A cis-element within the *ARF* locus mediates repression of *p16*^{INK4A} expression via long-range chromatin interactions

Yang Zhang[a,1], Judith Hyle[a,1], Shaela Wright[a,1], Ying Shao[b], Xujie Zhao[c], Hui Zhang[d] and Chunliang Li[a,*]

[a]Department of Tumor Cell Biology, St. Jude Children's Research Hospital, 262 Danny Thomas Place, Memphis, TN 38105, USA

[b]Department of Computational Biology, St. Jude Children's Research Hospital, 262 Danny Thomas Place, Memphis, TN 38105, USA

[c]Department of Pharmaceutical Sciences, St. Jude Children's Research Hospital, 262 Danny Thomas Place, Memphis, TN 38105, USA

[d]Department of Hematology and Oncology, Guangzhou Women and Children's Medical Center, Guangzhou Medical University, 9 Jinsui Road, Guangzhou, Guangdong 510623, People's Republic of China.

[1]These authors contributed equally to this study.

[*]To whom correspondence should be addressed. Chunliang Li, PhD. Department of Tumor Cell Biology, St. Jude Children's Research Hospital, 262 Danny Thomas Place, Memphis, TN 38105, USA. Email: chunliang.li@stjude.org; Telephone:1-901-595-6530.

www.pnas.org/cgi/doi/10.1073/pnas.1909720116

METHODS and MATERIALS

Cell culture

SEM cells (ACC-546, DSMZ) and SK-N-SH cells (HTB11, ATCC) were maintained in RPMI-1640 medium (Hyclone) containing 10% fetal bovine serum (FBS) (HyClone), 2 mM glutamine (Sigma), and 1% penicillin/streptomycin (Thermo Fisher Scientific). All cells were maintained at 37°C in a 5% CO_2 atmosphere and 95% humidity. All cells used in this study were mycoplasma free. Cell identity was confirmed by STR analysis.

Vector construction

All target-specific sgRNAs were predicted using online software (http://crispr.mit.edu/). Oligomers containing 20-bp of the selected target sequence were synthesized by Thermo Fisher Scientific and cloned into the all-in-one vector (1) between BsmBI sites. All oligomer information is listed in *SI Appendix*, Dataset S1. Correct clones were screened and confirmed by Sanger sequencing with the U6-Forward sequencing primer. A two-step cloning protocol was used to generate the long-HA donor vector. Primers were designed to amplify the 800-bp 5′ HA flanking the endogenous sgRNA target. Overhangs of 23-bp, including the target sgRNA and PAM sequences, were included on the 5′ end of the HA sequence. The 3′ HA was generated according to a similar design but included an additional overhang for in-fusion cloning. The HAs were amplified from SEM cell genomic DNA. The *P2A-mCherry* DNA fragment was amplified from pEGFP-C1 (Clontech). SnapGene software (SnapGene) was used to design all primers used for in-fusion cloning (Clontech). The PCR reactions were performed using CloneAmp

polymerase (Clontech), and the cycling parameters were as follows for all cloning: 98°C for 2 min, followed by 98°C for 15 s, 55°C for 20 s, and 72°C for 20 s for 40 cycles. First, the amplified 5′ HA for knock-in was cloned into pBluescript-SK by using the TA Cloning Kit (Thermo Fisher Scientific). The DNA was purified from colonies and screened by Sanger sequencing with the primers M13F and M13R. The pBluescript–SK–sgRNA–PAM–5′ HA was then linearized and ligated with *P2A-mCherry* and the 3′ HA–sgRNA–PAM through in-fusion cloning (Clontech). Sanger sequencing was performed to ensure that the knock-in DNA was cloned in-frame with the HAs. The Lenti-dCas9-KRAB-Blast plasmid (#89567) and the LentiGuide-Puro plasmid (#52963) were purchased from Addgene. The primer sequences used for cloning and sequencing are listed in *SI Appendix*, Dataset S1.

Establishment of a *p16$^{INK4A\text{-}P2A\text{-}Cherry}$* reporter cell line

SEM were electroporated by using the Nucleofector-2b device (Lonza) with the V-kit and program X-001. For *p16$^{INK4A\text{-}P2A\text{-}Cherry}$* knock-in delivery, 2.5 μg of the donor plasmid and 2.5 μg of the CRISPR/Cas9-p16^{INK4A}-C-terminus-gRNA all-in-one plasmid were used for 5 million SEM cells. Twenty-four hours after transfection, cells were sorted for the CRISPR/Cas9-p16^{INK4A}-C-terminus-gRNA vector GFP fluorescent marker to enrich the transfected cell population. After sorted cells recovered in culture for up to 3 weeks, a second sort was performed to select cells for successful knock-in by sorting for cells expressing the knockin mCherry fluorescent marker. Two weeks later, a third sort was repeated for the mCherry fluorescent marker. Single-cell derived colonies were picked up and expanded for further characterization.

Characterization of successful knock-in events by PCR and Sanger Sequencing

DNA from single-cell derived colonies was extracted with a PureLink Genomic DNA Mini Kit (Thermo Fisher Scientific). Combinatorial primer sets designed to recognize the 5′ and 3′ knock-in boundaries were used with the following PCR cycling conditions: 98°C for 2 mins, followed by 40 cycles of 98°C for 15 s, 55°C for 20 s, and 72°C for 20 s. The bands shown at the expected size in electrophoresis were cut out, purified, and ligated to the pCR-Blunt II-TOPO vector (Thermo Fisher Scientific) for Sanger sequencing with the M13F and M13R primers. Heterozygous knockin clones carrying both wildtype allele and the $p16^{INK4A\text{-}P2A\text{-}mCherry}$ knockin allele were selected for following functional characterization.

CRISPR library construction and screening

A set of 2,029-sgRNA oligos that target on H3K27ac and ATAC-seq positive peaks defined in the TAD containing *INK4*/*ARF* in IMR90, HCT116 and SEM cells as well as an additional 20 non-targeting control sgRNAs with no detectable match to the human genome were designed for array-based oligonucleotide synthesis (CustomArray). Unique binding of each sgRNA was verified by sequence blast against the whole human genome. To construct a sgRNA pooled library targeting on human TFs, seven gRNAs against each of the 1,639 human TFs were obtained from validated sgRNA libraries published previously (2-10). For each library, the synthesized oligo pool was amplified by PCR and cloned into LentiGuide-Puro backbone (#52963) by in-fusion assembly (Clontech). The $p16^{INK4A\text{-}P2A\text{-}mCherry}$ reporter cell line was overexpressed with lentiviral Cas9 or dCas9-

KRAB followed with infection of pooled sgRNA library at low M.O.I (~0.3). Infected cells were selected by blasticidine and puromycin and sorted for mCherry^{High} and mCherry^{Low} populations. The sgRNA sequences were recovered by genomic PCR and Deep Sequencing using MiSeq for single-end 150-bp (Illumina). The primer sequences used for cloning and sequencing are listed in *SI Appendix*, Dataset S1. The sgRNA sequences are described in *SI Appendix*, Datasets S2 and S3. High-titer lentivirus stocks were generated in 293T cells as previously described (1).

