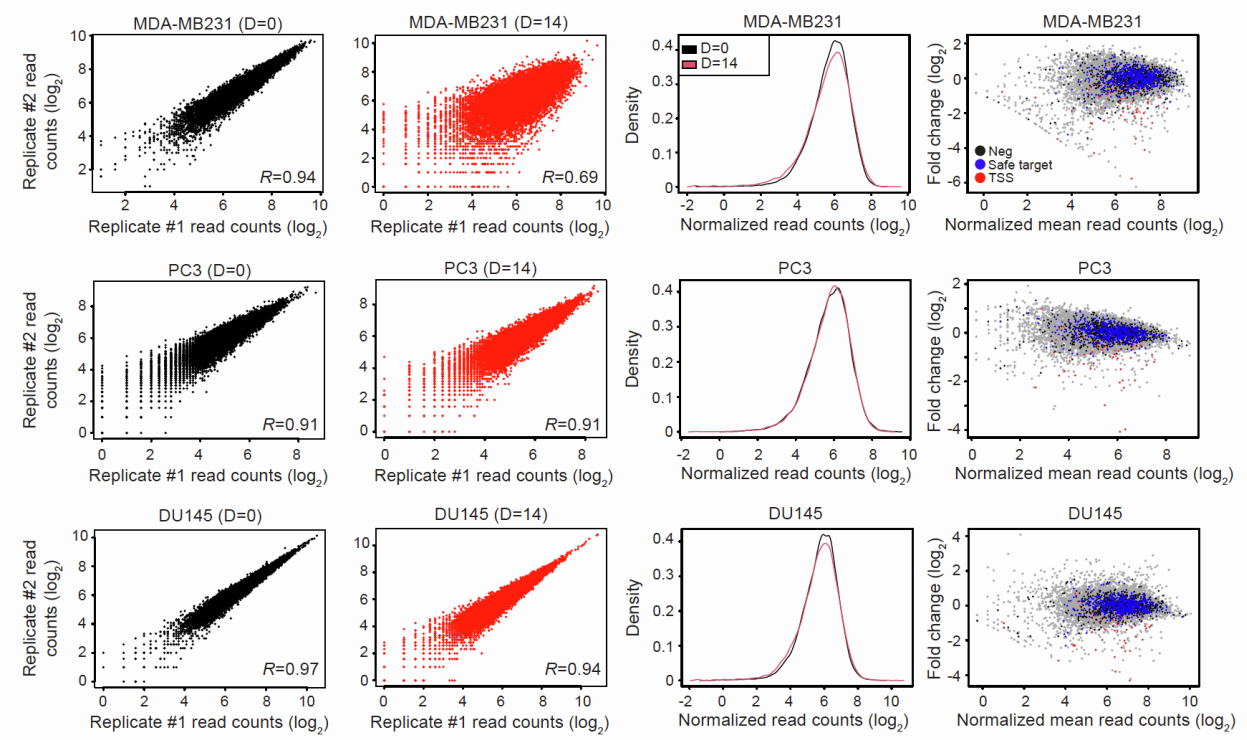


Figure S4. Quality assessment of proliferation-based CRISPRi screen with a tiling pgRNA library in 10 human cancer cell lines, Related to Figure 2

Left: Scatter plot represents normalized read count of each pgRNA at the indicated time point. Pearson's *R* was calculated between two biological replicates in each cell line. Doubling time 0, D=0; 14 doubling time, D=14. Middle: Density plots of pgRNA abundance between different time-points. Right: The read counts of pgRNAs from a cell population after 14 doubling times (D=14) compared with those from the initial control population (D=0). Fold change represents the ratios between read counts in D=14 population and the control population. pgRNAs targeting TSS of essential genes, negative controls, and safe-targeting genomic regions are labeled in red, black, and blue, respectively.

