CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes

Luke A. Gilbert[1,2,4,5], Matthew H. Larson[1,2,4,5], Leonardo Morsut[1,2], Zairan Liu[10], Gloria A. Brar[1,2,4,5], Sandra E. Torres [1,2,4,5], Noam Stern-Ginossar[1,2,4,5], Onn Brandman[1,2,4,5], Evan H. Whitehead[1,3,4], Jennifer A. Doudna[2,4-9], Wendell A. Lim[1,2,3,4], Jonathan S. Weissman[1,2,3,4,5]*, Lei S. Qi[1,3,4]

[1]Department of Cellular & Molecular Pharmacology

[2]Howard Hughes Medical Institute

[3]UCSF Center for Systems and Synthetic Biology

University of California, San Francisco, San Francisco, CA 94158, USA

[4]California Institute for Quantitative Biomedical Research, San Francisco, CA 94158, USA

[5]Center for RNA Systems Biology

[6]Department of Molecular & Cellular Biology

[7]Department of Chemistry

[8]Department of Bioengineering

University of California, Berkeley, Berkeley, CA 94720, USA

[9]Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

[10]Peking University, Beijing China

* Correspondence: weissman@cmp.ucsf.edu (J.S.W)

Figure S1, Related to Figure 1C. **CRISPRi mediated suppression of transcription is stable over time.**

(A) A plot displaying the level of GFP knockdown at day 3 and day 6 mediated by dCas9 for 16 sgRNAs targeting SV40-GFP. **(B)** A plot displaying the level of GFP knockdown at day 3 and day 6 mediated by dCas9-KRAB for 16 sgRNAs targeting SV40-GFP.

Figure S2, Related to Figure 1C, 1E. **CRISPRi can repress and activate transcription in human cells.**

(A) dCas9 function is limited by sgRNA expression. A graph showing GFP fluorescence as a function of mCherry fluorescence in GFP+HEK293 stably expressing either dCas9 or dCas9-KRAB and transiently transfected with the indicated sgRNAs. mCherry signal is an indirect measurement of sgRNA expression. Fluorescence was analyzed by flow cytometry 6 days following transfection. The data is binned and error represents the mean ± standard deviation. The data is representative of 3 independent experiments. **(B)** A diagram of the reporter construct used to measure gene activation by GAL4BD-VP64 or dCas9-VP64. sgRNA binding sites map to the Gal4 UAS. The design of the dCas9-VP64 constructs greatly impacts the activity of dCas9-VP64 fusion proteins. HEK293 cells are transfected with a GAL4BD-VP64 positive control or a dCas9-VP64 fusion construct and the indicated sgRNA targeting the Gal4 UAS. GFP fluorescence was measured by flow cytometry 48 hours later. The data are displayed as mean ± standard deviation for 2 independent experiments.

Figure S3, Related to Figure 2. CRISPRi is highly specific in human cells.

(A) A box and whisker plot shows the mean value, 1 and 99 percentiles for two independent RNA sequencing experiments. Outliers are depicted as black dots. The GFP is highlighted in green. RNA was collected 11 or 15 days following viral transduction. The fold change in RPKM

is displayed as a log normalized value for all expressed genes with reads more than 100. **(B)** sgRNAs expressed alone results in few on- or off-target changes in gene expression RNA sequencing data are displayed for GFP+HEK293 cells expressing sgGAL4-1 or sgGFP-NT1. RNA was collected 11 days following viral transduction. **(C)** dCas9-KRAB expressed alone results in few on- or off-target changes in gene expression. RNA sequencing data for GFP+HEK293 cells expressing dCas9-KRAB + sgGAL4-1 or sgGAL4-1. RNA was collected 11 days following viral transduction.

Figure S4, Related to Figure 3. CRISPRi can knock down endogenous genes in human and yeast cells.

(A) A graph displaying the level of CD71 or CXCR4 knockdown in HeLa cells stably expressing dCas9 or dCas9-KRAB and transiently transfected with sgRNAs targeting either CD71 or CXCR4. After 3 days cells are dissociated, stained with a FITC conjugated antibody for either CD71 or CXCR4 and fluorescence is measured by flow cytometry. The coefficient of determination (R^2) is 0.89 and the slope of the linear fit is 0.82 (smaller than 1), suggesting a strong correlation between the degree of knockdown for dCas9 and dCas9-KRAB and dCas9-KRAB on average is more repressive. **(B)** Quantitative PCR measurement of transcript abundance for double gene (CD71 and CXCR4) knockdown in HeLa cells. The y-axis shows transcript abundance values that are normalized to control cells transfected with a non-cognate negative control RNA (sgGAL4-1). RNA abundance is normalized to GAPDH. The data is displayed as mean ± standard deviation for 3 technical replicates. **(C)** A dCas9 fusion protein was constructed with 2 nuclear localization sequences with or without an Mxi1 domain under control of the TDH3 promoter. The sgRNAs are expressed from the *SNR52* promoter.