Data analysis of CRISPR screening

The raw FASTQ data were de-barcoded and mapped to the original reference sgRNA library. The differentially enriched sgRNAs were defined by comparing normalized counts between sorted cells of top 10% and bottom 10% of mCherry-expressing bulk populations. At least three independent replicate screenings were performed with reporter cell lines stably expressing Cas9, dCas9-KRAB and no effector control. Normalized counts for each sgRNA were extracted and used for identifying differentially enriched sgRNA by DESeq2 (11). The combined analysis of seven sgRNAs against each human TF was conducted by MAGeCK algorithm (12). Detailed screening results were included in *SI Appendix*, Datasets S2 and S3.

Flow cytometry

Suspension cultured SEM were collected by centrifugation at 800X*g*, filtered through a 70-μm filter, and analyzed for mCherry on a BD FACS Aria III flow cytometer with a

115 negative control. The 4,6-diamidino-2-phenylindole (DAPI) staining was conducted prior
116 to sort to exclude dead cells.
117
118 **Capture-C**
119 Next-generation Capture-C was performed essentially as described by others (13). In
120 brief, about $1X10^7$ cells were collected per Capture-C reaction. Cells were freshly fixed in
121 1% formaldehyde for 10 mins at room temperature by rotation then quenched by 0.125
122 M glycine followed by resuspension in 5 ml lysis buffer and incubated on ice for 20 mins.
123 The pellets were re-suspended in 1 mL of 1X DpnII buffer and transferred to a 2 mL tube
124 for homogenization followed with centrifugation at 14,000 rpm for 5 mins at 4°C. The
125 supernatant was removed and discarded. The remaining pellets underwent digestion,
126 ligation, and de-crosslinking. After de-crosslinking and DNA precipitation, a small portion
127 of DNA was run on a gel to check for digestion and ligation efficiency. Once the 3C library
128 construction was complete, 5-6 µg of the 3C DNA library was sheared ~200-bp followed
129 with end-repair, adaptor ligation, and PCR enrichment of adaptor ligated DNA. Two
130 rounds of captures were conducted using 1.5-2 µg of DNA, 5 µg COT DNA, and 1 µl xGen
131 Universal Blocking Oligo i7 and 1 µl xGen Universal Blocking Oligo i5 in a 1.5 mL tube.
132 Samples were transferred to thermal cycler at 47 °C for 72 hours for the first capture and
133 16-20 hours for the second capture. The constructed library was sequenced using the
134 Illumina MiSeq System for paired-end 150-bp. The paired-end reads were reconstructed
135 into single reads with FLASH, digested in silico with the DpnII2E.pl script, and aligned to
136 the human genome using Bowtie 1.0.0 (http://bowtie-bio.sourceforge.net/index.shtml).
137 Interaction frequencies were determined using CCanalyser2.pl

(https://github.com/telenius/captureC/releases). To quantify interaction frequency between bait sequence and target regions, signal value from .bw file was analyzed by "bigWigToBedGraph", and then normalized by probe signal for each experiment (14).

CUT&RUN assay

Three million cells were collected for each sample. Cells were pelleted at 600X*g* for 3 mins at room temperature. Cell pellets were washed twice with 1.5 ml room temperature wash buffer. Cell pellets were resuspended in 1 ml wash buffer at room temperature. While gently mixing, 100 µl concanavalin beads (BioMagAPlus Concanavalin A, Polysciences, 86057-3; prewashed and resuspended in 100 µl binding buffer) were added to the samples, followed with gentle rotation for 10 mins at room temperature. Samples were placed on a magnetic stand and the supernatant was removed. Beads were resuspended in 200 µl AB buffer (Dig-wash buffer with 0.02% digitonin and 0.5 M EDTA) with either ZNF217 antibody (Thermo Fisher Scientific, 720352), Cas9 antibody (Clontech, 632606) or YY1 antibody (Active Motif, 61779). Samples were rotated at 4 °C for 2 hours. Samples were placed on a magnetic stand and supernatant was removed. Beads were gently resuspended in 200 µl Dig-wash buffer with 0.02% digitonin and 2 µl PA-MN (home-made reagent kindly provided by Dr. Steven Henikoff), followed by rotation at 4 °C for 1 hour. Samples were placed on a magnetic stand and supernatant removed. Beads were washed for a total of three times in 1 ml Dig-wash buffer, pipetting gently. Following the final wash, beads were resuspended in 100 µl Dig-wash buffer by gentle mixing and transferred to a new tube. On ice, 2 µl 100 mM $CaCl_2$ was added to the sample with gentle mixing. Samples were incubated on ice for 30 mins. Stop buffer (100 µl) was

added to the samples and samples were incubated for 10 mins at 37 °C followed by centrifugation at 4 °C, 12,000 rpm. Samples were placed on a magnetic stand and the supernatant containing the cut and run fragments was collected. To the samples, 2 µl 10% SDS and 2.5 µl proteinase K were added and the samples were incubated at 50 °C overnight. DNA was extracted by phenol/chloroform/isopropanol protocol. Library construction was performed using the NEBNext UltraII DNA Library Prep Kit from NEB (E7645S). Indexed samples were run using the Illumina MiSeq V3 600-cycle kit (MS-102-3003). CUT&RUN raw reads were mapped to genome hg19. by bowtie 2.3.4 with default parameter. Then mapping file were converse to .bw file by bamCoverage (15, 16).

Immunoblotting

Cells were treated with RIPA buffer, then the lysates were subjected to SDS-PAGE (Thermo Fisher Scientific) and transferred to a PVDF membrane (Bio-Rad) in accordance with the manufacturer's protocols at a constant 100 V for 1 hour. After incubation with 5% nonfat milk in TBS-T (10 mM Tris, pH 8.0, 150 mM NaCl, 0.5% Tween-20) for 1 hour at room temperature, the membrane was incubated with antibodies against GAPDH (Thermo Fisher Scientific, AM4300, diluted 1:5,000), ARF (abcam, ab185620, diluted 1:1,000), $p16^{INK4A}$ (abcam, ab189034, diluted 1:1,000), p-RB (Ser807/811, Cell Signaling, 8516, diluted 1:1,000), p-RB (Ser780, Cell Signaling, 8180, diluted 1:1,000), RB (Cell Signaling, 9309, diluted 1:1,000) and YY1 antibody (Active Motif, 61779, diluted 1:1,000) at 4°C for 48 hours with gentle shaking. Membranes were washed three times for 30 min and incubated with a 1:2,000 dilution of horseradish peroxidase–conjugated secondary antibodies for 2 hours at room temperature. Blots were washed with TBS-T three times

for 30 mins and developed with the ECL system (Amersham Biosciences) in accordance with the manufacturer's instructions.

Fluorescence *in situ* hybridization

An 800-bp (homology arms removed) purified DNA (*P2A-mCherry*) sequence was labeled with a red-dUTP (AF594, Molecular Probes) by nick translation, and an *INK4*/*ARF* fosmid clone (WI2-2622K01/9p21.3) was labeled with a green-dUTP (AF488, Molecular Probes). Both labeled probes were combined with sheared human DNA and independently hybridized to fix the interphase and metaphase nuclei derived from each sample using routine cytogenetic methods in a solution containing 50% formamide, 10% dextran sulfate, and 2XSSC. The cells were then stained with DAPI and analyzed.

