Maximizing CRISPRi efficacy and accessibility with dual-sgRNA libraries and optimal effectors

Joseph M. Replogle[1-4,†], Jessica L. Bonnar[2-4,†], Angela N. Pogson[2-4,†,‡], Christina R. Liem[2,‡], Nolan K. Maier[5], Yufang Ding[5], Baylee J. Russell[5], Xingren Wang[5], Kun Leng[1,6], Alina Guna[2,4], Thomas M. Norman[2,‡], Ryan A. Pak[2,‡], Daniel M. Ramos[7,8], Michael E. Ward[9], Luke A. Gilbert[2,10,11], Martin Kampmann[6,12], Jonathan S. Weissman[2-4,13,14,*], Marco Jost[2,5,*]

[1]Medical Scientist Training Program, University of California, San Francisco, San Francisco, CA 94158, USA

[2]Department of Cellular and Molecular Pharmacology, University of California, San Francisco, San Francisco, CA 94158, USA

[3]Howard Hughes Medical Institute, Massachusetts Institute of Technology, Cambridge, MA 02142, USA

[4]Whitehead Institute for Biomedical Research, Cambridge, MA 02142, USA

[5]Department of Microbiology, Harvard Medical School, Boston, MA 02115, USA

[6]Institute for Neurodegenerative Disease, University of California, San Francisco, San Francisco, CA 94158, USA

[7]Center for Alzheimer's Disease and Related Dementias, National Institutes of Health, Bethesda, MD 20892, USA

[8]National Institute on Aging, National Institutes of Health, Bethesda, MD 20892, USA

[9]National Institute of Neurological Disorders and Stroke, NIH, Bethesda, MD 20892, USA

[10]Department of Urology and Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, CA 94158, USA

[11]Arc Institute, Palo Alto, CA, 94304, USA

[12]Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, CA 94158, USA

[13]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02142, USA

[14]David H. Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02142, USA

[†]Co-first authors

*Address correspondence to: J.S.W. (weissman@mit.edu); M.J. (marco_jost@hms.harvard.edu)

[‡]Present Addresses:

Angela N. Pogson: Department of Developmental Biology, Stanford University School of Medicine, Stanford, CA 94305, USA

Christina R. Liem: Division of Biological Sciences, Section of Cell and Developmental Biology, University of California San Diego, La Jolla, CA 92093, USA

Ryan A. Pak: Department of Neuroscience, Scripps Research, La Jolla, CA 92037, USA

Thomas M. Norman: Computational & Systems Biology Program, Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA

Abstract

CRISPR interference (CRISPRi) enables programmable, reversible, and titratable repression of gene expression (knockdown) in mammalian cells. Initial CRISPRi-mediated genetic screens have showcased the potential to address basic questions in cell biology, genetics, and biotechnology, but wider deployment of CRISPRi screening has been constrained by the large size of single guide RNA (sgRNA) libraries and challenges in generating cell models with consistent CRISPRi-mediated knockdown. Here, we present next-generation CRISPRi sgRNA libraries and effector expression constructs that enable strong and consistent knockdown across mammalian cell models. First, we combine empirical sgRNA selection with a dual-sgRNA library design to generate an ultra-compact (1-3 elements per gene), highly active CRISPRi sgRNA library. Next, we rigorously compare CRISPRi effectors to show that the recently published Zim3-dCas9 provides the best balance between strong on-target knockdown and minimal nonspecific effects on cell growth or the transcriptome. Finally, we engineer a suite of cell lines with stable expression of Zim3-dCas9 and robust on-target knockdown. Our results and publicly available reagents establish best practices for CRISPRi genetic screening.

Introduction

CRISPR interference (CRISPRi) enables programmable repression of gene expression with broad applications in genome engineering, genetic screening, and cell biology (Doench, 2018). In mammalian cells, CRISPRi requires two components: (i) an effector protein of catalytically dead Cas9 (dCas9) fused to one or more transcription repressor domains, which recruits endogenous epigenetic modulators to the genome, and (ii) a single guide RNA (sgRNA), which directs the effector protein to target DNA (Gilbert et al., 2013). When the sgRNA is targeted to a gene promoter, CRISPRi leads to repressive epigenome editing and knockdown of the gene (Gilbert et al., 2014; Horlbeck et al., 2016a, 2016b).

Several features distinguish CRISPRi from Cas9 nuclease-mediated DNA cutting, the major alternative CRISPR/Cas-based approach for loss-of-function genetic studies: i) Unlike Cas9, CRISPRi does not rely on introduction of double-stranded DNA breaks and therefore does not cause genomic rearrangements (Kosicki et al., 2018) and DNA damage-associated toxicity (Meyers et al., 2017), which may be especially limiting in primary and stem cells (Bowden et al., 2020; Haapaniemi et al., 2018; Ihry et al., 2018). ii) CRISPRi tends to confer more homogeneous loss of gene function compared to Cas9, which often generates subpopulations of cells bearing active in-frame indels (Smits et al., 2019). iii) CRISPRi is reversible and thus affords temporal control over gene expression levels (Gilbert et al., 2014; Mandegar et al., 2016). iv) CRISPRi enables titration of gene expression, which for example allows for partial depletion of genes essential for cell growth and interrogation of the resulting phenotypes (Bosch et al., 2021; Hawkins et al., 2020; Jost et al., 2020). v) In turn, one can directly measure the extent of on-target knockdown as well as the corresponding responses in individual cells, for example using single-cell RNA-seq (Perturb-seq), allowing for evaluation of the extent and potential biological significance of cell-to-cell heterogeneity. vi) CRISPRi enables loss-of-function studies for non-coding RNAs, which are difficult to inactivate or repress through CRISPR cutting and the introduction of indels as they are insensitive to frame-shifting mutations (Liu et al., 2017).

Like other CRISPR approaches, CRISPRi has been paired with large-scale sgRNA libraries to conduct systematic genetic screens. Such screens have been deployed to identify essential protein-coding and non-coding genes (Gilbert et al., 2014; Haswell et al., 2021; Horlbeck et al., 2016a; Liu et al., 2017; Raffeiner et al., 2020), to map the targets of regulatory elements (Fulco et al., 2019, 2016; Gasperini et al., 2019; Kearns et al., 2015; Klann et al., 2017;

Thakore et al., 2015), to identify regulators of cellular signaling and metabolism (Coukos et al., 2021; Liang et al., 2020; Luteijn et al., 2019; Semesta et al., 2020), to uncover stress response pathways in stem cell-derived neurons (Tian et al., 2021, 2019), to uncover regulators of disease-associated states in microglia and astrocytes (Dräger et al., 2022; Leng et al., 2022), to decode regulators of cytokine production in primary human T-cells (Schmidt et al., 2022), to define mechanisms of action of bioactive small molecules (Jost et al., 2017; Morgens et al., 2019; Sage et al., 2017), to identify synthetic-lethal genetic interactions in cancer cells (Du et al., 2017; Horlbeck et al., 2018), and to identify genetic determinants of complex transcriptional responses using RNA-seq readouts (Perturb-seq) (Adamson et al., 2016; Replogle et al., 2022, 2020; Tian et al., 2021, 2019), among others.

Despite these successes, two technical factors have limited wider adoption of CRISPRi. First, CRISPRi screening is constrained by the large size of sgRNA libraries. Previous machine learning efforts yielded guide design rules which substantially increased the activity of sgRNA libraries (Horlbeck et al., 2016a; Sanson et al., 2018). Nonetheless, commonly used libraries (e.g., Dolcetto, CRISPRi v2) target each gene with three or more sgRNAs to decrease false negative results in screens. The development of a more compact, highly active sgRNA library would enable CRISPRi screens in new cell types and for more complex phenotypes, especially when cost, time, and/or cell numbers are limiting. Second, there is no clear consensus guiding the use of the different reported CRISPRi effector proteins, complicating the generation of CRISPRi cell models (Alerasool et al., 2020; Carleton et al., 2017; Gilbert et al., 2014; Yeo et al., 2018).

Here we present a suite of tools to enable high-quality CRISPRi genetic screening in a broad range of cell models. Based on empirical data aggregated from 126 screens, we design and validate an ultra-compact, highly active CRISPRi library in which each gene is targeted by a single library element encoding a dual sgRNA cassette. Next, we comprehensively compare published CRISPRi effector proteins based on their on-target efficacy and non-specific effects on transcription and cell proliferation. We find that the recently published Zim3-dCas9 provides the best balance between strong on-target knockdown and minimal nonspecific effects. Finally, we generate K562, RPE1, Jurkat, HT29, HuTu-80, and HepG2 cell lines engineered to stably express Zim3-dCas9 and demonstrate robust on-target knockdown across these cell lines. Our results and reagents establish best practices for CRISPRi genetic screening.

