PNAS
www.pnas.org

Supplementary Information for

An MXD1-derived repressor peptide identifies noncoding mediators of MYC-driven cell proliferation

Philipp Raffeiner, Jonathan R. Hart, Daniel García-Caballero, Liron Bar-Peled, Marc S. Weinberg, Peter K. Vogt

Corresponding author: pkvogt@scripps.edu

This PDF includes:

- SI Materials and Methods
- Supplemental Figures and Figure Legends
  Figure S1: CRISPR-Cas9-induced promoter deletions fail to inhibit MALAT1 expression.
  Figure S2: Transcriptome comparison between P493-6 and RAMOS, identification and validation of additional KRAB domains and sequence preference of sgRNAs.
  Figure S3: CRISPRi-based competition assay.
  Figure S4: Validation of selected noncoding loci.
  Figure S5: Conditional repression of endogenous loci.
- Supplemental Tables
  Table S1: Selected oligodeoxynucleotide sequences.
  Table S2: PCR-primers used to generate guide library for next generation sequencing.
  Table S3: Oligodeoxynucleotide and gBlock sequences for cloning of paired sgRNAs.

Other supplementary materials for this manuscript include the following:

Dataset S1: sgRNA sequences and their genomic location
Dataset S2: RNAseq of RAMOS and P493-6
Dataset S3: Results of CRISPR-sgRNA screens
Dataset S4: Results of TRIM28 proteomics

www.pnas.org/cgi/doi/10.1073/pnas.1921786117

SI Materials and Methods

DNA constructs. To construct lenti-2×Guide constructs, oligodeoxynucleotides containing the paired guide sequences separated by *BsmB*I restriction sites (Table S3) were PCR-amplified and Gibson assembly-cloned into the *BsmB*I sites of lentiGuide-Puro (Addgene # 52963) (1). The resulting vectors were cut with *BsmB*I to insert the PCR-amplified gBlock containing the scaffold sequence of sgRNA A and the mouse U6 (mU6) promoter sequence (Table S3).

For conditional MCP-SID expression, the coding sequence of MXD1 (aa 2 to 32) was PCR-amplified and cloned into pLVX-TetOne-Puro (Clontech, Takara Bio), cut with *EcoR*I/*BamH*I. Lenti-guide(MS2)_Zeo was prepared by replacing the puromycin resistance gene of lenti-sgRNA(MS2)_Puro for a Zeocin resistance gene, using *BsiW*I/*Mlu*I sites. TRIM28 coding sequence was PCR-amplified from P493-6 cDNA using a forward primer containing the Flag-tag sequence, and the PCR product was ligated into the lentiCas9_Blast vector backbone, cut with *Xba*I/*BamH*I to replace the Cas9 coding sequence (lenti-Flag-TRIM28_Blast).

CRISPR-Cas9-mediated deletion of promoter sequences. P493-6 cells stably expressing Cas9 were prepared as described in the manuscript using LentiCas9_Blast. Cells were then transduced with lenti-2×Guide constructs and selected with puromycin. CRISPR-Cas9-induced deletions were checked by PCR using genomic DNA as template and *MALAT1*-specific primers (Table S1).

TRIM28 proteomics. 1×10^7 P493-6 cells per sample (high or low MYC) stably expressing FLAG-tagged TRIM28 or FLAG-METAP2 (control) were washed with ice-cold PBS and resuspended in IP lysis buffer (20 mM HEPES pH 7.5, 200 mM NaCl, 5% glycerol, 0.75% Triton X-100) supplemented with protease inhibitors (Thermo Fisher Scientific). After sonication, cell debris was removed by centrifugation, and lysates were precleared using protein A sepharose beads. Cleared lysates were incubated with FLAG-M2 magnetic beads (Sigma, 100 µL 50:50 slurry) overnight while rotating at 4°C. Beads were washed four times with IP lysis buffer, and proteins were eluted from beads with 8 M urea. Protein samples were reduced in 10 mM DTT for 30 min at 65°C and subjected to alkylation with iodoacetamide (IAA) for 30 min at 37°C, all in 6 M urea. After proteolytic digestion with trypsin (Promega), digests were analyzed by liquid chromatography tandem mass spectrometry (LC-MS/MS). MS2 spectra data were extracted from the raw file using RAW Convertor (version 1.000; available at http://fields.scripps.edu/downloads.php). MS2 spectra data were searched using the ProLuCID algorithm (publicly available at http://fields.scripps.edu/downloads.php) using a reverse concatenated, non-redundant variant of the Human UniProt database (release-2012_11). Cysteine residues were searched with a static modification for carboxyamidomethylation (+57.02146) and one differential modification for oxidized methionine (+15.9949). Spectral counts for proteins from FLAG-TRIM28 immunoprecipitates were compared to spectral counts for proteins from FLAG-METAP2 immunoprecipitates. Interacting proteins were classified as those proteins whose corresponding peptides were enriched by greater than 20-fold in FLAG-TRIM28 immunoprecipitates compared to FLAG-METAP2 immunoprecipitates.