Quantitative real-time PCR

Total RNA was collected from 1 million cells of SEM or SK-N-SH by using TRIzol (Thermo Fisher Scientific, 15596026). Reverse transcription was performed using a High-Capacity cDNA Reverse Transcriptase Kit (Applied Biosystems, 4374966). Real-time PCR was performed using FAST SYBR Green Master Mix (Applied Biosystems, 4385612) in accordance with the manufacturer's instructions and primers to detect $p16^{INK4A}$, *ARF,* *mCherry*, and *GAPDH*. Relative gene expression was determined by the ΔΔ-CT method (17). qRT-PCR primers were listed in *SI Appendix*, Dataset S1.

Statistics

All values are shown as the mean ± SEM. Statistical analyses were performed with GraphPad Prism software, version 6.0. *P*-values were calculated by a two-tailed *t*-test.

Abbreviations

CRISPR, clustered regularly interspaced short palindromic repeats; DSB, DNA double-strand break; PAM, protospacer adjacent motif; INDEL, insertion and deletion; bp, base pair; ChIP, chromatin immunoprecipitation; TAD, topologically associated domain; TF, transcription factor; TSS, transcription start site.

SI Appendix Figure Legends

Fig. S1 Functional validation of *p16$^{INK4A\text{-}P2A\text{-}mCherry}$* allele by CRISPR interference (CRISPRi)

(A) The *p16$^{mCherry/+}$* reporter cell line was transduced with lentiviral dCas9-KRAB and either NT-sgRNA or p16^{INK4A}-sgRNA-1 that targeted against the *p16^{INK4A}* promoter. Flow cytometry analysis of mCherry fluorescence was demonstrated in comparison to NT-sgRNA targeted cells along with parental SEM cells.

(B) qRT-PCR was performed to quantify *p16^{INK4A}* mRNA expression in *p16$^{mCherry/+;dCas9\text{-}KRAB}$* cells targeted with either NT-sgRNA or p16^{INK4A}-sgRNA-1 (N=4).

(C) Immunoblotting was performed on *p16$^{mCherry/+;dCas9\text{-}KRAB}$* cells targeted with either p16^{INK4A}-sgRNA-1 or NT-sgRNA by using antibodies against p16^{INK4A}, p-RB1, RB1 and GAPDH.

(D) Cell number was counted at day 0, 1 and 3 on *p16*$^{mCherry/+;dCas9\text{-}KRAB}$ cells targeted with either p16^{INK4A}-sgRNA-1 or NT-sgRNA.

Fig. S2 Characterization of non-coding CRISPR sgRNA library by Deep-Sequencing

(A) Maxiprep plasmid DNA containing the CRISPR sgRNA library designed against the entire *INK4*/*ARF* locus was amplified and sequenced by next-generation sequencing. Counts of each sgRNA were illustrated as $\log_2$ ratio.

(B) Deep-sequencing and raw count analysis of the distribution of all sgRNAs 48 hours after library was infected into human IMR90 and HCT116 cells.

(C) Pearson's Correlation analysis of the distribution of all sgRNAs amplified from library-infected human IMR90 and HCT116 cells in comparison to maxiprep plasmid DNAs.

Fig. S3 Optimization of the sorting gate of the *p16*$^{INK4A\text{-}P2A\text{-}mCherry}$ population in non-coding CRISPR screening

(A) Flow cytometry sorting was performed in CRISPR sgRNA library–targeted *p16*$^{mCherry/+;dCas9\text{-}KRAB}$ cells for mCherry^{High} and mCherry^{Low} populations, which were gated at the top and bottom of 5%, 10%, 12%, and 15% sorting gates.

(B) The number of sgRNAs that were enriched in the positive-control locus (*p16*^{INK4A} promoter) in each sort was calculated and compared to the total sgRNAs in these regions to determine the enrichment percentage.

(C) Pearson's Correlation of the global sgRNA distribution was calculated among the 5%, 10%, 12% and 15% sort experiments.

Fig. S4 Enriched sgRNAs in CRISPR/dCas9-KRAB screenings

(A) Volcano plot illustration of CRISPR/Cas9 screening result in combination with arrayed sgRNAs spanning the ATAC-seq and H3K27ac peaks in the *INK4*/*ARF* locus. The distribution of each sgRNA in the top 10% and bottom 10% of the mCherry sorted population was calculated for a given fold change of $\log_2$(High/Low) (x-axis) and *P*-value (y-axis). Each screening was performed three times to calculate the *P*-value. Under the cut-off of an adjusted *P*-value of ≤0.01 and a fold-change of $\log_2^{\text{High/Low}}$ ≥1 or ≤−1, purple dots represent sgRNAs designed to target the $p16^{INK4A}$ promoter. Red dots represent the non-targeting negative-control sgRNAs. Green dots indicate sgRNAs enriched in the *ARF* promoter. Pink dots indicate sgRNAs enriched in the $p15^{INK4B}$ promoter.

(B) Volcano plot demonstration of CRISPR/dCas9-KRAB screening result in combination with arrayed sgRNAs spanning the ATAC-seq/H3K27ac peaks in the *INK4*/*ARF* locus.

(C) Volcano plot demonstration of the results of no effector screening in combination with arrayed sgRNAs spanning the ATAC-seq/H3K27ac peaks in the *INK4*/*ARF* locus.

(D) The distribution of sgRNAs in selected viewpoints at *INK4*/*ARF* locus and a distal non-coding region in three screenings using dCas9-KRAB, Cas9, and no effector control.

Fig. S5 Distribution of the full array of sgRNAs from dCas9-KRAB and Cas9 screening. All of the 2029 sgRNAs represented in the screening were plotted to genome browser based on enrichment score between $\text{mCherry}^{\text{High}}$ and $\text{mCherry}^{\text{Low}}$.

Fig. S6 Distal repressive elements residing 3' and adjacent to the *ARF* promoter contribute to the repression of *p16^{INK4A}*

(A) The physical location of all four sgRNAs targeting the *p16^{INK4A}* repressive element adjacent to the *ARF* promoter (ARF-sgRNA-1, 2, 3 and 4).

(B) Validation of four randomly selected sgRNAs targeting the *p16^{INK4A}* promoter (p16^{INK4A}-sgRNA-1, 2, 3 and 4) and two sgRNAs targeting the *ARF* promoter (ARF-sgRNA-1 and 2) outside the 42-bp core sequence. Lentiviral sgRNA infection into *p16$^{mCherry/+;dCas9\text{-}KRAB}$* cells was performed, followed by flow cytometry analysis of the mCherry density. The *p*-values were calculated by a two-tailed *t*-test.

Fig. S7 Acute depletion of CTCF in SEM cells by auxin-inducible degron (AID) system did not affect expression of *p16^{INK4A}*

(A) ChIP-seq tracks of CTCF, SMC3 and YY1 on the *INK4*/*ARF* locus based on ENCODE data from GM12878 and SK-N-SH cells.