Supplemental data for sequences used in the study

For human sgRNAs the complete sequence is the N_{20-25} sequence in the table below followed by the sequence below where:

Underlined text is the region complementary to the gene being targeted without the target PAM site.

Italics text is a BstXI and an XhoI restriction site.

Bold text denotes the beginning and end of transcription from the U6 promoter.

We note that the sgRNA starts with the G in bold.

*CCACCTTGTTG***G**<u>NNNNNNNNNNNNNNNNNNNN</u>**GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGGTGCTTTTTTT***CTCGAG***

For yeast sgRNAs the complete sequence is the N_{20-25} sequence in the table below followed by the sequence below where:

The A in bold is the transcription start site.

The underlined C is a splice site generating an sgRNA which starts at N_1 of the targeting sequence.

TCTTTGAAAAGATAATGTATGATTATGCTTTCACTCATATTTATACAGAAACTTGATGTTTTCT
TTCGAGTATATACAAGGTGATTACATGTACGTTTGAAGTACAACTCT**<u>A</u>**GATTTTGTAGTGCC
CTCTTGGGCTAGCGGTAAAGGTGCGCATTTTTTCACACCCTACAATGTTCTGTTCAAAAGAT
TTTGGTCAAACGCTGTAGAAGTGAAAGTTGGTGCGCATGTTTCGGCGTTCGAAACTTCTCC
GCAGTGAAAGATAAATGAT<u>C</u>NNNNNNNNNNNNNNNNNNNNGTTTTAGAGCTAGAAATAGC
AAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGGTGCTT
TTTTTGTTTTTTATGTCT

Table S1. Sequences of sgRNAs used in the study.

sgRNAs for human cells	Sequence
sgGFP.NT1	GAATAGCTCAGAGGCCGAGG
sgGFP.NT2	GACCAGGATGGGCACCACCC
sgGFP.NT3	GGTGGTGCAGATGAACTTCA
sgGFP.NT4	GTGGTCACGAGGGTGGGCCA
sgGFP.NT5	GCACGGGGCCGTCGCCGATG
sgGFP.T1	GGGCGAGGAGCTGTTCACCG

sgGFP.T2	GGCCACAAGTTCAGCGTGTC
sgGFP.T3	GTGAACCGCATCGAGCTGAA
sgSV40.P1	GCATACTTCTGCCTGCTGGGGAGCCTG
sgSV40.P2	GAAAGTCCCCAGGCTCCCCAGC
sgSV40.P3	GCATCTCAATTAGTCAGCAACC
sgSV40.P4	GGATGGGCGGAGTTAGGGGC
sgSV40.P5	GGGCGGAGTTAGGGGC
sgSV40.P6	GTTCCGCCCATTCTCCGCCCCA
sgCXCR4.1	GACTTACACTGATCCCCTCCA
sgCXCR4.2	GCAGGTAGCAAAGTGACGCCGA
sgCXCR4.3	GAACCAGCGGTTACCATGGA
sgCXCR4.4	GCCAACAAACTGAAGTTTCTGGCCG
sgCXCR4.5	GAAGTTTCTGGCCGCGGC
sgCXCR4.6	GCGCATGCGCCGCTGGGGCG
sgCXCR4.7	GGAGGGAGAAGGCGGGGTGG
sgCXCR4.8	GCGGTGGCTACTGGAGCACTC
sgCXCR4.9	GCCGCTTCTGCCCGCTCGGAGA
sgCXCR4.10	GCGGCGCATGCGCCGCGCT
sgTRFC.1	GATATCCCGACGCTCTGAGGGGA
sgTRFC.2	GGACGCGCTAGTGTGAGTGC
sgTRFC.3	GGGATATCGGGTGGCGGCTC
sgTRFC.4	GGGCTGTGCGTCACTTCCTG
sgTRFC.5	GTGTACCTGCAGCCGCTCG
sgTRFC.6	GTTTATAGCCTGGCCCCGCCCC
sgTRFC.7	GGGCGGGGATGCGCGCGCAGC
sgTRFC.8	GCGGGCGGGGCCCTATGCGGA
sgTRFC.9	GAGGTGCTCTGACAGATCGCC
sgTRFC.10	GTGTCCTCCCTTCATCCTGC

sgRNAs for yeast cells	Sequence
sgTEF	TTGATATTTAAGTTAATAAACGG
sgTET	TATCAGTGATAGAGAAAAGTCCC