Analysis of ChIP-seq data. MYC ChIP-seq data for P493-6 was downloaded from NCBI SRA accession numbers SRR444429 and SRR444430 (2). The reads were aligned to human genome version hg38.d1.vd1 using BWA aligner (3). The aligned reads were sorted using Sambamba (4). The regions surrounding the targeted transcriptional start sites were selected using biomaRt (5). Histograms of the selected regions were produced using HOMER annotatePeaks (6) and graphed in R (7).

References:

1. Sanjana NE, Shalem O, & Zhang F (2014) Improved vectors and genome-wide libraries for CRISPR screening. Nat Methods 11(8):783-784.
2. Lin CY, et al. (2012) Transcriptional amplification in tumor cells with elevated c-Myc. Cell 151(1):56-67.
3. Li H & Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25(14):1754-1760.
4. Tarasov A, Vilella AJ, Cuppen E, Nijman IJ, & Prins P (2015) Sambamba: fast processing of NGS alignment formats. Bioinformatics 31(12):2032-2034.
5. Durinck S, Spellman PT, Birney E, & Huber W (2009) Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. Nat Protoc 4(8):1184-1191.
6. Heinz S, et al. (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell 38(4):576-589.
7. R Core Team (2014) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/

Supplemental Figures

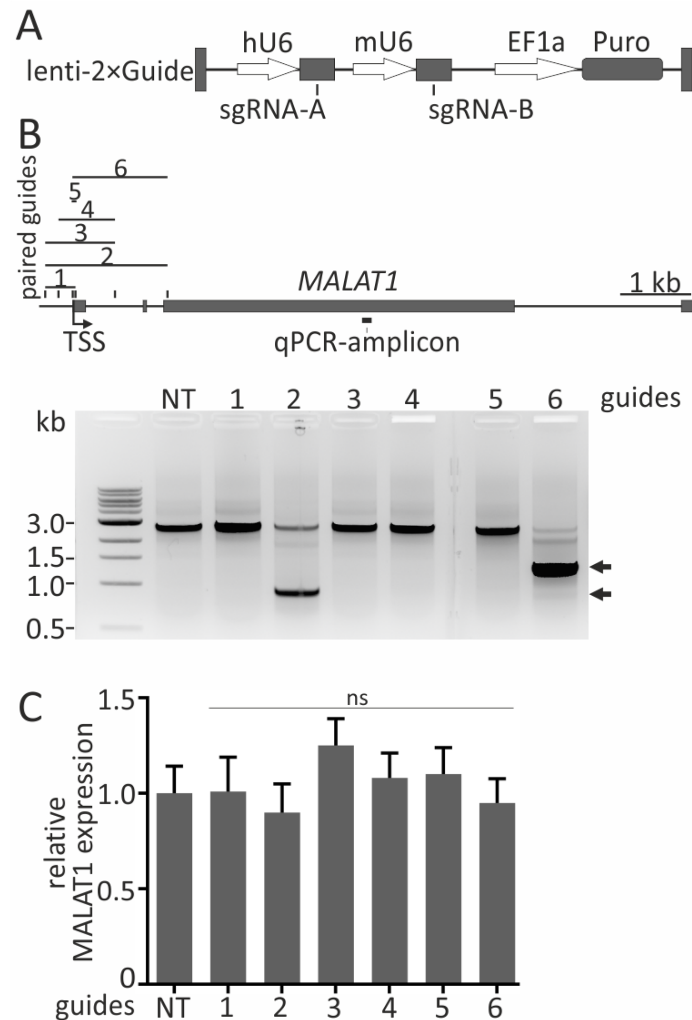

Fig. S1. CRISPR-Cas9-induced promoter deletions fail to inhibit MALAT1 expression. (A) Depiction of the lentiviral construct for expression of paired sgRNAs. EF1a, EF-1 alpha promoter; hU6, human U6 promoter; mU6, mouse U6 promoter; Puro, puromycin resistance gene. (B) Depiction of the *MALAT1* locus. Positions of sgRNAs and resulting deletions are indicated (upper panel). CRISPR-Cas9-induced deletions upon stable expression of Cas9 and indicated paired sgRNAs in P493-6 were investigated by PCR and agarose gel electrophoresis (lower panel). Arrows indicate PCR amplicons of successful deletions. NT, non-targeting guides; TSS, transcription start site. (C) Expression analysis by RT-qPCR using RNA from cells described in (B). Position of qPCR amplicon within MALAT1 is indicated in (B). ns, not significant.