(B) Schematic illustration of working mechanism of auxin-inducible degron system.

(C) Schematic diagram of donor vector design for miniAID-mClover3 knock-in to C-terminus of human *CTCF* before the stop codon. HA: homology arm. Single-guide RNA (sgRNA) and PAM sequences were added to each end of the HA.

(D) A Bi-allelic knock-in clone and a mono-allelic knock-in clone were used to detect protein degradation after IAA treatment for 24 hours. The CTCF-miniAID-mClover3 fusion protein was 35 KD larger than endogenous CTCF. GAPDH was included as loading control.

(E) CUT&RUN of CTCF was conducted in SEM$^{CTCF\text{-}AID}$ knockin cells (clones 35 and 42) treated with or without IAA for 48 hours. A representative set of consistent data from one clone (clone 42) was shown. Tracks were shown at the viewpoint of *INK4*/*ARF* locus.

(F) qRT-PCR analysis of *p16*^{INK4A} was conducted in three bi-allelic knock-in clones in response to CTCF depletion.

(G) Immunoblotting was performed using antibody against p16^{INK4A} in four individual bi-allelic knock-in clones in response to CTCF depletion.

REFERENCES

1. Vo BT*, et al.* (2017) Inactivation of Ezh2 Upregulates Gfi1 and Drives Aggressive Myc-Driven Group 3 Medulloblastoma. *Cell Rep* 18(12):2907-2917.
2. Wang T*, et al.* (2015) Identification and characterization of essential genes in the human genome. *Science* 350(6264):1096-1101.
3. Doench JG*, et al.* (2016) Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol* 34(2):184-191.
4. Sanjana NE, Shalem O, & Zhang F (2014) Improved vectors and genome-wide libraries for CRISPR screening. *Nat Methods* 11(8):783-784.
5. Ma H*, et al.* (2015) A CRISPR-Based Screen Identifies Genes Essential for West-Nile-Virus-Induced Cell Death. *Cell Rep* 12(4):673-683.
6. Tzelepis K*, et al.* (2016) A CRISPR Dropout Screen Identifies Genetic Vulnerabilities and Therapeutic Targets in Acute Myeloid Leukemia. *Cell Rep* 17(4):1193-1205.
7. Hart T*, et al.* (2015) High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities. *Cell* 163(6):1515-1526.
8. Hart T*, et al.* (2017) Evaluation and Design of Genome-Wide CRISPR/SpCas9 Knockout Screens. *G3 (Bethesda)* 7(8):2719-2727.
9. Smith JR*, et al.* (2008) Robust, persistent transgene expression in human embryonic stem cells is achieved with AAVS1-targeted integration. *Stem Cells* 26(2):496-504.
10. Park RJ*, et al.* (2017) A genome-wide CRISPR screen identifies a restricted set of HIV host dependency factors. *Nat Genet* 49(2):193-203.
11. Love MI, Huber W, & Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15(12):550.
12. Li W*, et al.* (2014) MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol* 15(12):554.

13. Davies JO*, et al.* (2016) Multiplexed analysis of chromosome conformation at vastly improved sensitivity. *Nat Methods* 13(1):74-80.
14. Kent WJ, Zweig AS, Barber G, Hinrichs AS, & Karolchik D (2010) BigWig and BigBed: enabling browsing of large distributed datasets. *Bioinformatics* 26(17):2204-2207.
15. Langmead B & Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9(4):357-359.
16. Ramirez F, Dundar F, Diehl S, Gruning BA, & Manke T (2014) deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res* 42(Web Server issue):W187-191.
17. Schmittgen TD & Livak KJ (2008) Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc* 3(6):1101-1108.