Results

Comparison of single and dual sgRNA CRISPRi libraries for genetic screening

A critical mediator of the potential applications of CRISPRi screening is the on-target efficacy and size of the sgRNA library. In recent work, we found that targeting individual genes with dual-sgRNA constructs substantially improved CRISPRi-mediated gene knockdown (Replogle et al., 2020). Building on this result, we asked whether a dual-sgRNA strategy could be used to generate an ultra-compact, genome-wide CRISPRi library.

To assess the potential utility of dual sgRNA libraries in systematic genetic screens, we began by cloning two pilot libraries for comparison: (i) one targeting each human gene with two distinct sgRNAs expressed from a tandem sgRNA cassette (dual sgRNA) and (ii) one targeting each human gene by only the single best sgRNA (see *Methods*; Table S1; Supplementary Note 1). We also optimized a protocol to amplify and sequence dual-sgRNA cassettes from lentivirally integrated genomic DNA (see *Methods*; Supplementary Note 2). Next, we compared the performance of our single- and dual-sgRNA libraries in a genome-wide growth screen (Figure 1A). We transduced K562 cells stably expressing dCas9-KRAB(Kox1) with our libraries, used puromycin to select for cells with lentiviral integration, and harvested cells at day 8 (T_0) and day 20 (T_{final}) post-transduction. We amplified sgRNA cassettes from extracted genomic DNA, sequenced to quantify sgRNA abundance in the two populations, and calculated growth phenotypes for each library element by comparing changes in abundance between T_0 and T_{final} (Figure 1C). The growth phenotypes produced by the single- and dual-sgRNA libraries were well-correlated with previously published CRISPRi growth screens using 5 sgRNAs per gene (single sgRNA $r = 0.81$; dual sgRNA $r = 0.82$; Figure S1A-C) and produced near-perfect recall of essential genes (Figure S1D) (AUC>0.98 for both single- and dual-sgRNA libraries). Yet, for essential genes previously identified by the Cancer Dependency Map (DepMap) (Behan et al., 2019; Tsherniak et al., 2017), the dual-sgRNA library produced significantly stronger growth phenotypes (median 28% decrease in the growth rate [γ]) than the single-sgRNA library (n = 1,847 genes; single-sgRNA mean $\gamma = -0.19$; dual-sgRNA mean $\gamma = -0.24$; Mann Whitney p-value = $8 \cdot 10^{-13}$; Figure 1C, 1D), suggesting that the dual-sgRNA library confers stronger depletion of target genes.

A well-recognized challenge for the use of dual-sgRNA libraries is that the lentiviral reverse transcriptase can undergo template switching between the two copies of the lentiviral genome packaged into each capsid (Adamson et al., 2018, 2016; Feldman et al., 2018; Hill et al., 2018; Horlbeck et al., 2018; Xie et al., 2018). These two copies generally bear two different sgRNA pairs in a pooled dual-sgRNA library, such that template switching can produce a recombined element with sgRNAs targeting different genes. Our sequencing strategy allowed us to directly identify such recombined elements (Figure 1B), which occurred with a frequency of 29.4% across replicates, consistent with prior reports (Horlbeck et al., 2018; Replogle et al., 2020). In our downstream analyses, we exclude all recombined elements such that they do not impact phenotypes, although in principle these recombined elements could be used to assess independent effects of the two sgRNAs targeting each gene.

Design and validation of ultra-compact, dual sgRNA CRISPRi libraries

Having validated the performance of dual-sgRNA libraries in a systematic genetic screen, we sought to optimize the activity and utility of dual-sgRNA CRISPRi libraries (Figure 1E). To optimize sgRNA selection for each gene, we aggregated empirical sgRNA activity data from 126 CRISPRi genetic screens (Table S2) and implemented a three-tiered selection system. First, for genes that are essential in K562 cells, we ranked sgRNAs by growth phenotype. Second, for genes that produced a significant phenotype in one of our previous CRISPRi screens, we ranked sgRNAs by relative *z*-scored phenotype averaged across screens in which the target gene was identified as a hit. Finally, for genes without any empirical effect in a prior screen, we ranked sgRNAs according to predicted activities from the hCRISPRi v2.1 algorithm (see *Methods*) (Horlbeck et al., 2016a). To allow users to select the library size suitable to their application, we cloned sublibraries of the best single element (guide ranked 1+2; referred to as hCRISPRi_dual_1_2), the second best element (guides ranked 3+4; referred to as hCRISPRi_dual_3_4), or the third best element (guides ranked 5+6; referred to as hCRISPRi_dual_5_6) (Table S3).

Further examination of the phenotypes from our screens revealed that a small number of elements produced discordant effects between screens, which may arise from bottlenecking or amplification bias (Figure S1A-C). For libraries with multiple elements targeting each gene, such discordant effects can often be mitigated by comparing phenotypes across elements, but this

option is not available with single-element libraries. In previously reported CRISPR cutting libraries, incorporation of barcodes into the sgRNA cassette enabled marking and tracing populations of cells derived from individual lentiviral integrations, which allowed for detection of bottlenecking events and amplification bias and thereby improved screen sensitivity and robustness (Michlits et al., 2017; Zhu et al., 2019). Building on these results, we incorporated a set of 215 8-nucleotide barcodes, which we term Integration Barcodes (IBCs), in the tandem sgRNA cassette of our final hCRISPRi_dual_1_2, hCRISPRi_dual_3_4, and hCRISPRi_dual_5_6 libraries (Methods, Table S4). We then optimized a sequencing strategy for simultaneously sequencing the two sgRNAs, the IBC, and a sample index on Illumina sequencers (Figure 1B, Supplementary Note 2).

Finally, we sought to test our optimized dual-sgRNA library side-by-side with the recently reported Dolcetto CRISPRi library, which was designed with a differently prioritized sgRNA selection algorithm and uses single-sgRNAs (Sanson et al., 2018). We used direct capture Perturb-seq (Replogle et al., 2020), pooled CRISPR screens with single-cell RNA-seq readout, to measure the on-target knockdown mediated by the top three elements in our dual sgRNA library (guides 1+2, guides 3+4, or guides 5+6) or the three Dolcetto Set A sgRNAs for 128 randomly selected genes that are expressed in K562 cells (Table S5). Our dual-sgRNA library significantly outperformed the Dolcetto library, as quantified by the average knockdown (dual-sgRNA median knockdown 82%; Dolcetto median knockdown 65%; Mann Whitney *p*-value = $2.4 \cdot 10^{-7}$) as well as the strongest knockdown per gene (dual-sgRNA median knockdown 90%; Dolcetto median knockdown 87%; Mann Whitney *p*-value = $2 \cdot 10^{-4}$; Figure 1F). Indeed, the top-ranked element of our dual-sgRNA library (guides 1+2) alone produced comparable knockdown to the best of all three Dolcetto sgRNAs (dual sgRNA element 1+2 median knockdown 86%; best Dolcetto sgRNA median knockdown 87%; Mann Whitney *p*-value = 0.43) (Figure 1G). We note that an analogous dual-sgRNA approach may improve knockdown for the Dolcetto library. Nonetheless, from these data we conclude that our dual-sgRNA library improves on-target knockdown compared to gold-standard CRISPRi libraries.

Design of CRISPRi effector expression constructs for systematic comparisons

We next sought to compare different CRISPRi effectors, with the goal of identifying an effector with strong activity and minimal non-specific effects on global transcription and cell growth. We

selected four repressor domains that had been described to mediate strong and specific knockdown in dCas9 fusions: 1) the KRAB domain from KOX1 (*ZNF10*), which was used in the original conception of CRISPRi (Gilbert et al., 2013); 2) the KRAB domain from ZIM3, which was recently reported to mediate stronger knockdown than KRAB(KOX1) (Alerasool et al., 2020); 3) the SIN3A interacting domain of MAD1 (SID4x) (Carleton et al., 2017); and 4) the transcription repression domain of MeCP2 (Yeo et al., 2018).

To enable direct comparisons, we embedded each effector in a standardized lentiviral expression construct (Figure 2A, Table S6). Briefly, in this construct, expression is driven by a spleen focus-forming virus (SFFV) promoter, with an upstream ubiquitous chromatin opening element (UCOE) to minimize silencing, internal nuclear localization signals (NLSs) and an internal HA tag, a GFP marker linked at the C-terminus via a P2A ribosomal skipping sequence to allow for stable cell line generation by FACS, and a Woodchuck Hepatitis Virus post-transcriptional regulatory element (WPRE) in the 3′ UTR to increase mRNA stability. Where necessary, we included linker sequences derived from the functionally innocuous XTEN domain (Schellenberger et al., 2009) to minimize proteolytic cleavage between dCas9 and fused repressor domains. We attempted to maximize the activity for each repressor domain based on our previous data and data in the literature, although we note that our evaluation is not exhaustive. The final designs of the four effector expression constructs are depicted in Figure S2, with further rationale in the Methods section. We then compared the four effectors with regards to two key criteria: on-target activity and absence of non-specific effects on cell viability and gene expression.