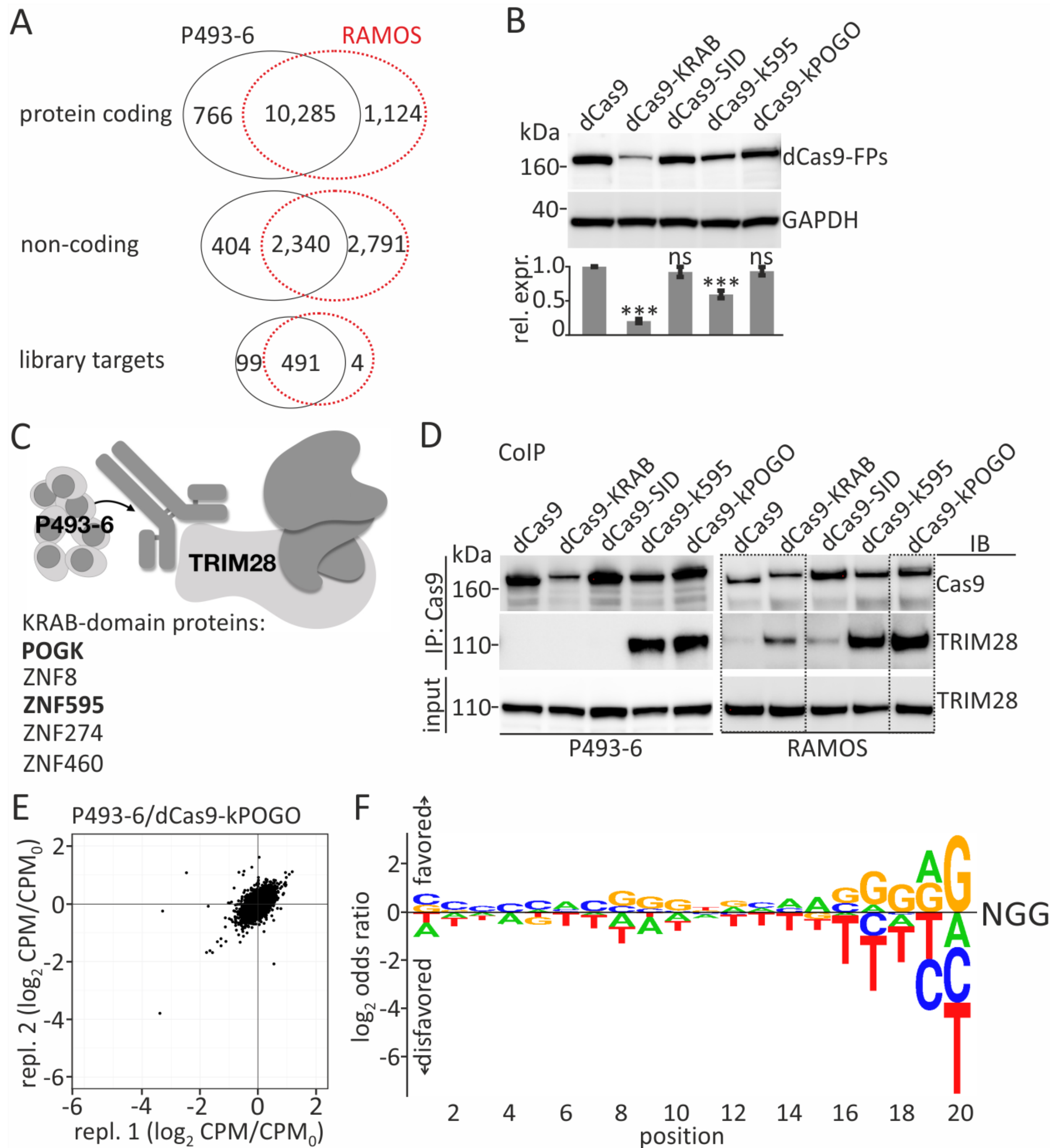

Fig. S2. Transcriptome comparison between P493-6 and RAMOS, identification and validation of additional KRAB domains and sequence preference of sgRNAs. (A) Venn diagram illustrating the transcriptome overlap between P493-6/MYC-ON and RAMOS, based on RNAseq data. (B) Expression of dCas9-fusion proteins in P493-6 cells was analyzed by immunoblotting. Densitometric quantification of three independent experiments, mean ± SD. ***, $p<0.001$ using unpaired t-test; ns, not significant. (C) TRIM28 was immunoprecipitated from P493-6 lysates, and interacting proteins were