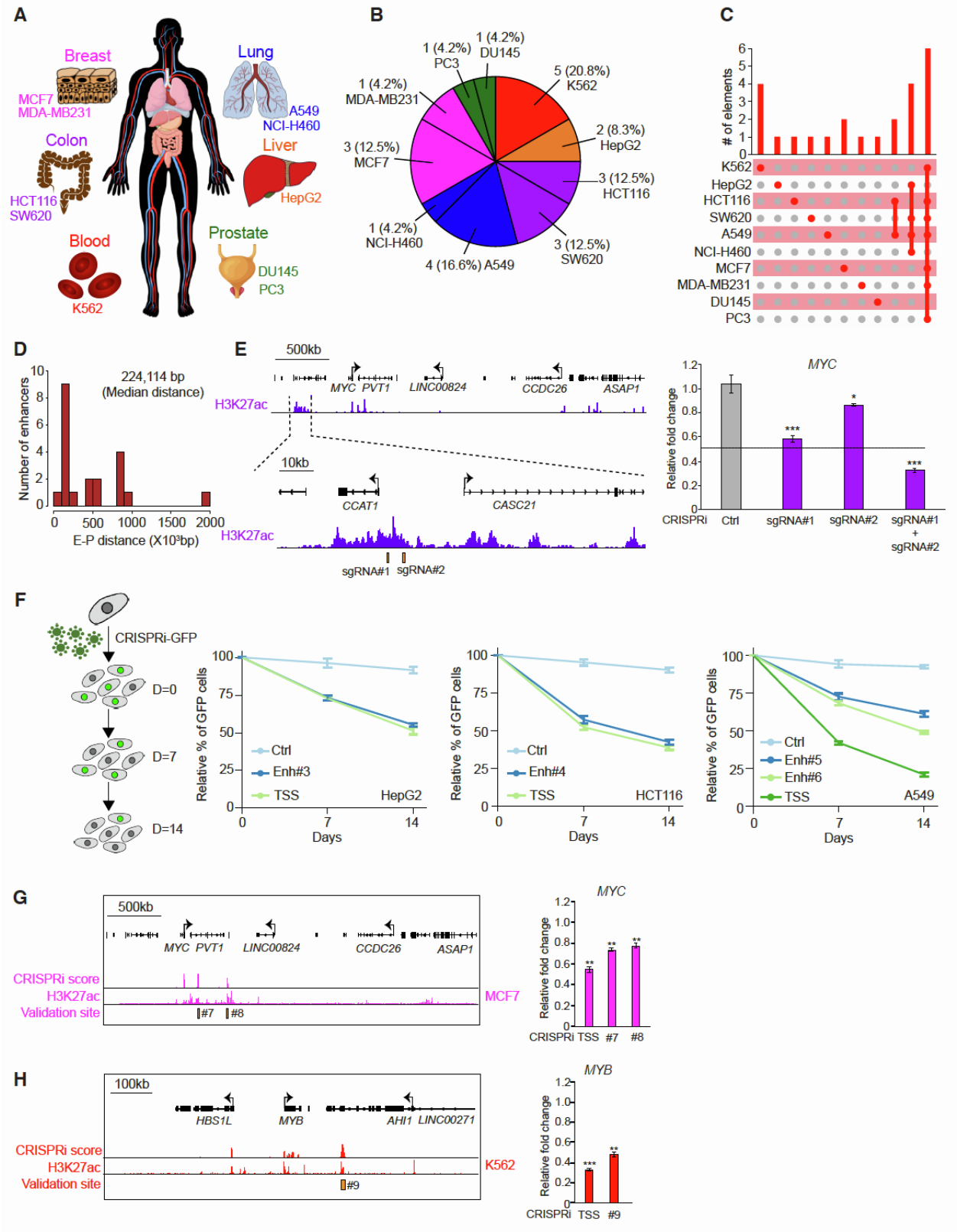


Figure S5. Summary of essential enhancers identified in 10 human cancer cell lines, Related to Figure 2