CRISPRi effectors containing SID or MeCP2 domains have non-specific effects on cell viability and gene expression

The repressor domain of each CRISPRi effector is a transcription factor domain whose overexpression has the potential to cause non-specific (i.e., not mediated by dCas9 targeting) and potentially detrimental effects on transcription or cell proliferation. To test for effects on proliferation, we generated K562 cell lines stably expressing each effector by lentiviral transduction followed by FACS (Figure 2B) and then quantified the effect of each effector on cell proliferation using an internally normalized competitive growth assay. We mixed cells bearing each effector ~1:1 with cells expressing mCherry and quantified growth defects of

effector-expressing cells by measuring the ratio of mCherry-negative to mCherry-positive cells over time by flow cytometry. We used mCherry-expressing cells as a reference population instead of parental, GFP-negative cells because some of the effector-expressing cells convert to GFP-negative over time due to silencing, which is difficult to separate from true dropout of effector-expressing cells due to growth defects. Over 19 days, cells expressing dCas9 only, dCas9-Kox1, or Zim3-dCas9 proliferated at the same rate as cells expressing GFP only or non-transduced control cells, suggesting that expression of these effectors is not toxic over this time span (Figure 2C). By contrast, cells expressing SID-dCas9-Kox1 had a strong growth defect (~6% per day), and cells expressing dCas9-Kox1-MeCP2 had a mild growth defect (~1% per day, Figure 2C).

To assess non-specific effects of effectors on transcription, we performed global transcriptome profiling of K562 cells stably transduced with these effectors by RNA-seq (Figure 2D,E). Consistent with the growth assay, cells expressing SID-dCas9-Kox1 had globally perturbed transcription, with 4282 genes differentially expressed compared to control cells expressing GFP only at $p<0.05$ (Figure 2E). Indeed, these samples clustered separately from every other control and effector-expressing sample (Figure 2D). In addition, 53 genes were differentially expressed in cells with dCas9-Kox1-MeCP2, suggesting that constitutive expression of this effector also leads to minor non-specific effects on transcription (Figure 2E). No more than 3 genes were detected to be differentially expressed in cells expressing any of the other effectors, suggesting that these effectors do not non-specifically perturb transcription (Figure 2E). Together, these results suggest that (over)expression of SID-dCas9-Kox1 is toxic and globally perturbs transcription at least in K562 cells. We therefore excluded this effector from further analysis.

Zim3-dCas9 and dCas9-Kox1-MeCP2 mediate strongest knockdown

We next sought to measure the efficacy of each effector in knocking down targeted genes with two complementary approaches: (i) measurement of growth phenotypes resulting from knockdown of essential genes, i.e. genes required for the growth or survival of dividing human cells, and (ii) direct measurement of knockdown of cell surface proteins (Figure 3A, Table S7). In both assays, we used single-sgRNA expression cassettes, which allowed us to use previously validated strong and intermediate-activity sgRNAs (Jost et al., 2020). We included intermediate-

activity sgRNAs for two reasons: First, activity differences between effectors are more apparent when knockdown is not saturated. Second, as it can be challenging to identify sgRNAs with high activity across genes and cell types, effectors that mediate strong knockdown even with imperfect sgRNAs could reduce false negative rates in genetic screens.

We measured growth phenotypes resulting from knockdown of essential genes using internally normalized competitive growth assays. We transduced K562 cell lines stably expressing each CRISPRi effector with vectors simultaneously expressing an sgRNA and a fluorescent marker (mCherry) at a low multiplicity of infection (0.2-0.5). We then monitored the ratio of sgRNA-expressing cells (mCherry+) and unperturbed cells (mCherry-) by flow cytometry, with the expectation that cells with an essential gene-targeting sgRNA would deplete at a rate proportional to CRISPRi activity. We targeted three genes, alanyl-tRNA synthetase (*AARS*), the mitochondrial inner membrane import factor DNAJC19, and subunit D of RNA polymerase I and III (*POLR1D*), with three different sgRNAs each. For all sgRNAs tested, sgRNA-expressing cells depleted at the fastest rate with Zim3-dCas9 and at the second-fastest rate with either dCas9-Kox1 or dCas9-Kox1-MeCP2 (Figures 3B, S3A).

Next, to directly measure depletion of targeted proteins, we knocked down the non-essential cell surface proteins CD55 (complement decay-accelerating factor), CD81 (TAPA-1/ TSPAN28), and CD151 (TSPAN24) and measured staining intensity with fluorescently labeled antibodies by flow cytometry as a proxy for protein levels. We transduced K562 lines stably expressing the different CRISPRi effectors with vectors expressing either targeting or non-targeting sgRNAs at a low multiplicity of infection (0.2-1). Six days after transduction, we stained cells with fluorescently labeled antibodies against the different cell surface proteins and assessed knockdown by comparing the median antibody staining intensity in cells expressing a targeting sgRNA and cells expressing a non-targeting control sgRNA. With strong sgRNAs, Zim3-dCas9, dCas9-Kox1, and dCas9-Kox1-MeCP2 all mediated strong depletion of each cell surface protein (>96.8% median depletion for all effectors and sgRNAs). With weak sgRNAs, dCas9-Kox1-MeCP2 mediated the strongest knockdown closely followed by Zim3-dCas9, whereas dCas9-Kox1 mediated weaker knockdown (Figures 3C, 3D, S3B-D).

Importantly, flow cytometry reports on expression at the single-cell level, allowing us to assess cell-to-cell heterogeneity in knockdown, which is missed when quantifying median expression. As a proxy for heterogeneity, we calculated the fraction of cells without evidence of

knockdown despite the use of a strong sgRNA (Figure 3E). For Zim3-dCas9, knockdown was largely homogeneous, with only ~5% of cells without detectable knockdown (Figures 3D, 3E, S3B-D), perhaps due to the presence of some senescent cells in the population in which lack of cell division limits protein dilution. By contrast, for dCas9-Kox1-MeCP2, 15-20% of cells did not achieve knockdown (Figures 3D, 3E, S3B-D). This result may be explained by the toxicity of the effector protein leading to selection against effector expression (Figure 2C) or may be indicative of an intrinsic property of MeCP2 activity. The observed heterogeneity in MeCP2 knockdown may help explain why dCas9-Kox1-MeCP2 appears to mediate the strongest median knockdown while Zim3-dCas9 leads to faster dropout of sgRNA-expressing cells in the essential gene growth assay; in the growth assay, heterogeneity would lead to worse performance due to selection against strong knockdown. In sum, these results suggest that the Zim3-dCas9 effector confers strong knockdown that is homogeneous across a cell population.

A versatile collection of Zim3-dCas9 constructs for robust knockdown across cell types

To assess the general utility of the Zim3-dCas9 effector, we measured knockdown efficacy in different cell types. For each cell type, we constructed cell lines stably expressing Zim3-dCas9 (see *Methods*; Supplementary Note 4) and measured knockdown of cell surface proteins by flow cytometry. In both RPE1 (retinal pigment epithelium) and Jurkat (acute T-cell leukemia) cells, cells expressing Zim3-dCas9 had stronger knockdown than previously reported cell lines expressing dCas9-Kox1 (Figure 4A) (Horlbeck et al., 2018; Jost et al., 2017). Zim3-dCas9 also conferred strong and homogeneous knockdown in HepG2 (hepatocellular carcinoma), HT29 (colorectal adenocarcinoma), and HuTu-80 (duodenal adenocarcinoma) cells (Figure 4B, S4).

To further maximize utility of the Zim3-dCas9 effector, we generated a panel of constructs for expression of Zim3-dCas9 from the SFFV or EF1α promoters linked to BFP, GFP, or mCherry (Table S6). We also generated backbones to express effectors from additional promoters (CMV, EFS) and with different types of C-terminal fluorescent protein linkages (P2A, IRES, direct fusion) (Table S6). In addition, as the bright fluorescence from the fluorescent proteins may be undesirable in some settings, we generated a construct in which expression of Zim3-dCas9 is linked to a hygromycin resistance marker [Zim3-dCas9 (Hygro)]. K562 cells stably transduced with Zim3-dCas9 (Hygro) and selected with hygromycin for 4 weeks had strong and homogeneous knockdown that was indistinguishable from knockdown in a cell line

generated by FACS (Figure 4C). Finally, we generated constructs in which the fluorescent protein is flanked by LoxP sites, such that the fluorescent protein can be removed by transient delivery of Cre once a stable cell line has been generated (Table S6). All of our constructs are available via Addgene. Our collection of Zim3-dCas9 expression constructs and streamlined protocols enables robust CRISPRi across a broad range of cell models.