identified by mass spectrometry. (D) Protein-protein interaction between indicated dCas9-fusion proteins and TRIM28 was determined by co-IP in P493-6 and RAMOS cells. Dotted lines highlight sections of the RAMOS blot that are shown in main Fig. 1. IB, immunoblot; IP, immunoprecipitation; k595, KRAB domain of ZNF595; kPOGO, KRAB domain of POGK, KRAB, KRAB domain of ZNF10; SID, SIN3 interacting domain. (E) Pooled MYCncLibrary screen in P493-6/dCas9-kPOGO. CPM, counts per million of CRISPRi sample; CPM_0, counts per million of control sample (P493-6, no dCas9). (F) Logo indicating the preference of nucleotide sequences that impact sgRNA efficiency. The heights of the letters are based on the $\log_2$ odds ratio of nucleotide frequency of efficient guides in the P493-6/dCas9-SID screen.

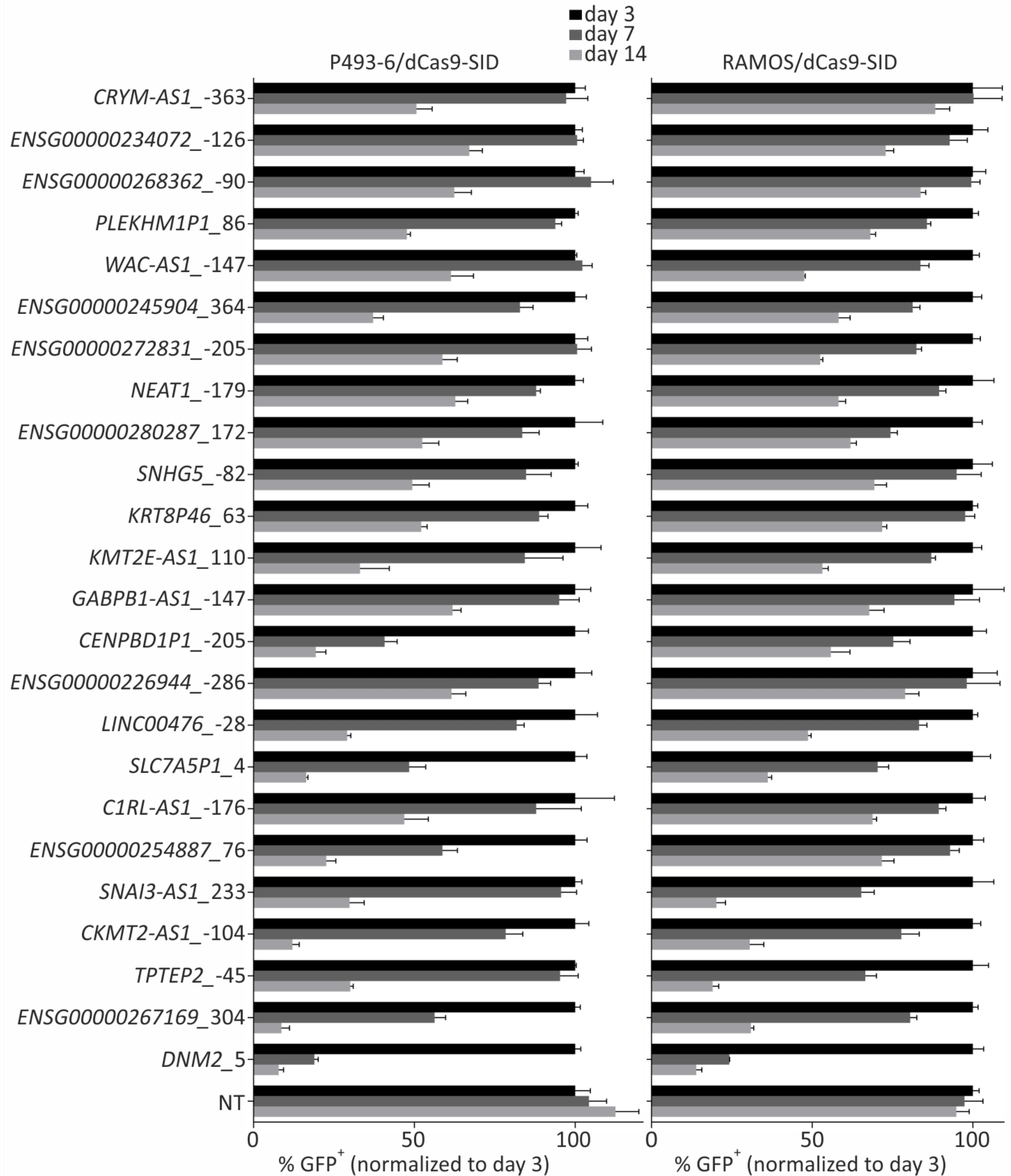

Fig. S3. CRISPRi-based competition assay. Cells were transduced with GFP-expressing sgRNA vectors targeting indicated genes (numbers indicate guide positions relative to TSS of gene) or a non-targeting (NT) control guide. GFP^+ cells were quantified by flow cytometry at indicated time points. Mean ± SD of three replicates.