(A) Cancer types are included in this study. CRISPRi screen was carried out in a total of ten different cancer cell lines representing six cancer types. (B) The number of essential enhancers identified from each cell line. (C) The number of cell-line-specific and pan-cancer essential enhancers. Red dots and gray dots indicate the presence and absence of essential enhancers, respectively. (D) The distance between essential enhancers and the TSS of target genes. (E) Validation of pgRNA design achieved greater silencing effect in one previously identified enhancer (Hnisz et al., 2013) in HCT116. Messenger RNA expression level was measured using RT-qPCR after targeting different genomic loci with KRAB-dCas9. Relative fold changes represent the ratios of gene expression from the CRISPRi silenced essential enhancer to the control samples. Data shown are mean ± SD from three biological replicates performed. *P*-values were determined by two-tailed Student's *t*-test (*: $p < 0.1$; ***: $p < 0.001$). The line indicates the expected fold-change by targeting two genomic loci. The orange bar presents the target site of each sgRNA with tracks showing the H3K27ac signal at the selected genomic locus. (F) Validation of the effect of essential enhancers in cell proliferation. Left: Scheme of the experimental strategy to illustrate a competitive cell growth assay. Cell population was harvested at different time points and the percentage of GFP positive cells was determined by FACS. Right: The measurement of relative abundance changes in GFP positive cells after silencing indicated enhancers (genomic loci of these selected enhancers is illustrated in Fig. 2B) with KRAB-dCas9 in three cancer cell lines. sgRNA targeting the TSS of *MYC* gene is served as a positive control. (G-H) Left: Genome browser snapshot shows essential enhancers identified in MCF7 (G) and K562 (H) with tracks showing H3K27ac signal at selected genomic loci. Right: The measurement of relative fold change in *MYC* (G) and *MYB* (H) oncogenes by RT-qPCR after epigenetic silencing essential enhancers using KRAB-dCas9. Relative fold changes represent the ratios of gene expression from the CRISPRi silenced essential enhancer to the control samples. Data shown are mean ± SD from three biological replicates performed. *P*-values were determined by a two-tailed Student's *t*-test (**: $p < 0.01$; ***: $p < 0.001$).