Discussion

High-quality genetic screening approaches are catalysts for basic research and drug development. Among the available approaches, CRISPRi has several appealing features including independence of double-stranded DNA breaks, homogeneity and reversibility of perturbations, accessibility of partial loss-of-function phenotypes, and compatibility with direct measurements of target gene expression levels in both bulk populations and single cells. CRISPRi screens have indeed propelled biological discovery in several contexts, but broader deployment has been limited by difficulties in generating CRISPRi cell models and limited knockdown efficacy for a subset of genes. Here, we present a suite of tools and accompanying protocols to address these limitations and improve the efficacy and accessibility of CRISPRi.

Our ultra-compact, dual-sgRNA CRISPRi library confers stronger knockdown and growth phenotypes than previously reported libraries and thus should minimize false-negative rates in screens. Nonetheless, this library also has drawbacks. First, some library elements undergo intermolecular recombination during lentiviral transduction. We can detect and computationally remove such recombination events, such that they do not corrupt the resulting data. As a consequence, recombination primarily decreases effective library coverage, and in return cell numbers need to be increased by ~20-30% to ensure coverage. In the future, recombination may be further mitigated using decoy vectors, different promoters, and alternatively processed guides (Adamson et al., 2016; Dong et al., 2017; Feldman et al., 2018; Knapp et al., 2019). Second and perhaps more importantly, screens will be inherently noisier and sensitive to off-target effects with only a single element per gene, such that in standard cell lines in which cell numbers are not a concern, existing single-sgRNA libraries may remain the approach of choice. Inclusion of the 3-4 and 5-6 element sublibraries in our dual-sgRNA library can mitigate this noise at the expense of some of the compactness. In cases in which cell numbers are limited by the model, time, or cost, however, the compactness of our dual-sgRNA

library will be transformative. Examples include screens in primary or stem-cell derived models or *in vivo* as well as screens with high-content readout such as Perturb-seq (Bock et al., 2022; Przybyla and Gilbert, 2021). Additionally, this dual-sgRNA strategy may provide similar benefits for other CRISPR modalities such as CRISPR-mediated overexpression (CRISPR activation, CRISPRa), as also described by others (Yin et al., 2022), and we have designed a dual-sgRNA CRISPRa library for this purpose (Table S8). Finally, the improved knockdown afforded by the dual-sgRNA approach will also be beneficial in arrayed experiments, in which recombination is not a concern, and we have included a protocol for cloning dual-sgRNA libraries in array (Supplementary Note 3). In sum, our dual-sgRNA libraries improve CRISPRi knockdown and complement existing libraries by broadening the scope of models in which CRISPRi screens are feasible.

In the realm of CRISPRi effectors, our work points to a clear consensus: Zim3-dCas9 is the effector of choice, as it appears equal or superior to other effectors in every test we performed and had no downsides. We had previously measured by Perturb-seq that Zim3-dCas9 afforded median mRNA knockdown of 91.6% across 2,285 genes in RPE1 cells (Replogle et al., 2022), and here we further found that Zim3-dCas9 mediates robust knockdown across a range of cell types. Our work highlights the importance of using multiple assays to assess effector function including single-cell assays to assess cell-to-cell heterogeneity, of directly measuring knockdown instead of relying on proxies such as growth phenotypes that conflate multiple factors, and of evaluating effectors in stably transduced cells rather than in transiently transfected cells to evaluate longer-term consequences for cell viability. To facilitate implementation of CRISPRi in additional cell models, we created a suite of effector expression constructs with different combinations of promoters and markers (Table S6) as well as a cell line generation protocol (Supplementary Note 4).

Nonetheless, there is more progress to be made in evaluating effectors and generating robust CRISPRi models. First, our comparison of the effectors was not exhaustive. For example, although we expressed all effectors from the same context, we did not control for potential differences in expression levels or nuclear localization across effectors. The Zim3-dCas9 expression constructs appear optimal as they are, but activities of other effectors may be boosted by optimizing these factors. Second, repression of gene expression is generally mediated through recruitment of endogenous cofactors; for KRAB domains such as those from Zim3 and Kox1,

this endogenous cofactor is TRIM28 (Ecco et al., 2017). TRIM28 expression varies by cell type, and efficacy of Zim3 and Kox1 may be limited in cell types with low TRIM28 expression. In such cell types, the MeCP2 effector may be a suitable alternative, but the selection against effector-expressing cells may increase false positive and false negative rates. Third, we did not measure if the effectors differed in propensity for sgRNA-dependent off-target effects. Previous work on dCas9-Kox1 had documented that well-designed sgRNAs have minimal off-target effects (Gilbert et al., 2013). The main source of off-target effects of CRISPRi are at bidirectional promoters, which likely is an inevitable consequence of the mechanism of CRISPRi. We note that the stronger activity of Zim3-based effectors may amplify such effects. For now, such off-target effects can be readily predicted and measured, for example by Perturb-seq (Replogle et al., 2022). Perhaps future efforts will identify strategies to limit knockdown of neighboring genes. Finally, in some cell types effector expression is silenced over time, leading to loss of CRISPRi activity. We described some strategies to counteract such silencing (Supplementary Note 4), but further protection against silencing remains as an area for improvement. In any case, the assays we describe may be used to test additional effectors in a streamlined and standardized fashion, with the goal of making CRISPRi universally available across cell models.

Altogether, our resources and best practices will guide both current implementations and future developments of CRISPRi. All our protocols, constructs, cell lines, and libraries are available as resources to the community.

Acknowledgements

We thank B. Adamson, J. Nuñez, J. Hussmann, M. Horlbeck, E. Chow, S. Gupta, and F. Urnov for helpful discussions. We thank M. de Vera, S. Sinha, C. Muresan, and J. Kanter for lab and administrative support. We thank I. Aifantis (NYU) for donating a vector containing SID4x and G. Ow and E. Collisson (UCSF) for donating the mCherry-marked sgRNA expression vector. We thank the UCSF Center for Advanced Technology for sequencing, the Whitehead Institute Flow Cytometry Core and the Harvard Immunology Flow Cytometry Core Facility for access to FACS machines and flow cytometers, and the Whitehead Institute Genome Technology Core for performing bulk RNA-seq library preparation.

Funding

This work was supported in part by the National Institutes of Health (grants R00GM130964 to MJ and RM1HG009490-01 to JSW), Springer Nature Global Grant for Gut Health 1772808 (MJ), a Charles H. Hood Foundation Child Health Research Award (MJ), DARPA HR0011-19-2-0007 (JSW), the Ludwig Center at MIT (JSW), and the Chan Zuckerberg Initiative (JSW). JMR was supported by NIH F31 Ruth L. Kirschstein National Research Service Award NS115380. KL was supported by NIH F30 Ruth L. Kirschstein National Research Service Award AG066418. BJR was supported by Harvard Bacteriology PhD Training Program T32 AI132120. AG was supported by Human Frontier Science Program grant 2019L/LT000858. MK was supported by the Chan Zuckerberg Initiative Ben Barres Early Career Acceleration Award. JSW is a Howard Hughes Medical Institute Investigator.

Competing Interests

JMR consults for Maze Therapeutics and Waypoint Bio. TMN consults for Maze Therapeutics. MK serves on the Scientific Advisory Boards of Engine Biosciences, Casma Therapeutics, Cajal Neuroscience, and Alector, and is an advisor to Modulo Bio and Recursion Therapeutics. MJ consults for Maze Therapeutics and Gate Bioscience. LAG declares outside interest in Chroma Medicine. JSW declares outside interest in 5 AM Venture, Amgen, Chroma Medicine, KSQ Therapeutics, Maze Therapeutics, Tenaya Therapeutics, Tessera Therapeutics, and Third Rock Ventures. The Regents of the University of California with TMN, MJ, LAG, and JSW as inventors have filed patent applications related to CRISPRi/a screening and Perturb-seq. LAG, MK, and JSW are inventors on US Patent 11,254,933 related to CRISPRi/a screening.

Data and Code Availability

Python scripts for alignment of sequencing data from dual-sgRNA screens with and without IBCs is available here: https://github.com/josephreplogle/CRISPRi-dual-sgRNA-screens. Sequencing data are available on NCBI GEO under accession number GSE205310 (Perturb-seq) and GSE205147 (bulk RNA-seq).