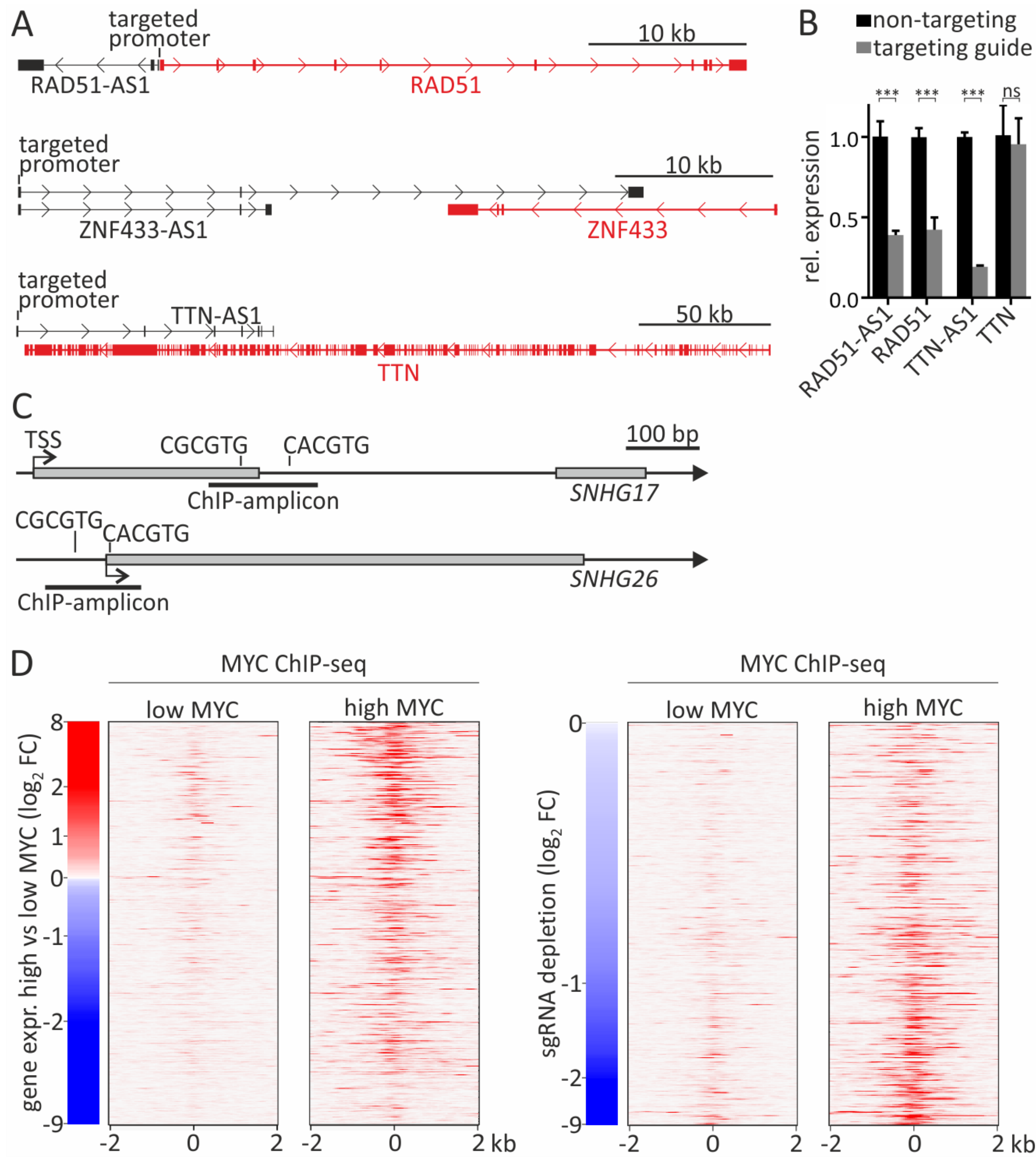

Fig. S4. Validation of selected noncoding loci. (A) Schematic depiction of sense-antisense pairs of the *RAD51-AS1*, *ZNF433-AS1* (two transcript variants are shown) and *TTN-AS1* loci. sgRNA-targeted promoters are indicated. (B) RT-qPCR experiments to determine CRISPRi-mediated repression of indicated sense-antisense gene pairs using guides that target the noncoding gene promoters in P493-6. Mean ± SD of three replicates. ***, p<0.001; ns, not significant; determined using unpaired t-test. (C) Schematic depiction of the proximal promoters of *SNHG17* and *SNHG26*, the positions of the E-boxes are indicated. TSS, transcription start site. (D) P493-6 MYC ChIP-seq data analyses of library target genes ranked by MYC-dependent gene expression (left panel) or by sgRNA depletion upon library screens in P493-6 (right panel).