Figure S1

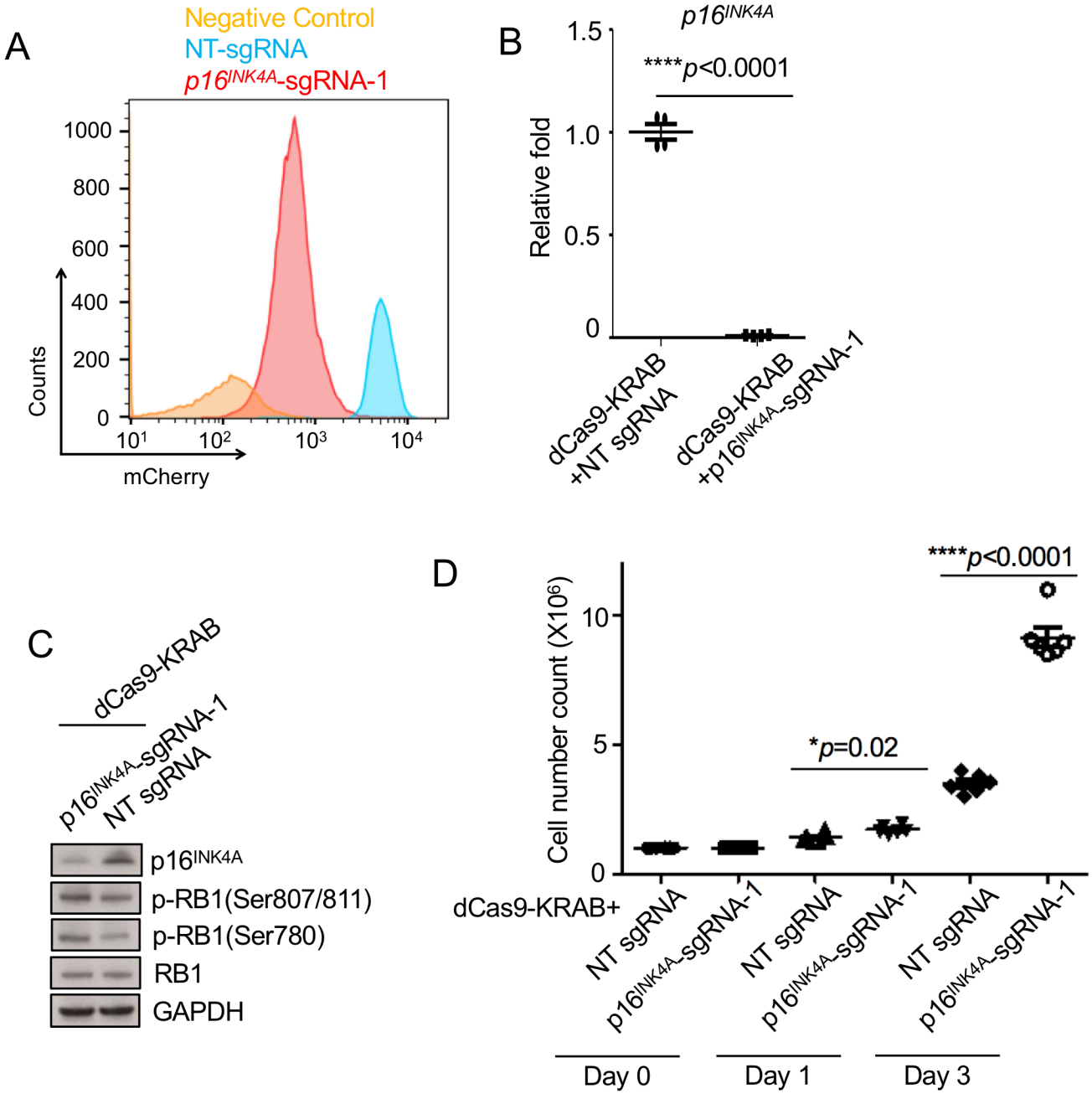

A
Negative Control
NT-sgRNA
p16INK4A-sgRNA-1
Counts
mCherry
B
p16INK4A
****p<0.0001
Relative fold
dCas9-KRAB +NT sgRNA
dCas9-KRAB +p16INK4A-sgRNA-1
C
dCas9-KRAB
p16INK4A-sgRNA-1
NT sgRNA
p16INK4A
p-RB1(Ser807/811)
p-RB1(Ser780)
RB1
GAPDH
D
****p<0.0001
*p=0.02
Cell number count (X10^6)
dCas9-KRAB+
NT sgRNA
p16INK4A-sgRNA-1
Day 0
Day 1
Day 3

Figure S2

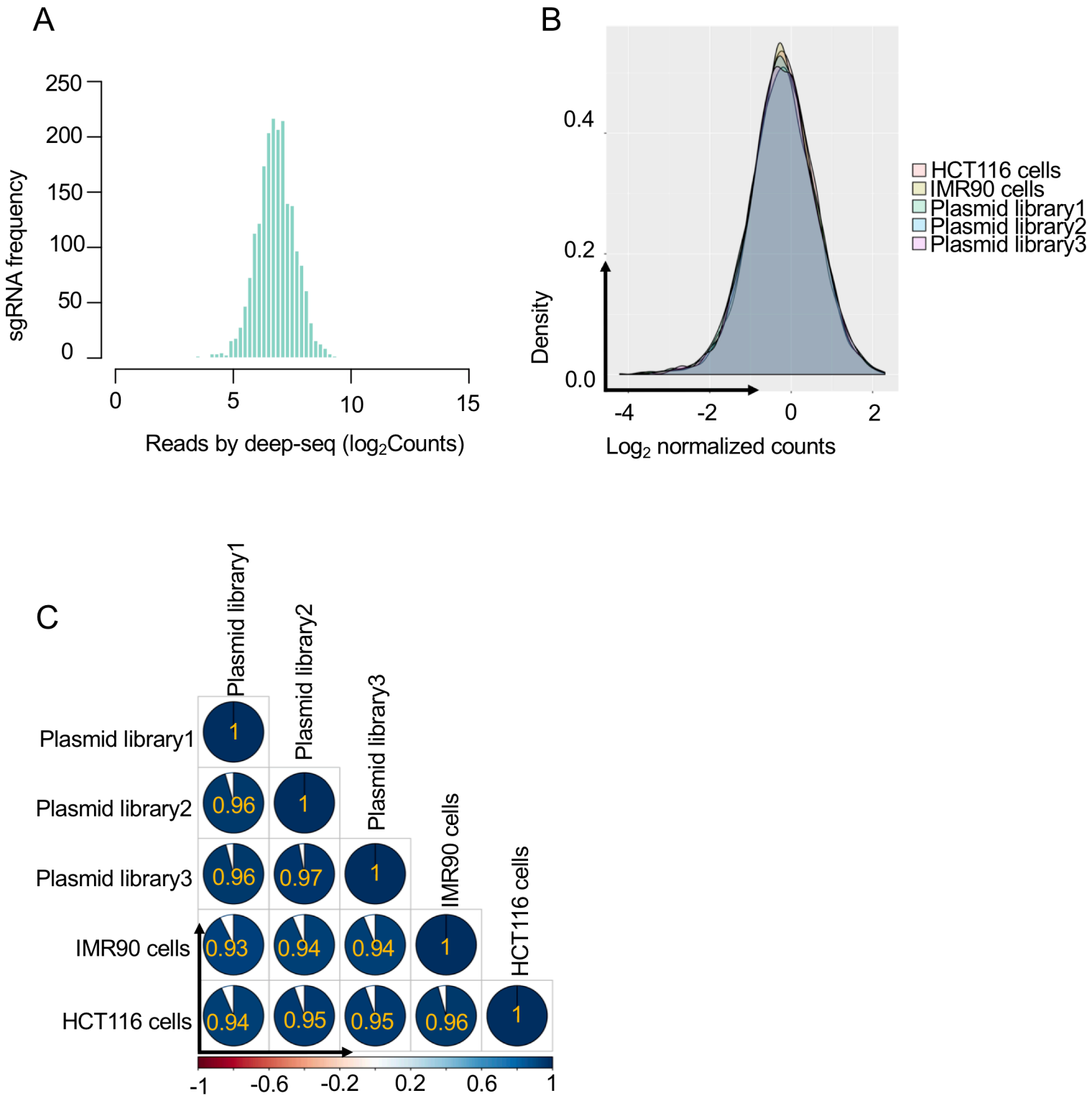

A
250
200
150
100
50
0
sgRNA frequency
0
5
10
15
Reads by deep-seq (log2Counts)
B
0.4
0.2
0.0
Density
-4
-2
0
2
Log2 normalized counts
HCT116 cells
IMR90 cells
Plasmid library1
Plasmid library2
Plasmid library3
C
Plasmid library1
Plasmid library2
Plasmid library3
IMR90 cells
HCT116 cells
1
0.96
1
0.96
0.97
1
0.93
0.94
0.94
1
0.94
0.95
0.95
0.96
1
-1
-0.6
-0.2
0.2
0.6
1