Materials and Methods

Cell line generation and maintenance

K562 cells were grown in RPMI 1640 medium with 25 mM HEPES, 2 mM L-glutamine, 2 g/L $NaHCO_3$ (Gibco) supplemented with 10% (v/v) standard fetal bovine serum (FBS, VWR), 100 units/mL penicillin, 100 µg/mL streptomycin, and 2 mM L-glutamine (Gibco). hTERT-immortalized RPE1 cells (ATCC, CRL-4000) were grown in DMEM:F12 (Gibco) supplemented with 10% (v/v) standard FBS (VWR), 0.01 mg/mL hygromycin B, 100 units/mL penicillin, and 100 µg/mL streptomycin. Jurkat cells (ATCC, Clone E6-1 TIB-152) were grown in RPMI 1640 medium with 25 mM HEPES, 2 mM L-glutamine, 2 g/L $NaHCO_3$ (Gibco) supplemented with 10% (v/v) standard FBS (VWR), 100 units/mL penicillin, 100 µg/mL streptomycin, and 2 mM L-glutamine (Gibco). HepG2 and HuTu-80 cells (both ATCC) were grown in Eagle's Minimum Essential Medium with 1.5 g/L NaHCO3, 110 mg/L sodium pyruvate, 292 mg/L L-glutamine (Corning) supplemented with 10% (v/v) standard FBS (R&D Systems), 100 units/mL penicillin, and 100 µg/mL streptomycin (Gibco). HT29 cells (ATCC) were grown in DMEM with 25 mM D-glucose, 3.7 g/L $NaHCO_3$, 4 mM L-glutamine (Gibco) supplemented with 10% (v/v) standard FBS (R&D Systems), 100 units/mL penicillin, and 100 ug/mL streptomycin (Gibco). HEK293T cells were grown in Dulbecco's modified eagle medium (DMEM) with 25 mM D-glucose, 3.7 g/L $NaHCO_3$, 4 mM L-glutamine (Gibco) and supplemented with 10% (v/v) standard FBS (VWR or R&D Systems), 100 units/mL penicillin, 100 µg/mL streptomycin, and 2 mM L-glutamine (Gibco). K562 (chronic myelogenous leukemia) and HT29 (colorectal adenocarcinoma) cells are derived from female patients. Jurkat (acute T-cell leukemia), HuTu-80 (duodenal adenocarcinoma), and HepG2 (hepatocellular carcinoma) cells are derived from male patients. HEK293T (embryonic kidney) cells are derived from a female fetus. RPE1 (immortalized retinal pigment epithelium) cells are derived from a female subject. All cell lines were grown at 37 °C in the presence of 5% CO_2.

To generate the K562 cell lines stably expressing various CRISPRi effectors, parental K562 cells were stably transduced with lentiviral vectors expressing the corresponding effectors linked to GFP via a P2A ribosome skipping sequence from an SFFV promoter with an upstream ubiquitous chromatin opening element (UCOE). Polyclonal populations of GFP-positive cells were selected using two rounds of fluorescence-activated cell sorting (FACS) on a Sony SH800S Cell Sorter.

To generate RPE1 cells stably expressing Zim3-dCas9, RPE-1 cells were infected with lentivirus containing UCOE-SFFV-Zim3-dCas9-P2A-BFP (pJB108) at low multiplicity of infection by centrifugation at 1000 × g. Polyclonal populations of BFP-positive cells were selected using two rounds of FACS on a Sony SH800S Cell Sorter. To generate Jurkat cells stably expressing Zim3-dCas9, Jurkat cells were infected with virus containing UCOE-EF1α-Zim3-dCas9-P2A-mCh (pJB109) at low multiplicity of infection by centrifugation at 1000 × g. Polyclonal populations of mCherry-positive cells were selected using two rounds of FACS on a Sony SH800S Cell Sorter. To generate HepG2, HuTu-80, and HT29 cells stably expressing Zim3-dCas9, cells were infected with lentivirus containing UCOE-EF1α-Zim3-dCas9-P2A-mCh (pJB109) at low multiplicity of infection. Polyclonal populations of mCherry-positive cells were selected using two rounds of FACS on a FACSAria II Cell Sorter (BD Biosciences). To generate K562 cells stably expressing Zim3-dCas9 without a fluorescent marker, K562 cells were infected with virus containing UCOE-SFFV-Zim3-dCas9-P2A-hygro (pAG389) at low multiplicity of infection by centrifugation at 1000 × g. To select for a polyclonal population, cells were treated 48 hours after infection with 200 µg/mL hygromycin for one week, followed by treatment 500 µg/ml hygromycin for three weeks.

Lentivirus production

Lentivirus was generated by transfecting HEK239T cells with the transfer plasmid and four packaging plasmids (for expression of VSV-G, Gag/Pol, Rev, and Tat) using TransIT-LT1 Transfection Reagent (Mirus Bio). Viral supernatant was harvested two days after transfection and filtered through 0.44 µm PES filters and/or frozen at –80 °C prior to transduction.

Design and cloning of pilot genome-wide single- and dual-sgRNA CRISPRi libraries

To compare the use of single- and dual-sgRNA CRISPRi libraries in systematic genetic screens, pilot genome-wide single- and dual-sgRNA CRISPRi libraries were designed and cloned. sgRNAs targeting each gene were selected from our previously published hCRISPRi v2 library by balancing empirical data from previous genetic screens with Horlbeck *et al.* predicted rankings (Horlbeck et al., 2016a) using a three-tiered approach:

Tier 1. For genes essential for growth in the K562 CRISPRi screen data (p-value < 0.001 and $\gamma < -0.2$) (Horlbeck et al., 2016a), sgRNAs were ranked by their growth phenotype.

Tier 2. As many genetic perturbations only cause a conditional cellular phenotype (e.g., in a particular cell type, chemical stressor, or reporter phenotype), we next aggregated data across multiple genetic screens (only a subset of the data in Table S2 was available for the pilot library design). For genes that were identified as a significant hit [FDR 0.05 based on MAGeCK RRA p-value (Li et al., 2014)] in previous CRISPRi screens, sgRNAs were ranked by the sum of z-scored phenotypes across screens.

Tier 3. For all other genes, sgRNAs were ranked by the regression scores in hCRISPRi v2.1 (Horlbeck et al., 2016a).

Using this ranking scheme, we selected the single best sgRNA for a single-sgRNA/single-element-per-gene library (dJR004) and the two best sgRNAs for a dual-sgRNA/single-element-per-gene library (dJR020). A list of sgRNA targeting sequences both the single and dual sgRNA libraries is available in Table S1.

The single-sgRNA library dJR004 was cloned using the protocol described here: https://weissman.wi.mit.edu/resources/Pooled_CRISPR_Library_Cloning.pdf. A modified CROP-seq sgRNA lentiviral expression vector (pJR107) was derived from the parental vector pBA950 (https://www.addgene.org/122239/) by incorporating a GFP fluorescent marker and a UCOE element upstream of the EF1alpha promoter to prevent marker silencing. sgRNA targeting sequences were appended with flanking sequence, BstX1/BlpI overhangs, and PCR adapters. The library was synthesized as an oligonucleotide pool (Twist Biosciences), PCR-amplified, BstX1/BlpI-digested, and inserted into pJR107 by ligation.

The dual-sgRNA library dJR020 was cloned using the protocol in Supplementary Note 1. Briefly, dual-sgRNA targeting sequences were spaced by a BsmBI-cut site and appended with flanking sequence, BstX1/BlpI overhangs, and PCR adapters with the structure: with the structure: 5'- PCR adaptor - CCACCTTGTTG – targeting sequence A -

gtttcagagcgagacgtgcctgcaggatacgtctcagaaacatg – targeting sequence B - GTTTAAGAGCTAAGCTG - PCR adaptor-3'. The library was synthesized as an oligonucleotide pool (Twist Biosciences), PCR-amplified, BstX1/BlpI-digested, and inserted into pJR104 by ligation. Next, the sgRNA CR3/hU6 promoter insert pJR98 was BsmBI-digested and ligated into the BsmBI-digested library to generate the final library. In the final library, each element expresses two unique sgRNAs from tandem U6 expression cassettes.

Genome-wide growth screens for library comparison

Parallel growth screens were performed to compare dJR004 versus dJR020. Lentivirus from dJR004 and dJR020 was produced in HEK293T as described above. CRISPRi K562 cells expressing dCas9-KOX1 KRAB were spinfected (1000G) with polybrene (8 ug/ml) with lentivirus from dJR004 and dJR020 in biological replicate. Throughout the screen, cells were maintained at a density between 250,000 – 1,000,000 cells per ml and 1000X coverage per library element. On day 3 post-transduction, an infection rate of 11%-18% was measured by GFP fluorescence. On day 3 through day 6 post-transduction, puromycin at 1 ug/ml was used to select for infected cells, and cells were allowed to recover for two days. On day 8 post-transduction, a cell pellet was frozen for each replicate representing the initial sample (T_0) of the screen. Throughout the screen, the number of cell doublings was recorded, and final samples (T_{final}) were collected on day 20 post-transduction.