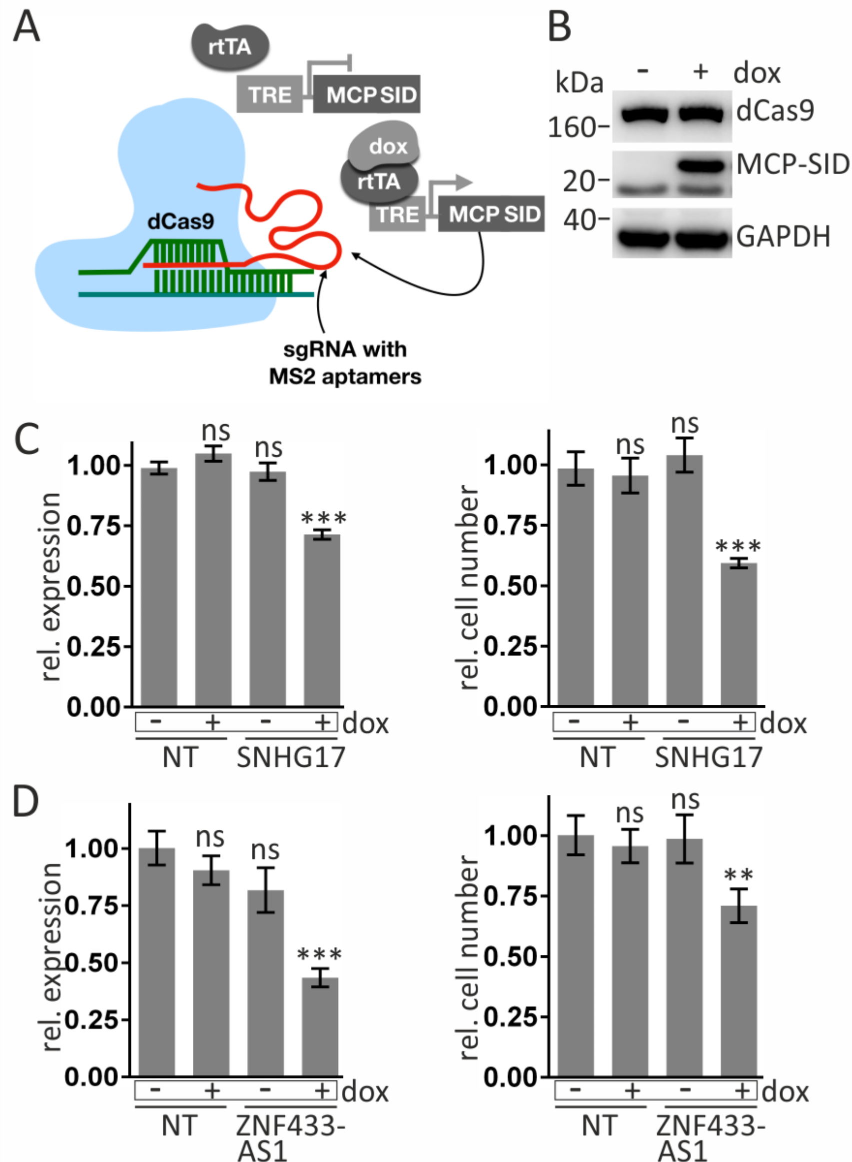

Fig. S5. Conditional repression of endogenous loci. (A) Schematic depiction of Doxycycline (dox)-regulatable (Tet-On) expression of MCP-SID that binds to the sgRNA-dCas9 complex to silence gene expression (rtTA, tetracycline transactivator; TRE, tetracycline response element). (B) dCas9-expressing RAMOS cells were stably transduced with a lentiviral vector that encodes for MCP-SID, controlled by tetracycline-response elements. Cells were incubated for 24 h with or without 0.1 µg/ml dox and then subjected to Western blot analysis using antibodies directed against Cas9, MS2 and GAPDH. (C) and (D) Cells described in (B) were stably transduced with lentiviral vectors for expression of a non-targeting (NT) guide or guides targeting *SNHG17* or *ZNF433-AS1*. Cells were grown with or without dox for 4 d followed by RNA isolation and RT-qPCR analysis (left panels). Cell counts after 4 d incubation with or without dox (right panels). Mean ± SD of three independent experiments. ***, p-value < 0.001; **, p-value < 0.01, using unpaired t-test; ns, not significant.

Supplemental Tables

Table S1. Selected oligodeoxynucleotide sequences.

Target	Forward sequence (5' to 3')	Reverse sequence (5' to 3')
Guide sequences (Figs. 2B, 3, S3 and S4)		
Non-targ. (NT)	CACCGCGAATTAGAAGCGTATTCCG	AAACCGGAATACGCTTCTAATTCGC
DNM2	CACCGCGCCTGAGAACCGGATGAGG	AAACCCTCATCCGGTTCTCAGGCGC
MALAT1	CACCGGAGGGACTGCGCAACCGGTG	AAACCACCGGTTGCGCAGTCCCTCC
RAD51-AS1	CACCGCGGGGTGAAGTCGGAGCGCG	AAACCGCGCTCCGACTTCACCCCGC