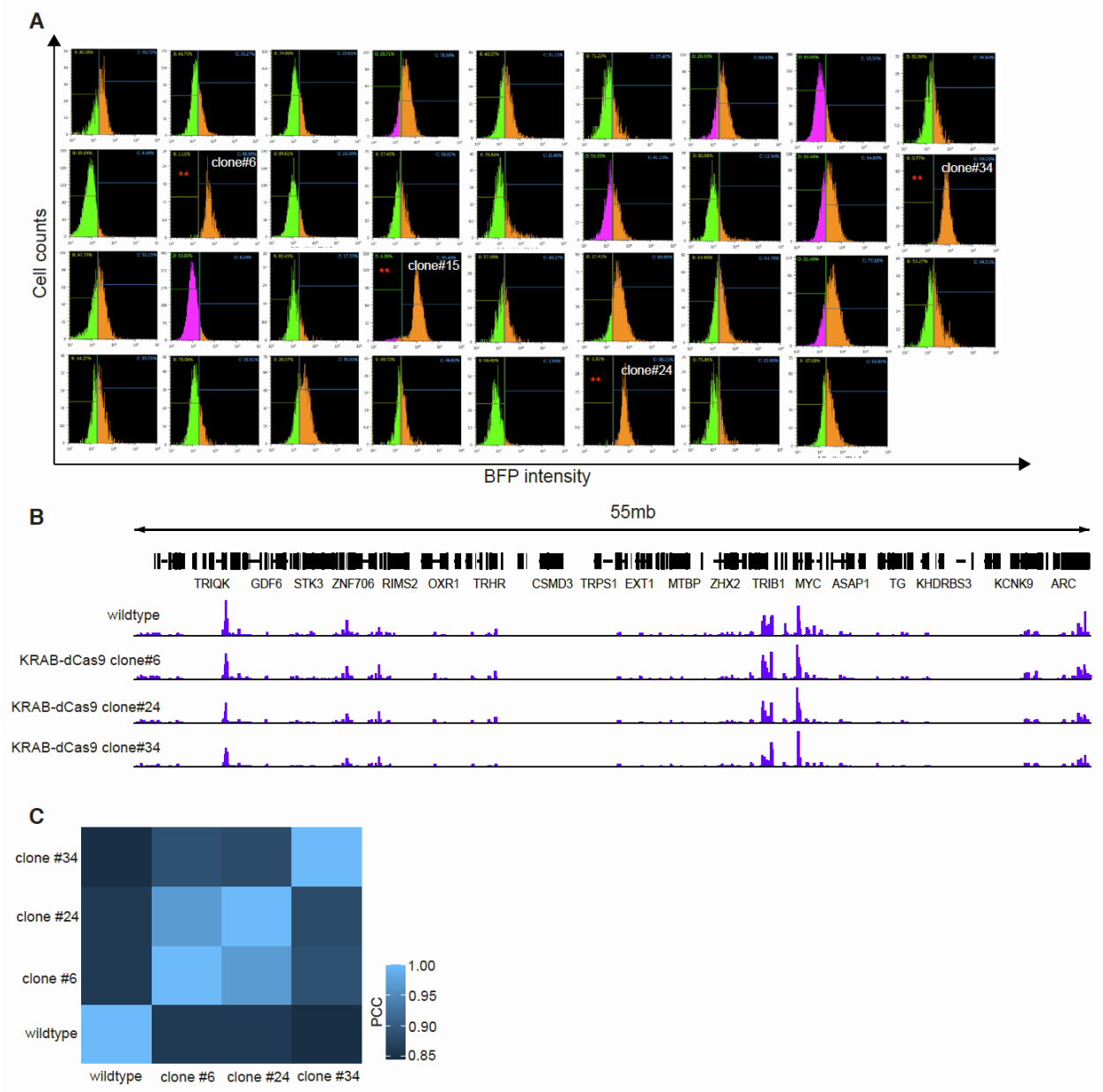


Figure S6. H3K27ac signals in KRAB-dCas9 expressing HCT116 cell clones are similar to the parental cells, Related to Figure 3

(A) The selection of KRAB-dCas9 clones based on the expression level of BFP (dCas9-TagBFP-KRAB) by FACS analysis. (B) Genome browser snapshot shows H3K27ac ChIP-seq signal between parental HCT116 (wildtype) and three KRAB-dCas9 stable clones (KRAB-dCas9 clone #6, #24, #34) within the selected 55-Mbp genomic locus. (C) Heatmap shows Pearson correlation of H3K27ac signal in distal enhancers between parental HCT116 (wildtype) and three KRAB-dCas9 stable clones (KRAB-dCas9 clone #6, #24, #34).