Screen library preparation, sequencing, and analysis

Amplicon DNA libraries were prepared from cell pellets as previously described (Nuñez et al., 2021). Genomic DNA was isolated using a NucleoSpin Blood XL kit or NucleoSpin Blood L kit (Macherey–Nagel) depending on pellet size. Purified genomic DNA was directly amplified by 22 cycles of PCR using NEBNext Ultra II Q5 PCR MasterMix (NEB). Sequencing was performed on a NovaSeq 6000 (Illumina) using a 19 bp Read 1, 19 bp Read 2, and 5 bp Index Read 1 with custom sequencing primers.

After sequencing, sgRNA sequencing reads were aligned to the single and dual sgRNA libraries using a custom Python script without allowing mismatches. Reads for which the Read 1 and Read 2 sgRNA sequences did not target the same gene likely arose from lentiviral recombination and were discarded from downstream analysis. For both replicates of the dual-

sgRNA library, 29.4%, of mapped reads contained sgRNAs targeting different genes. For each sgRNA or sgRNA pair, the growth phenotype was defined as the $\log_2$(sgRNA count T_{final} / sgRNA count T_0) – median non-targeting control $\log_2$(sgRNA count T_{final} / sgRNA count T_0) divided by the replicate total cell doublings (Gilbert et al., 2014). For the analysis of the Cancer Dependency Map (DepMap) Common Essential genes, the 20Q1 Common Essential genes were downloaded from https://depmap.org/portal/download/. For receiver operating characteristic (ROC) curve analysis, "positives" were defined as genes with a K562 CRISPRi growth screen *p*-value < 0.001 and $\gamma < -0.05$ (Horlbeck et al., 2016a), and "negatives" were defined as non-targeting control guide pairs.

Empirical sgRNA selection, incorporation of integration barcodes, and validation of finalized dual-sgRNA CRISPRi libraries

While the pilot dual-sgRNA library dJR020 enabled validation of the dual-sgRNA strategy, finalized dual-sgRNA libraries were designed with additional considerations. An expanded set of aggregated CRISPR screen data was used to optimize guide selection (Table S2). Optimal sgRNAs targeting each gene were selected using an updated set of rules. First, sgRNAs containing a BsmBI target sequence (CGTCTC or GAGACG) were removed to avoid dropout during cloning. Second, each transcript per gene in Horlbeck *et al.*, 2016a was targeted independently. Genes were separated into three tiers, similar to the tiers described for the pilot library but with additional considerations:

Tier 1 (n=662 genes). For genes essential for growth in the K562 CRISPRi screen data (*p*-value < 0.001 and $\gamma < -0.2$) (Horlbeck et al., 2016a), sgRNAs were ranked by their growth phenotypes (calculated relative to the best-performing sgRNA targeting each gene per screen in which the gene was a significant hit at FDR 0.05).

Tier 2 (n=4,033 genes): The ranking strategy used to generate the pilot library (dJR020) included any gene identified as a significant hit in any previous CRISPRi screen for empirical guide selection and as such did not control for the increased chance that a gene may score as a false positive in a screen as the number of screens increases (the equivalent of multiple comparisons). To control for such false positives, the 320 olfactory genes served as a negative control set. None of the 320 olfactory genes were a significant hit [FDR 0.05 based on MAGeCK RRA *p*-value (Li et al., 2014)] in greater than four previous CRISPRi screens. Therefore, as a

first cutoff, any gene that was identified as a significant hit in five or more previous CRISPRi screens, regardless of the strength of the phenotype, was included in this tier.

This cutoff misses genes that score strongly, and as such are high-confidence hits, in a small number of screens. To also include such genes, each gene that was a significant hit [FDR 0.05 based on MAGeCK RRA *p*-value (Li et al., 2014)] in one to four screens was assigned a score based on the maximum absolute value discriminant score (calculated as the $-\log_{10}$ *p*-value multiplied by the mean *z*-scored phenotype of the top three sgRNAs), summed across screens in which the gene scored as a hit. As a comparison, this same score was calculated for olfactory genes. Genes were included in this tier if the discriminant score was greater than a threshold calculated from the olfactory gene scores for the same number of screens in which a gene was identified as a hit.

For all genes included in this tier, sgRNAs were ranked by the average of phenotypes across screens in which the gene was identified as a hit. Only sgRNAs that were identified as a hit at FDR<0.01 in at least one screen were ranked. sgRNA phenotypes were calculated relative to the best performing sgRNA targeting each gene per screen in which the gene was a significant hit at FDR 0.05.

Tier 3 (n=14,493 genes): For all other genes, sgRNAs were ranked by the regression scores in hCRISPRi v2.1 (Horlbeck et al., 2016a).

Using this ranking scheme, we selected the first and second ranked sgRNAs for a dual-sgRNA/single-element-per-gene sublibrary (hCRISPRi_dual_1_2), the third and fourth ranked sgRNAs for a second dual-sgRNA/single-element-per-gene sublibrary (hCRISPRi_dual_3_4), and the fifth and sixth ranked sgRNAs for a final dual-sgRNA/single-element-per-gene sublibrary (hCRISPRi_dual_5_6). Each library also contains a set of non-targeting control dual sgRNAs representing 5% of the total library elements. A list of sgRNA targeting sequences for all libraries is available in Table S3.

Integration barcodes (IBCs) were incorporated between the tandem sgRNA cassettes in the dual-sgRNA library in four steps. First, a library of 215 8-nucleotide IBCs were designed with a Hamming distance ≥ 4 and between 25-75% GC content (Table S4). Second, the library of IBCs were cloned into pJR98 in an arrayed format. pJR98 was digested by AscI and ssDNA oligo donors of the sequence 5' CTCTTCCTGCCCGACCTTGGGG – reverse complement IBC – CAGCGCCATAGCTGAGTGTAGATTCGAGC – 3' were cloned into the vector using

NEBuilder HiFI DNA Assembly Master Mix (NEB). Third, the library of cloned IBCs were Sanger verified and pooled at a equimolar ratio for all barcodes. Fourth, the library was cloned into the dual-sgRNA library by BsmBI-digestion and ligation. Sequencing was performed on a NovaSeq 6000 (Illumina) using a 19 bp Read 1, 19 bp Read 2, 8 bp Index Read 1, and 8 bp Index Read 2 with custom sequencing primers as described in Supplementary Note 2. Demultiplexing on only the i5 index using the i7 index (IBC) as a read was performed as detailed: https://gist.github.com/sumeetg23/a064a36801d2763e94da2e191699fb9f.

Perturb-seq comparison of dual sgRNA libraries versus Dolcetto

Direct capture Perturb-seq (Replogle et al., 2020) was used to directly compare the knockdown produced by the dual-sgRNA libraries versus the Dolcetto Set A CRISPRi library. N=128 genes were randomly selected from the 4,000 most highly-expressed genes in K562 cells based on RNA-seq (https://www.encodeproject.org/experiments/ENCSR000AEL/). Two parallel libraries were cloned: a library containing the three dual-sgRNA elements targeting each gene and a library containing the three Dolcetto Set A guides targeting each gene, plus non-targeting control guides. For Dolcetto sgRNAs, the 5′ base was replaced with a G to enable expression from the U6 promoter. The Dolcetto single-sgRNA library was cloned as described above into pJR101 guide expression vector containing a Perturb-seq capture sequence in stem loop 2. The dual-sgRNA library cloned as described above into pJR101 with a pJR98 insert cassette containing a Perturb-seq capture sequence in stem loop 2 of guide B. After library verification by sequencing, lentivirus was prepared in HEK293T as described above.

For Perturb-seq, CRISPRi K562 cells expressing dCas9-KOX1 KRAB (Gilbert et al., 2014) were spinfected (1000 × *g*) with polybrene (8 µg/ml) with lentivirus from both libraries in parallel. Throughout the screen, cells were maintained at a density between 250,000 – 1,000,000 cells per ml and 1000× coverage per library element. On day 3 post-transduction, an infection rate of 5% was measured for both screens, and infected cells were sorted by FACS (BD FACS Aria). On day 7 post-transduction, cells were prepared for single-cell RNA-sequencing as detailed in the 10x Genomics Single Cell Protocols Cell Preparation Guide (10x Genomics, CG00053 Rev C) and separated into droplet emulsions using the Chromium Controller (10x Genomics) with Chromium Single-Cell 3′ Gel Beads v3.1 (10x Genomics, PN-1000121 and PN-1000120) across 12 lanes/gemgroups with the goal of recovering ~15,000 cells per GEM group before filtering.

Sequencing libraries were prepared following the 10x Genomics Chromium Single Cell 3' Reagent Kits User Guide (v3.1 Chemistry) with Feature Barcoding technology for CRISPR Screening (CG000205; Rev C). Libraries were sequenced on a NovaSeq 6000 (Illumina) according to the 10x Genomics User Guide.