SNHG17	CACCGGCCCTGGCGTCGAGTGGCGG	AAACCCGCCACTCGACGCCAGGGCC
SNHG26	CACCGCCCTAACGGAGGCTTCCCCG	AAACCGGGGAAGCCTCCGTTAGGGC
TTN-AS1	CACCGACTCCTGGCGTTGCGGAAGG	AAACCCTTCCGCAACGCCAGGAGTC
ZNF433-AS1	CACCGCGGAGCCGGAAACCCCGAGG	AAACCCTCGGGGTTTCCGGCTCCGC
RT-qPCR primers (Figs. 2B, 4, S3 and S4)		
DNM2	GATGAAGTCCGGCAGGAGATTG	GTCGATGAGGGTCAAGTTCAAC
MALAT1	GGAGAACTTCAGAAGAGCTTG	CACCAATCCCAACCGTAACAG
MYC	GCAAAAGCTCATTTCTGAAGAGG	TTCATAGGTGATTGCTCAGGAC
RAD51	GTGGCATAAATGCCAACGATG	GCTTTGGCTTCACTAATTCCC
RAD51-AS1	CTCCAGGAATGCGAGTAGGAG	GAGTTCTCACCATCGCAGCTG
SNHG17	GCAGAGAATGGAGAGTGAGGC	CTCTGAGGCAGCTGAAGTCTC
SNHG26	GGAATGTGAAGCTCGTGGTTC	GATGCCTAGAACCTAGGCTG
SNORA71A	GTGATCATAGGCTGCCTGTG	GTGTGAAAGCTTCAGGGTTCG
TTN	CTTCCCTTTCTTGGTAAGCCTC	TGAAGAGATTAAGGTAGAAGCTA
TTN-AS1	GAGCTCCAACTTTAGCAAGTGC	GGTGATGCCTAAGGATGATTCC
ZNF433-AS1	CAGCACTTTGGGAAGCAGAG	GATTACAGGCATGTGCCATGAG
ChIP-qPCR primers (Fig. 4)		
SNHG17	CAGAGTAGCTGCGCGGACAG	CAGTGTCTCGTCCTCTTTGCC
SNHG26	CTTGATCTCCGCAGTGCCCTC	GAGAGGCGCGCCTAGGAAGC
control	GTGGAGCTCAAGACCACGTC	CTGACTGGTCAAGGGTGTGG
shRNA sequences (Fig. 4)		
control	CCGGTCCTAAGGTTAAGTCGCCCTCTCG AGAGGGCGACTTAACCTTAGGATTTTTG	AATTCAAAAATCCTAAGGTTAAGTCGCCCT CTCGAGAGGGCGACTTAACCTTAGGA
ZNF433-AS1_1	CCGGTTGTCTCTCTATTGCCTGCCTCGA GGCAGGCAATAGAGAGACAATTTTTG	AATTCAAAAATTGTCTCTCTATTGCCTGCC TCGAGGCAGGCAATAGAGAGACAA
ZNF433-AS1_2	CCGGTGCAGTTCTGTAGTATCATCTCGA GATGATACTACAGAACTGCATTTTTG	AATTCAAAAATGCAGTTCTGTAGTATCATC TCGAGATGATACTACAGAACTGCA
PCR primers used to check genomic deletions (Fig. S1)		
MALAT1_locus	CCAGAGCAAGCATCAGGACTGTG	CCTTCCTCATGCTACTCTTCTAAGTC