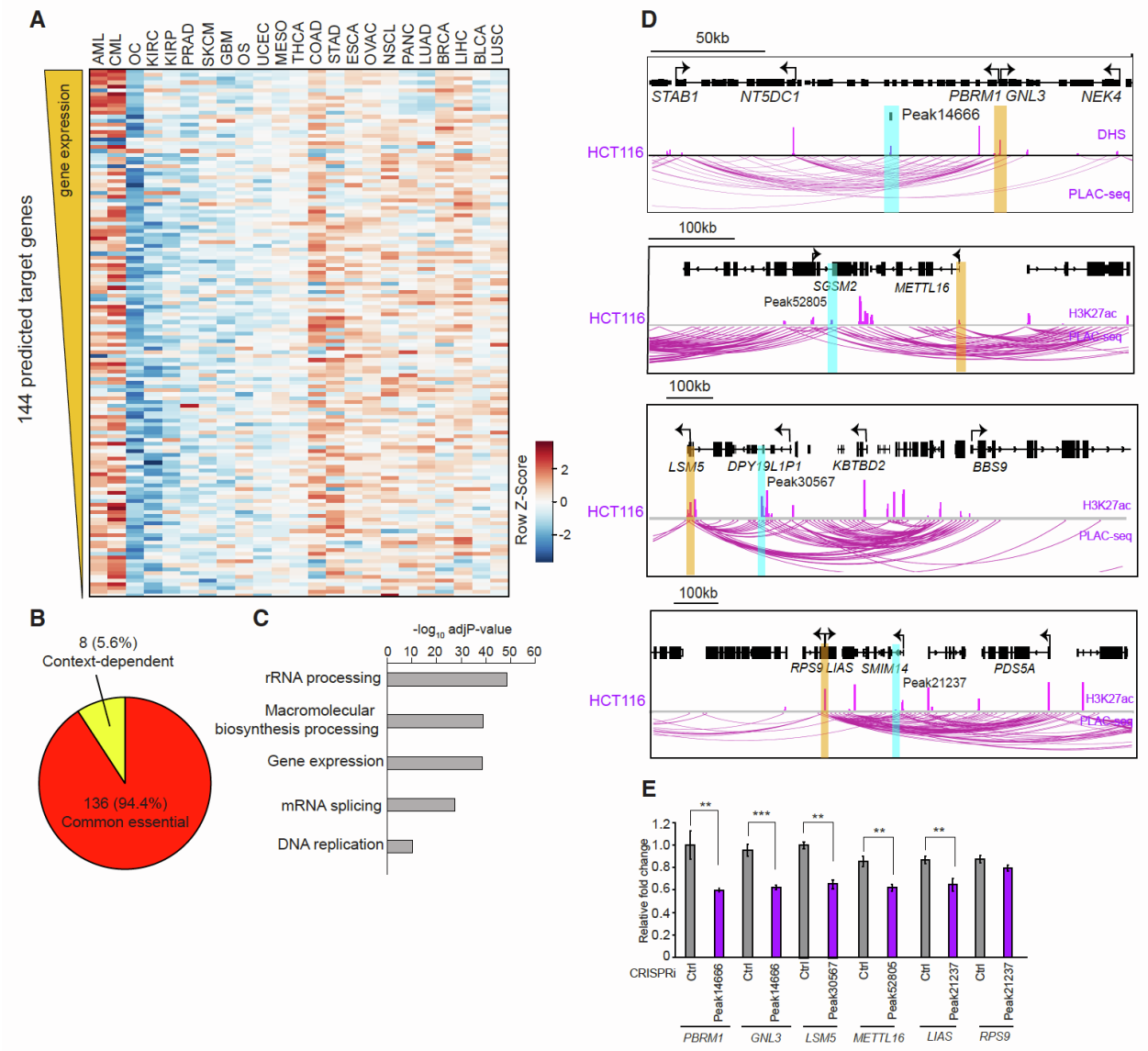


Figure S7. Expression and function of predicted target genes, Related to Figure 3

(A) Ubiquitous expression of predicted target genes. Heatmap representation of gene expression of 144 predicted target genes regulated by 150 essential enhancers across 23 cell types. Each row represents an individual gene. The median gene expression level [$\log_2$ (TPM+1)] from each cell type is used. Predicted target genes are sorted by decreasing expression levels from high to low. Color represents the relative RNA-seq gene expression from the Cancer Cell Line Encyclopedia (Ghandi *et al.*, 2019) for each predicted target as a z-score. (B) The percentage of context-dependent or common essential genes in the predicted target genes. (C) Gene ontology (GO) analysis of the predicted target genes regulated by essential enhancers using Enrichr (Kuleshov et al., 2016). GO terms that were partially redundant with those listed were eliminated for brevity, and the top five GO terms in the biological process were shown. (D) Chromatin interactions identified by H3K4me3 PLAC-seq between essential enhancers and predicted target genes in HCT116. The essential enhancers are highlighted in cyan and the predicted target genes are highlighted in orange. (E) Relative RNA levels of predicted target gene determined by RT-qPCR after silencing indicated essential enhancers with KRAB-dCas9 in HCT116. Relative fold changes represent the ratios between gene expression after CRISPRi indicated enhancer and the control silenced samples. Data shown are mean ± SD from three biological replicates performed. *P*-values were determined by a two-tailed Student's *t*-test (**: $p < 0.01$; ***: $p < 0.001$).