After sequencing, mRNA and sgRNA counts were obtained from Cell Ranger 4.0.0 software (10x Genomics). To assign guides to cells, we used a Poisson-Gaussian mixture model as previously described (Replogle et al., 2020). Only cells bearing a single Dolcetto sgRNA or a single dual-sgRNA guide B sgRNA were used for downstream calculation of CRISPRi efficacy. For each guide, the on-target knockdown was calculated as the fraction of mRNA remaining (target gene expression in targeting cells relative to cells bearing non-targeting control guides).

Design and cloning of constructs for CRISPRi effector expression

All CRISPRi effectors were cloned into a lentiviral backbone containing a ubiquitous chromatin opening element and a spleen focus forming virus (SFFV) promoter (pMH0001, Addgene # 85969). Briefly, dCas9, effector domains, linker domains, and GFP were PCR amplified and inserted into backbone linearized by digest with MluI and NotI using Gibson assembly (NEBuilder HiFI DNA Assembly Master Mix, NEB). P2A sequences were incorporated into primer overhangs. The following additional considerations were incorporated into the final construct designs:

1) For KRAB from Kox1, the KRAB(KOX1) domain from dCas9-BFP-KRAB (Addgene # 46911) was fused to the C-terminus of dCas9, because C-terminal fusions of KRAB(KOX1) have historically produced the highest activity, linked by an 80-amino acid linker (XTEN80). XTEN80-KRAB(KOX1) was synthesized as a gBlock (IDT). We chose XTEN80 because we previously found that inclusion of a linker increases activity and the original dCas9-BFP-KRAB(KOX1) construct (Gilbert et al., 2013) underwent proteolytic cleavage between dCas9 and KRAB(KOX1) in some cell types, giving rise to free dCas9, a dominant negative for CRISPRi. The final construct is dCas9-XTEN80-KRAB(KOX1) or dCas9-Kox1 for short.
2) KRAB(ZIM3) was fused to the N-terminus of dCas9 with a 6-amino acid GS linker, which had produced the highest activity in a previous report, including when compared to C-terminal fusions (Alerasool et al., 2020). KRAB(ZIM3) was PCR-amplified from

pLX303-ZIM3-KRAB-dCas9 (Addgene # 154472). The final construct is KRAB(ZIM3)-dCas9 or Zim3-dCas9 for short. Note that this construct contains an additional nuclear localization signal between Zim3 and dCas9.

3) For SID4x, SID4x was fused to the N-terminus of dCas9-XTEN80-KRAB(Kox1), because SID4x had previously only been evaluated for CRISPRi in the context of a dual fusion (Carleton et al., 2017). A shorter 16-aa linker (XTEN16) was included between SID4x and dCas9, which has been a sufficient linker length at the N-terminus in the past. SID4x was amplified from a construct generously donated by the Aifantis lab (New York University). The final construct is SID4x-XTEN16-dCas9-XTEN80-KRAB(KOX1) or SID-dCas9-Kox1 for short.

4) For MeCP2, the previously reported dCas9-KRAB(Kox1)-MeCP2 construct (Addgene # 110821; Yeo et al., 2018) was PCR-amplified and transferred into the common backbone, giving rise to dCas9-Kox1-MeCP2. Note that this construct contains no linker between dCas9 and KRAB(Kox1), such that the KRAB(Kox1) domain may be largely inactive, and that the dCas9 uses different codons. We separately also generated a construct in which we fused MeCP2 to the C-terminus of the dCas9-XTEN80-KRAB(KOX1) construct. We observed similar growth defects and non-specific effects on the transcriptome using this construct.

Additional constructs with expression driven by a EF1α promoter were generated by performing analogous assemblies in the pMH0006 backbone (Addgene # 135448). Constructs with expression driven by CMV or EFS promoters were generated by replacing the SFFV promoter in existing constructs. Constructs in which effector expression is marked with BFP, mCherry, or hygromycin resistance were generated by assembling with PCR products containing the desired markers. Constructs in which expression of the fluorescent protein is linked by an internal ribosome entry site (IRES) from encephalomyocarditis virus (EMCV) were generated by incorporating a PCR fragment generated from pHR-TRE3G-TUBB-IRES-mCherry (Jost et al., 2017) instead of the P2A site. Constructs in which EGFP is flanked by loxP sites were generated by PCR-amplifying EGFP with primers containing loxP 2272 sequences (ATAACTTCGTATAAaGTATcCTATACGAAGTTAT). The amplicon was inserted by Gibson Assembly into pJB069 or pJB109 linearized by digestion with NotI and AsiSI. Finally, constructs in which the fluorescent proteins are constitutively linked to dCas9 were generated by

omitting the P2A sequence from primer overhangs. A full list of generated constructs is included in Table S6. All constructs have been deposited to Addgene.

Evaluation of effects of CRISPRi effectors on growth and transcription

K562 cell lines stably expressing CRISPRi effectors from an SFFV promoter linked to GFP via P2A were generated by lentiviral transduction and FACS. Each effector expression construct was transduced in triplicate in parallel with all other constructs. 100,000 GFP-positive cells per replicate were isolated by FACS on a Sony SH800S Cell Sorter 5 d after transduction and allowed to recover.

To generate RNA-seq libraries of cells expressing each effector, 1×10^6 cells were harvested for each sample 6 d after FACS by centrifugation at 300 × *g* for 5 min and flash frozen in a dry ice and ethanol bath. RNA was extracted using the Direct-zol RNA Miniprep kit (Zymo Research) and quantified using the Qubit RNA BR Assay Kit (Life Technologies). RNA-seq libraries were prepared by the Whitehead Genome Technology Core facility using the Roche Diagnostics KAPA mRNA HyperPrep Kit. Paired-end 100 sequencing was performed on a NovaSeq (Illumina).

To evaluate growth of CRISPRi effector-expressing cells, a reference population of K562 cells stably expressing mCherry was generated by lentiviral transduction of pU6-sgRNA EF1Alpha-puro-T2A-mCherry (a gift from Gregory Ow and Eric Collisson, UCSF) and FACS. This was conducted in parallel with the generation of CRISPRi effector-expressing cells. 7 days after sorting, ~125,000 cells per GFP-sorted population (different CRISPRi effectors) were mixed with ~125,000 mCherry-sorted cells (reference population). The ratio of mCherry-positive to mCherry-negative cells was read out immediately after mixing and periodically for the next 19 days by flow cytometry on an Attune NxT (ThermoFisher).

RNA-seq data analysis

Sequencing reads were aligned strand-specifically to the human genome (GRCh38) and then aggregated by gene using only reads uniquely mapped to the reverse strand using the spliced read aligner STAR (Dobin et al., 2013), version 2.7.9, against an index containing features from Ensembl release 98 / GENCODE v32 (downloaded from 10x Genomics reference 2020-A). Replicate sample 2 for cells expressing dCas9-Kox1 had substantially fewer reads than expected

and was excluded from analysis. For clustering analysis, transcript counts were normalized to transcripts per million for each sample, filtered for the 2000 most highly expressed genes on average, and clustered using the Ward variance minimization algorithm implemented in scipy version 1.6.2. Differential expression analysis was carried out on gene counts using DESeq2 (Love et al., 2014). For Figure 2E, transcript counts were not filtered. The trends for numbers of differentially expressed genes were equivalent when only including genes with an average count > 2 across all samples.

Selection and cloning of individual sgRNAs

Strong sgRNAs against essential genes or cell surface markers were selected from the hCRISPRi-v2 library (Horlbeck et al., 2016a; Nuñez et al., 2021). Intermediate-activity sgRNAs were selected either from the hCRISPRi-v2 library or by incorporating defined mismatches in strong sgRNAs (Jost et al., 2020). All sgRNA sequences used for individual evaluation are listed in Table S6.

Individual sgRNA expression constructs were cloned as described previously (Gilbert et al., 2014). Briefly, two complementary oligonucleotides (IDT), containing the sgRNA targeting region as well as overhangs matching those left by restriction digest of the vector with BstXI and BlpI, were annealed and ligated into pCRISPRia-v2 [pU6-sgRNA EF1Alpha-puro-T2A-BFP with two SbfI sites flanking the sgRNA expression cassette, Addgene #84832 (Horlbeck et al., 2016a)] or pU6-sgRNA EF1Alpha-puro-T2A-mCherry [a gift from Gregory Ow and Eric Collisson, UCSF, (Jost et al., 2020)] digested with BstXI (NEB or Thermo Fisher Scientific) and BlpI (NEB) or Bpu1102I (Thermo Fisher Scientific). The ligation product was transformed into Stellar chemically competent *E. coli* cells (Takara Bio) and plasmid was prepared following standard protocols. The resulting sgRNA expression vectors were individually packaged into lentivirus as described above.