Figure S3

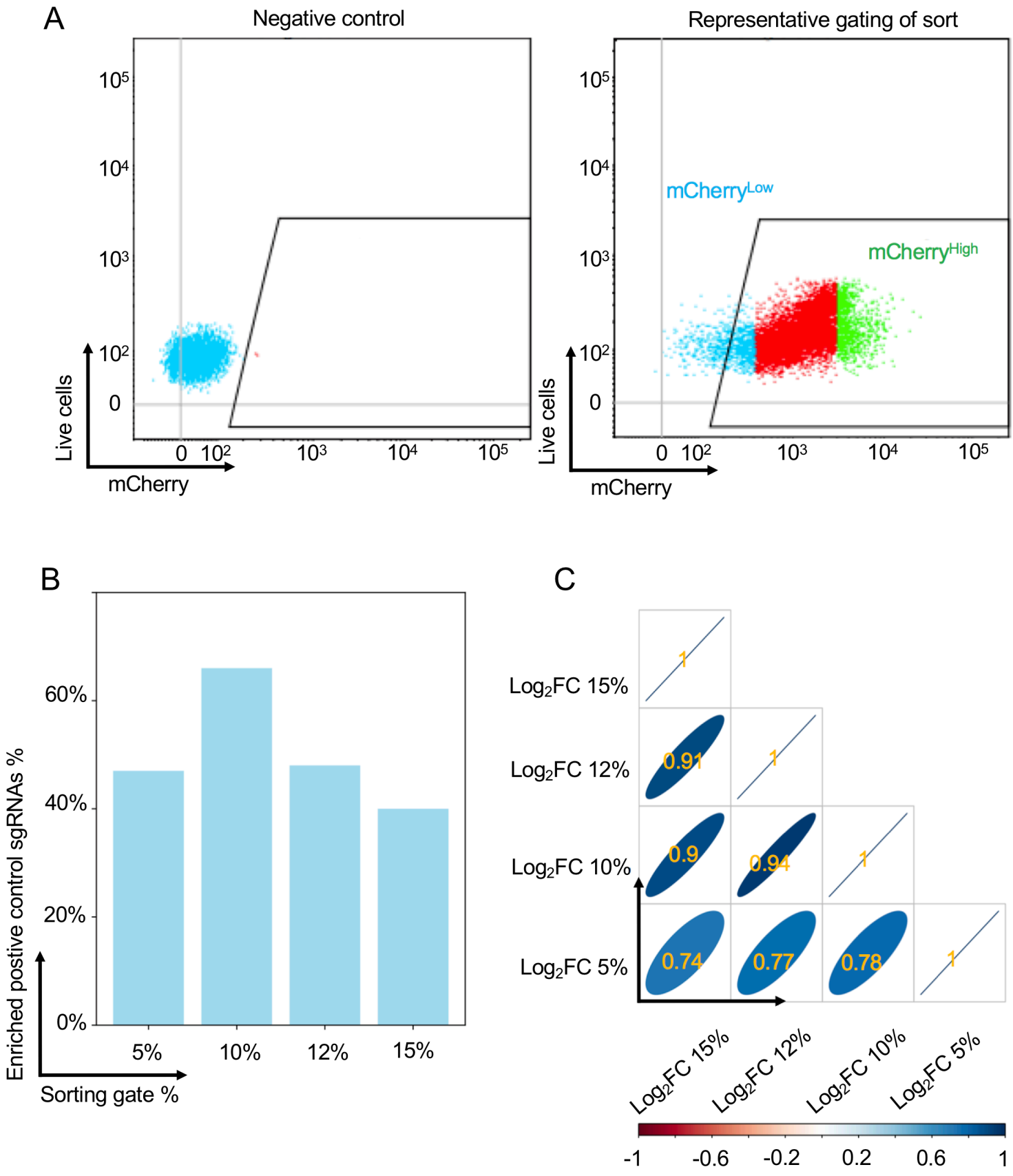

A
Negative control
Representative gating of sort
Live cells
mCherry
mCherryLow
mCherryHigh
B
Enriched postive control sgRNAs %
Sorting gate %
5%
10%
12%
15%
C
Log2FC 15%
Log2FC 12%
Log2FC 10%
Log2FC 5%
0.91
0.9
0.94
0.74
0.77
0.78
-1
-0.6
-0.2
0.2
0.6
1

Figure S4

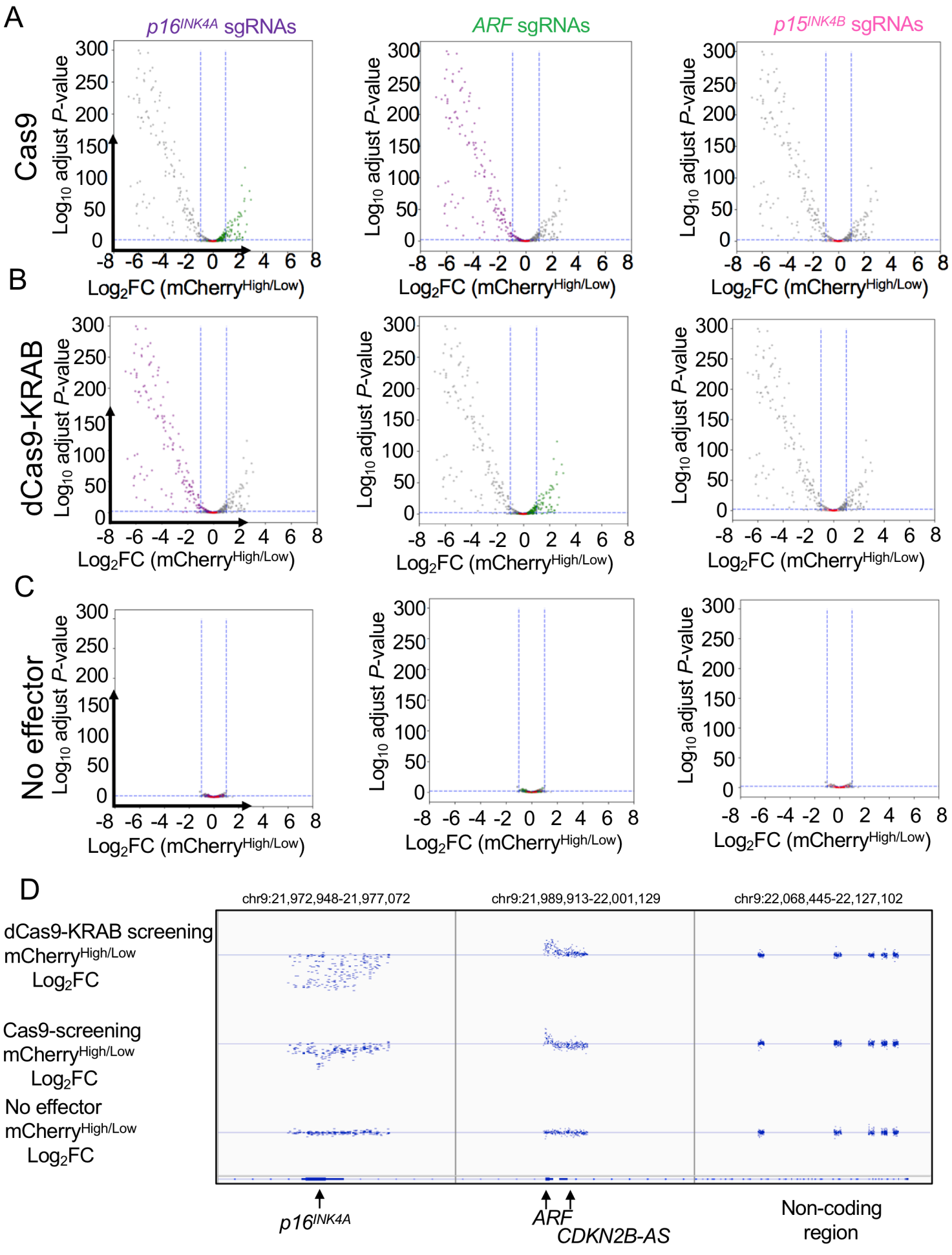

A
p16INK4A sgRNAs
ARF sgRNAs
p15INK4B sgRNAs
Cas9
dCas9-KRAB
No effector
Log10 adjust P-value
Log2FC (mCherryHigh/Low)
B
C
D
chr9:21,972,948-21,977,072
chr9:21,989,913-22,001,129
chr9:22,068,445-22,127,102
dCas9-KRAB screening
mCherryHigh/Low
Log2FC
Cas9-screening
mCherryHigh/Low
Log2FC
No effector
mCherryHigh/Low
Log2FC
p16INK4A
ARF
CDKN2B-AS
Non-coding
region