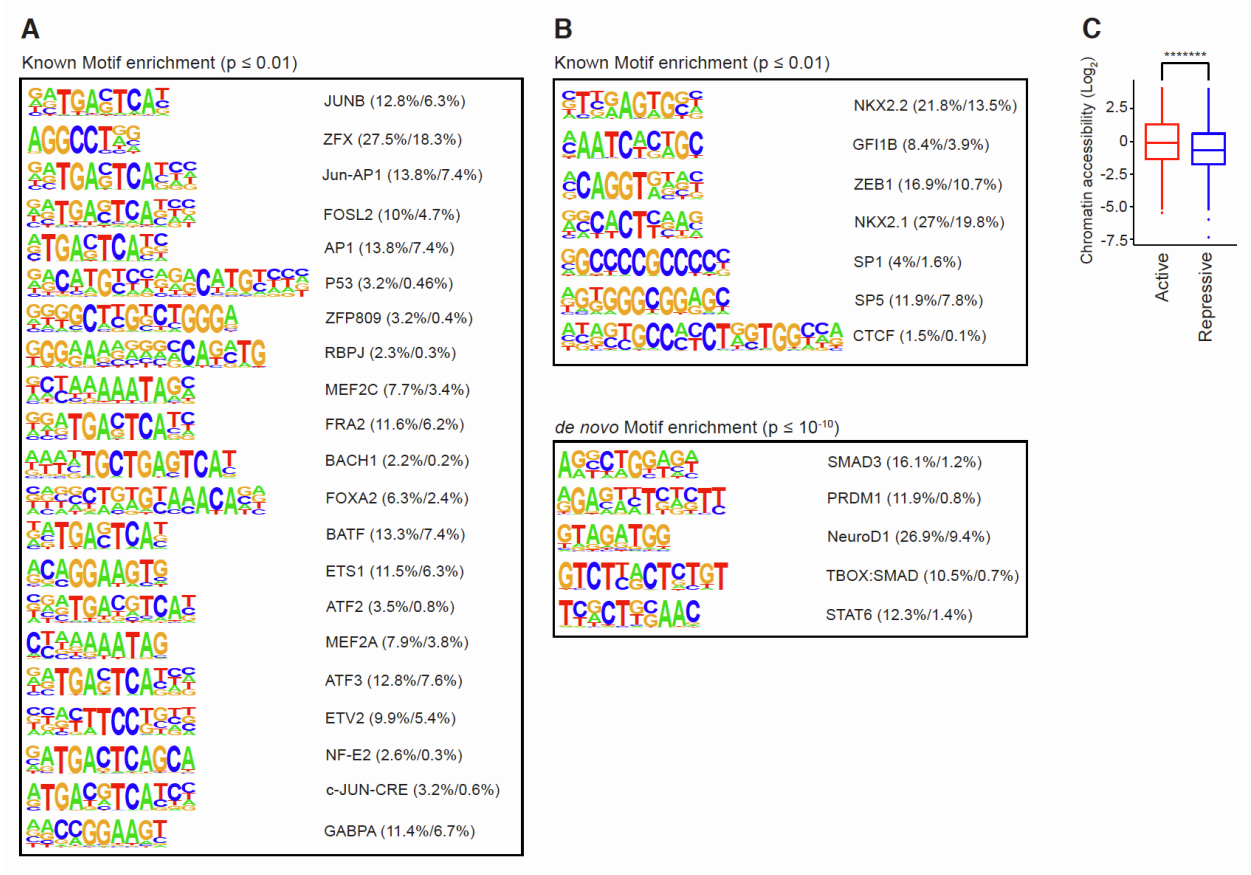


Figure S8. Motif enrichment analysis from STARR-seq screen, Related to Figure 4

(A-B) The motifs identified from (A) 687 activating or (B) 896 repressive elements. (C) Comparison of chromatin accessibility between activating and repressive elements. *P*-values were determined by the two-sided Wilcoxon test (*******: $p < 0.0000001$).