Table S2. PCR-primers used to generate guide library for next generation sequencing.

Primer name	Sequence (5' to 3')
Fwd_1 (i5-barcode)	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTAAGTAGAGATATCTTGTGGAAAGGACGAAACACCG
Fwd_2 (i5-barcode)	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTACACGATCATATCTTGTGGAAAGGACGAAACACCG
Fwd_3 (i5-barcode)	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTAACGCATTATATCTTGTGGAAAGGACGAAACACCG
Fwd_4 (i5-barcode)	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTCATGATCGATATCTTGTGGAAAGGACGAAACACCG
Fwd_5 (i5-barcode)	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTCGTTACCAATATCTTGTGGAAAGGACGAAACACCG
Fwd_6 (i5-barcode)	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTTCCTTGGTATATCTTGTGGAAAGGACGAAACACCG
Rev_1 (i7-barcode)	CAAGCAGAAGACGGCATACGAGATACAGGTATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGCCAAGTTGATAACGGACTAGCCTT
Rev_2 (i7-barcode)	CAAGCAGAAGACGGCATACGAGATAACGCATTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGCCAAGTTGATAACGGACTAGCCTT

Table S3. Oligodeoxynucleotide and gBlock sequences for cloning of paired sgRNAs.

Name	Sequence (5' to 3')
MALAT1_1	AAAGGACGAAACACCGATGACGTCGAGACTATAGTTTGGAGACG CGATCGCGTCTCGTTTGATGGCGCTGCGCTTAAGGTTTTAGAGCT AGAAATAGCA
MALAT1_2	AAAGGACGAAACACCGATGACGTCGAGACTATAGTTTGGAGACG CGATCGCGTCTCGTTTGCGCCCGAGCTGTGCGGTGTTTTAGAGC TAGAAATAGCA
MALAT1_3	AAAGGACGAAACACCGATGACGTCGAGACTATAGTTTGGAGACG CGATCGCGTCTCGTTTGAATTTCCGTGCGGGCCGGTTTTAGAGC TAGAAATAGCA
MALAT1_4	AAAGGACGAAACACCGGCACAGGCGTTAGGGCGGTTTGGAGAC GCGATCGCGTCTCGTTTGAATTTCCGTGCGGGCCGGTTTTAGAG CTAGAAATAGCA
MALAT1_5	AAAGGACGAAACACCGGGACTGCGCAACCGGTGGTTTGGAGAC GCGATCGCGTCTCGTTTGATGGCGCTGCGCTTAAGGTTTTAGAG CTAGAAATAGCA
MALAT1_6	AAAGGACGAAACACCGGGACTGCGCAACCGGTGGTTTGGAGAC GCGATCGCGTCTCGTTTGCGCCCGAGCTGTGCGGTGTTTTAGAG CTAGAAATAGCA
gBlock (scaffold of guide A and mU6 promoter)	CTCATTTACAACCGTCTCCGTTTAAGAGCTATGCTGGAAACAGCA TAGCAAGTTTAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTG GCACCGAGTCGGTGCTTTTTTGGATCCGATCCGACGCCGCCATC TCTAGGCCCGCGCCGGCCCCCTCGCACAGACTTGTGGGAGAAG CTCGGCTACTCCCCTGCCCCGGTTAATTTGCATATAATATTTCCTA GTAACTATAGAGGCTTAATGTGCGATAAAAGACAGATAATCTGTT CTTTTTAATACTAGCTACATTTTACATGATAGGCTTGGATTTCTATA AGAGATACAAATACTAAATTATTATTTTAAAAAACAGCACAAAAGG AAACTCACCCTAACTGTAAAGTAATTGTGTGTTTTGAGACTATAAA TATCCCTTGGAGAAAAGCCTTGTTTGAGAGACGTACAAAGATC