Figure S5

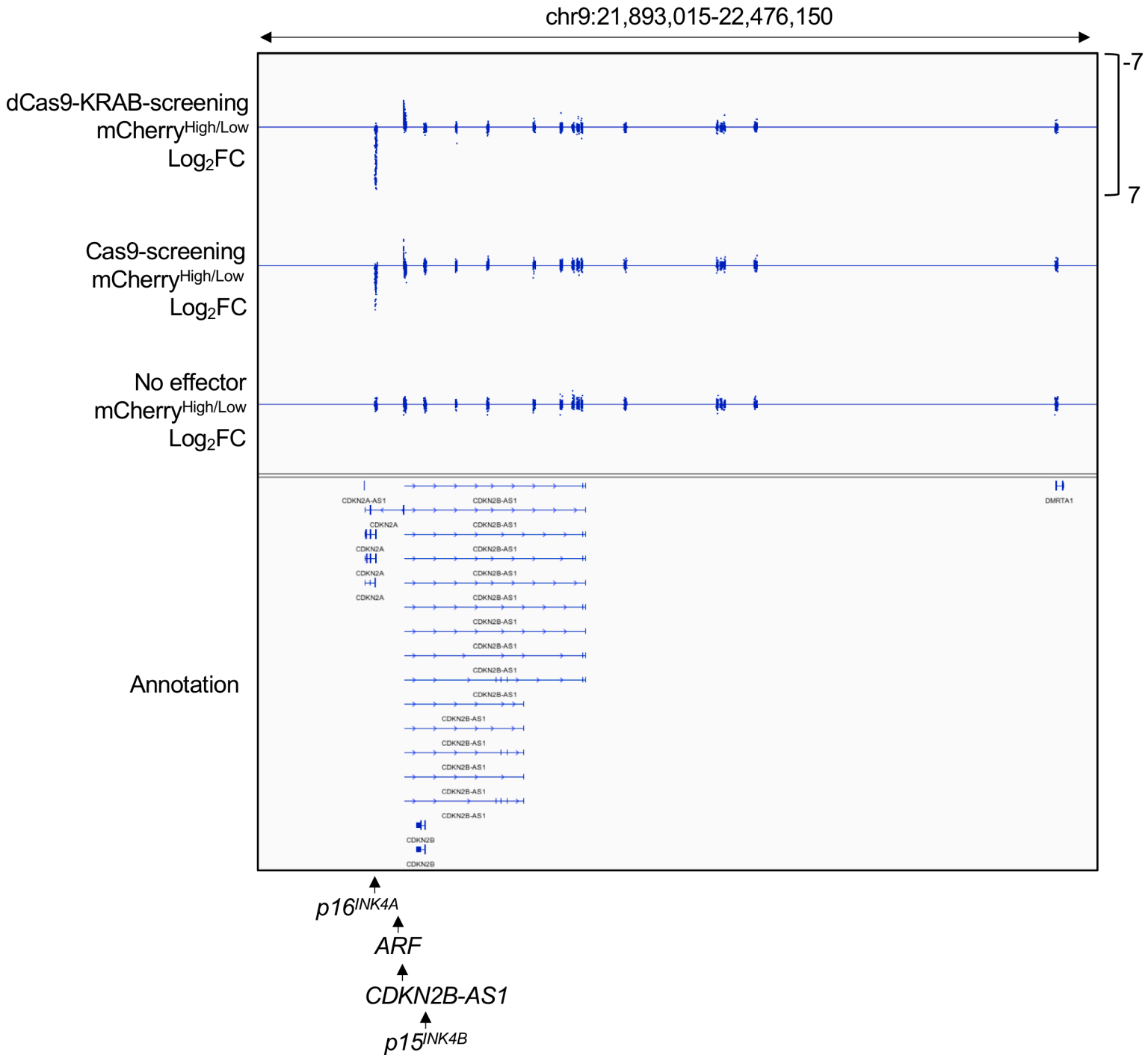

chr9:21,893,015-22,476,150
dCas9-KRAB-screening
mCherryHigh/Low
Log2FC
-7
7
Cas9-screening
mCherryHigh/Low
Log2FC
No effector
mCherryHigh/Low
Log2FC
Annotation
CDKN2A-AS1
CDKN2B-AS1
DMRTA1
CDKN2A
CDKN2B
p16INK4A
ARF
CDKN2B-AS1
p15INK4B