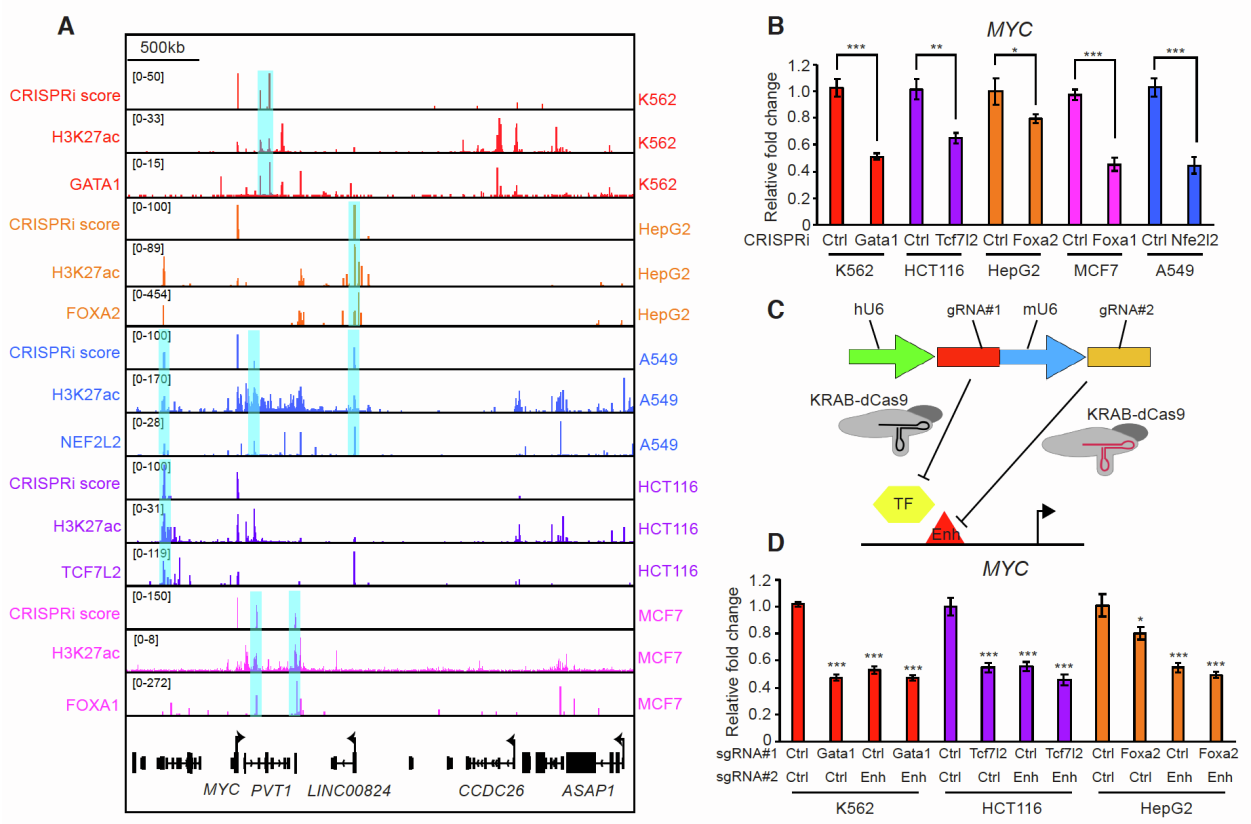


Figure S9. Lineage-specific transcription factors bind to the essential enhancers near MYC gene, Related to Figures 2 and 4

(A) ChIP-seq signal of lineage-specific TFs around *MYC* locus overlapping with essential enhancers. Genome browser snapshot shows the selective enrichment of GATA1, FOXA2, NEF2L2, TCF7L2, and FOXA1 bindings at essential enhancers identified from CRISPRi screen in K562, HepG2, A549, HCT116, and MCF7, respectively. Essential enhancers with strong enrichment of lineage-specific TF binding are highlighted in blue. (B) Gene expression of *MYC* measured by RT-qPCR after silencing the selected lineage-specific transcription factor with KRAB-dCas9 in various cancer cell lines. Data shown are mean ± SD from three biological replicates performed. *P*-value was determined by a two-tailed Student's *t*-test. (*: $p < 0.1$, **: $p < 0.01$ and ***: $p < 0.001$). (C) Scheme of the experimental strategy to silence lineage-specific TF, essential enhancer or both using KRAB-dCas9. (D) RT-qPCR for relative RNA levels of *MYC* expression in cells expressing sgRNAs targeting individual essential enhancer and lineage-specific TF or both. Relative fold changes represent the ratios of gene expression from the CRISPRi silenced essential enhancer to the control samples. Data shown are mean ± SD from three biological replicates performed. *P*-value was determined by a two-tailed Student's *t*-test. (*: $p < 0.1$ and ***: $p < 0.001$).