Figure S6

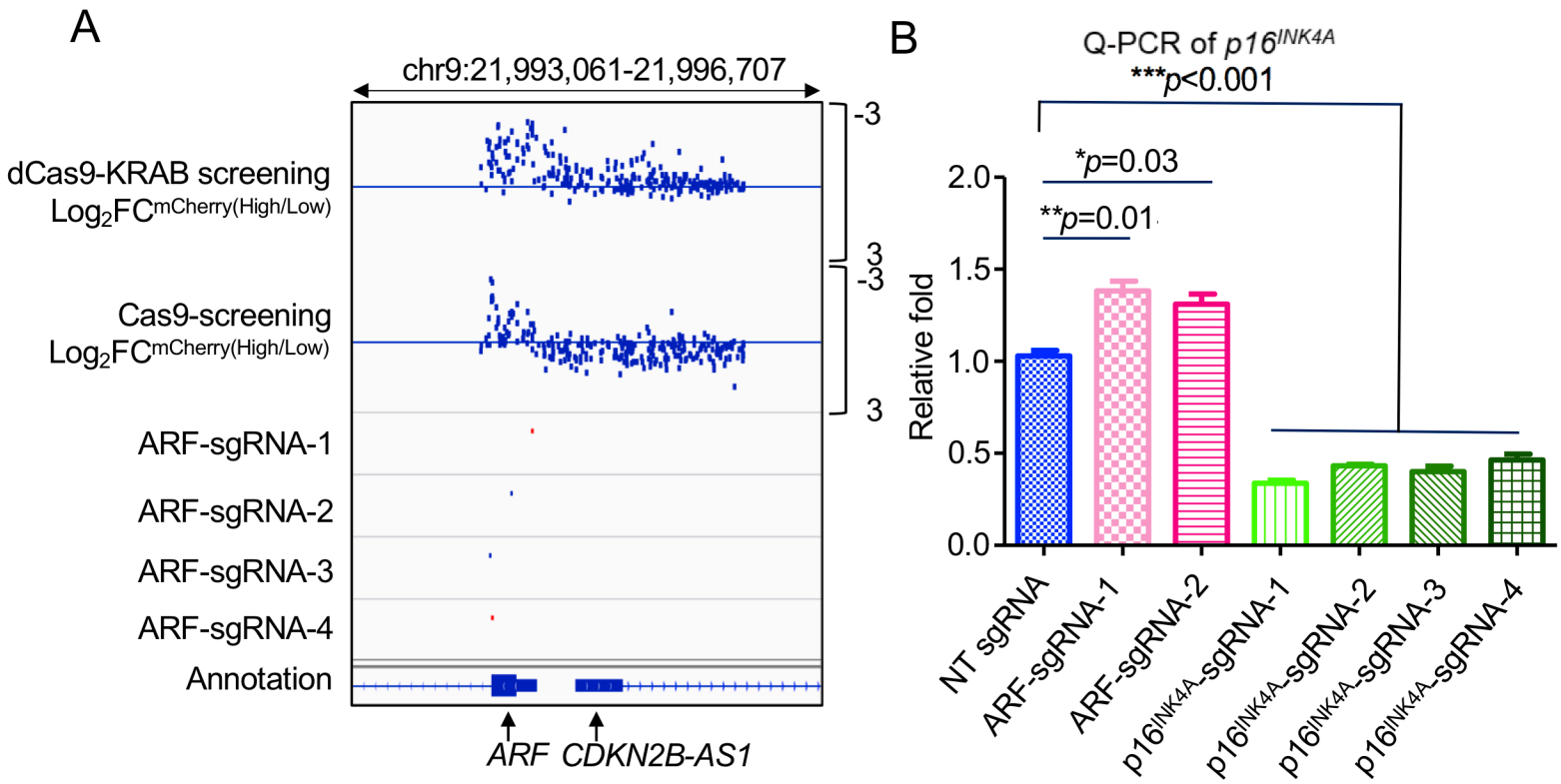
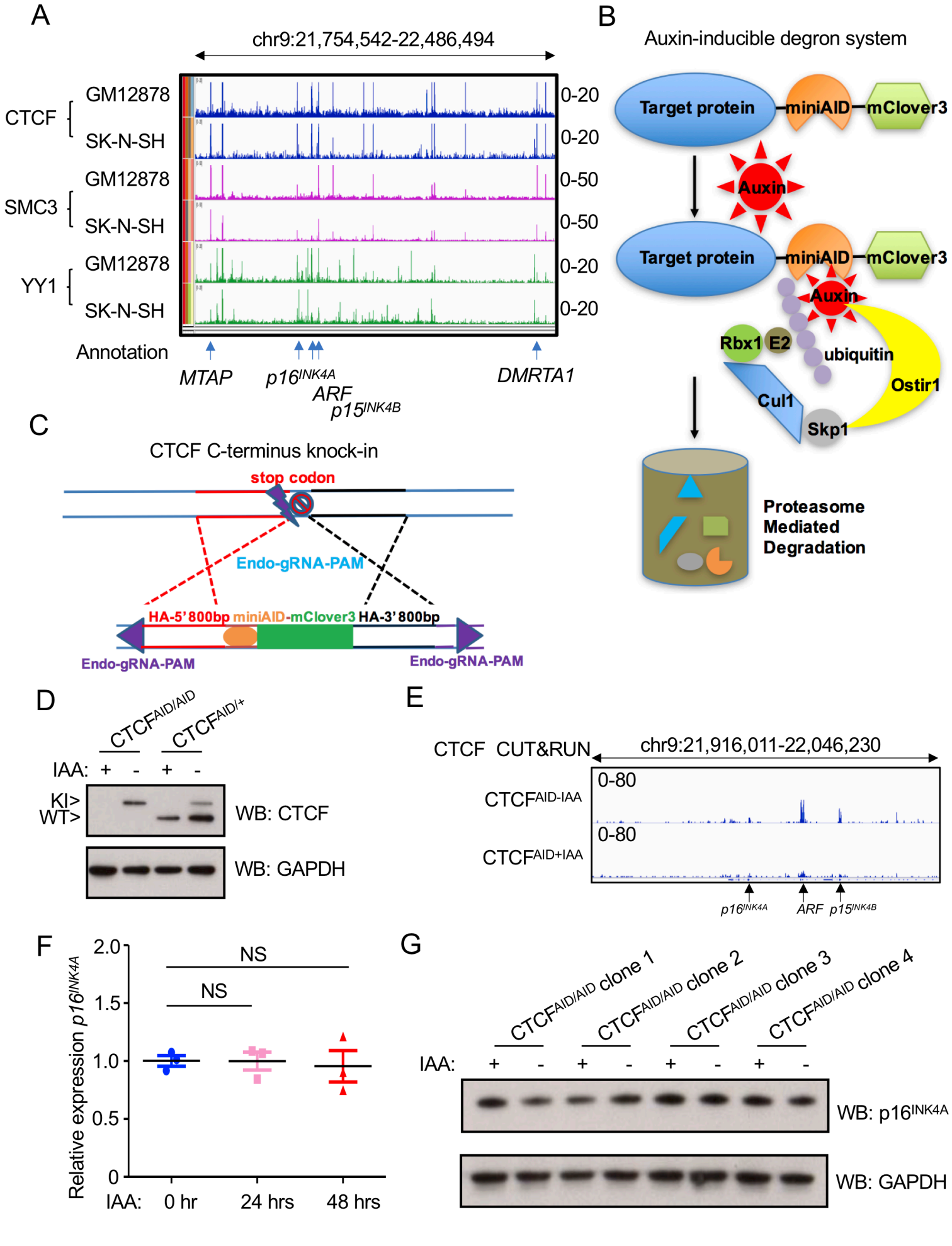

A
chr9:21,993,061-21,996,707
dCas9-KRAB screening
Log2FCmCherry(High/Low)
Cas9-screening
Log2FCmCherry(High/Low)
ARF-sgRNA-1
ARF-sgRNA-2
ARF-sgRNA-3
ARF-sgRNA-4
Annotation
-3
3
-3
3
ARF
CDKN2B-AS1
B
Q-PCR of p16INK4A
***p<0.001
*p=0.03
**p=0.01
Relative fold
2.0
1.5
1.0
0.5
0.0
NT sgRNA
ARF-sgRNA-1
ARF-sgRNA-2
p16INK4A-sgRNA-1
p16INK4A-sgRNA-2
p16INK4A-sgRNA-3
p16INK4A-sgRNA-4

Figure S7


A
chr9:21,754,542-22,486,494
CTCF
SMC3
YY1
GM12878
SK-N-SH
0-20
0-50
Annotation
MTAP
p16INK4A
ARF
p15INK4B
DMRTA1
B
Auxin-inducible degron system
Target protein
miniAID
mClover3
Auxin
Rbx1
E2
ubiquitin
Ostir1
Cul1
Skp1
Proteasome Mediated Degradation
C
CTCF C-terminus knock-in
stop codon
Endo-gRNA-PAM
HA-5' 800bp
miniAID-mClover3
HA-3' 800bp
D
CTCFAID/AID
CTCFAID/+
IAA:
KI>
WT>
WB: CTCF
WB: GAPDH
E
CTCF CUT&RUN
chr9:21,916,011-22,046,230
CTCFAID-IAA
CTCFAID+IAA
0-80
F
Relative expression p16INK4A
NS
2.0
1.5
1.0
0.5
0
0 hr
24 hrs
48 hrs
G
CTCFAID/AID clone 1
CTCFAID/AID clone 2
CTCFAID/AID clone 3
CTCFAID/AID clone 4
WB: p16INK4A
WB: GAPDH