Evaluation of individual sgRNA phenotypes

Effects of sgRNAs targeting essential genes on cell growth were measured in internally controlled growth assays by transducing cells with mCherry-marked sgRNA expression constructs at MOI < 0.5 (15 – 40% infected cells) and measuring the fraction of sgRNA-expressing cells 3-12 days after transduction as mCherry-positive cells by flow cytometry on an

Attune NxT (ThermoFisher). All experiments were performed in duplicates from the infection step.

Effects of sgRNAs on expression levels of cell surface proteins were measured by flow cytometry. K562 or Jurkat cell lines expressing CRISPRi effectors of interest were infected with lentivirus containing sgRNA expression vectors by centrifugation at 1000 × *g* for 1 h in 24-well plates in the presence of 8 µg/mL polybrene (Sigma-Aldrich). RPE1, HepG2, HuTu-80, and HT29 cell lines expressing Zim3-dCas9 were infected with lentivirus containing sgRNA expression vectors for 24 h in the presence of 8 µg/mL polybrene. 6-14 d after transduction, cells were harvested by centrifugation (suspension cells) or trypsin-free detachment (adherent cells; mechanical detachment or EDTA), washed once in flow cytometry buffer (PBS with 5% (v/v) FBS), and stained at room temperature for 15-30 min with APC-conjugated antibodies targeting CD55 (clone JS11, BioLegend 311311), CD81 (clone 5A6, Biolegend 349509), CD151 (clone 50-6, BioLegend 350405), CD29 (clone TS2/16, Biolegend 303007), or B2M (clone 2M2, Biolegend 316312) diluted 1:100 in flow cytometry buffer. After staining, cells were washed twice in 200 µL flow cytometry buffer and resuspended in flow cytometry buffer for measurement on an Attune NxT (ThermoFisher), LSR-II (BD Biosciences) or Symphony A3 (BD Biosciences).

Optimal dilutions for each antibody were determined by testing 1:20, 1:100, and 1:500 antibody titrations on K562 cells with epitope-targeting or non-targeting sgRNAs and choosing the titration with the maximum signal difference.

Flow cytometry data were analyzed using FlowCytometryTools 0.5.0 (https://eyurtsev.github.io/FlowCytometryTools/) and python 3.8. Briefly, the data were gated for cells (FSC-A versus SSC-A), FSC singlets (FSC-W versus FSC-H for data recorded on an Attune NxT and FSC-W versus FSC-A for data recorded on an LSR-II), SSC singlets (SSC-W versus SSC-H for data recorded on an Attune NxT and SSC-W versus SSC-A for data recorded on an LSR-II), and sgRNA-expressing cells (BFP- or mCherry-positive, depending on the experiment). Background APC fluorescence intensity from unstained cells or cells stained with an APC-conjugated Mouse IgG1, κ isotype control (BioLegend clone MOPC-21) was subtracted to correct for background fluorescence. Knockdown was quantified using median background-corrected APC fluorescence intensity in cells expressing a targeting sgRNA relative to intensity in cells expressing a non-targeting control sgRNA, with the exception of the Jurkat and RPE1

826 experiments, for which knockdown was quantified using median background-corrected APC
827 fluorescence intensity in cells expressing a targeting sgRNA relative to intensity in cells not
828 expressing an sgRNA in the same well.
829

References

Adamson B, Norman TM, Jost M, Cho MY, Nuñez JK, Chen Y, Villalta JE, Gilbert LA, Horlbeck MA, Hein MY, Pak RA, Gray AN, Gross CA, Dixit A, Parnas O, Regev A, Weissman JS. 2016. A Multiplexed Single-Cell CRISPR Screening Platform Enables Systematic Dissection of the Unfolded Protein Response. *Cell* 167:1867-1882.e21. doi:10.1016/j.cell.2016.11.048

Adamson B, Norman TM, Jost M, Weissman JS. 2018. Approaches to maximize sgRNA-barcode coupling in Perturb-seq screens. *Biorxiv* 298349. doi:10.1101/298349

Alerasool N, Segal D, Lee H, Taipale M. 2020. An efficient KRAB domain for CRISPRi applications in human cells. *Nat Methods* 17:1093–1096. doi:10.1038/s41592-020-0966-x

Behan FM, Iorio F, Picco G, Gonçalves E, Beaver CM, Migliardi G, Santos R, Rao Y, Sassi F, Pinnelli M, Ansari R, Harper S, Jackson DA, McRae R, Pooley R, Wilkinson P, Meer D van der, Dow D, Buser-Doepner C, Bertotti A, Trusolino L, Stronach EA, Saez-Rodriguez J, Yusa K, Garnett MJ. 2019. Prioritization of cancer therapeutic targets using CRISPR–Cas9 screens. *Nature* 568:511–516. doi:10.1038/s41586-019-1103-9

Bock C, Datlinger P, Chardon F, Coelho MA, Dong MB, Lawson KA, Lu T, Maroc L, Norman TM, Song B, Stanley G, Chen S, Garnett M, Li W, Moffat J, Qi LS, Shapiro RS, Shendure J, Weissman JS, Zhuang X. 2022. High-content CRISPR screening. *Nat Rev Methods Primers* 2:8. doi:10.1038/s43586-021-00093-4

Bosch B, DeJesus MA, Poulton NC, Zhang W, Engelhart CA, Zaveri A, Lavalette S, Ruecker N, Trujillo C, Wallach JB, Li S, Ehrt S, Chait BT, Schnappinger D, Rock JM. 2021. Genome-wide gene expression tuning reveals diverse vulnerabilities of M. tuberculosis. *Cell* 184:4579-4592.e24. doi:10.1016/j.cell.2021.06.033

Bowden AR, Morales-Juarez DA, Sczaniecka-Clift M, Agudo MM, Lukashchuk N, Thomas JC, Jackson SP. 2020. Parallel CRISPR-Cas9 screens clarify impacts of p53 on screen performance. *Elife* 9:e55325. doi:10.7554/elife.55325

Carleton JB, Berrett KC, Gertz J. 2017. Multiplex Enhancer Interference Reveals Collaborative Control of Gene Regulation by Estrogen Receptor α-Bound Enhancers. *Cell Syst* 5:333-344.e5. doi:10.1016/j.cels.2017.08.011

Coukos R, Yao D, Sanchez MI, Strand ET, Olive ME, Udeshi ND, Weissman JS, Carr SA, Bassik MC, Ting AY. 2021. An engineered transcriptional reporter of protein localization identifies regulators of mitochondrial and ER membrane protein trafficking in high-throughput CRISPRi screens. *Elife* 10:e69142. doi:10.7554/elife.69142

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29:15–21. doi:10.1093/bioinformatics/bts635

Doench JG. 2018. Am I ready for CRISPR? A user's guide to genetic screens. *Nat Rev Genet* 19:67–80. doi:10.1038/nrg.2017.97

Dong F, Xie K, Chen Y, Yang Y, Mao Y. 2017. Polycistronic tRNA and CRISPR guide-RNA enables highly efficient multiplexed genome engineering in human cells. *Biochem Bioph Res Co* 482:889–895. doi:10.1016/j.bbrc.2016.11.129

Dräger NM, Sattler SM, Huang CT-L, Teter OM, Leng K, Hashemi SH, Hong J, Aviles G, Clelland CD, Zhan L, Udeochu JC, Kodama L, Singleton AB, Nalls MA, Ichida J, Ward ME, Faghri F, Gan L, Kampmann M. 2022. A CRISPRi/a platform in iPSC-derived microglia

876 uncovers regulators of disease states. Biorxiv 2021.06.16.448639.
877 doi:10.1101/2021.06.16.448639
878 Du D, Roguev A, Gordon DE, Chen M, Chen S-H, Shales M, Shen JP, Ideker T, Mali P, Qi LS,
879 Krogan NJ. 2017. Genetic interaction mapping in mammalian cells using CRISPR
880 interference. *Nat Methods* 14:577–580. doi:10.1038/nmeth.4286
881 Ecco G, Imbeault M, Trono D. 2017. KRAB zinc finger proteins. *Development* 144:2719–2729.
882 doi:10.1242/dev.132605
883 Feldman D, Singh A, Garrity AJ, Blainey PC. 2018. Lentiviral co-packaging mitigates the effects
884 of intermolecular recombination and multiple integrations in pooled genetic screens. *Biorxiv*
885 262121. doi:10.1101/262121
886 Fulco CP, Munschauer M, Anyoha R, Munson G, Grossman SR, Perez EM, Kane M, Cleary B,
887 Lander ES, Engreitz JM. 2016. Systematic mapping of functional enhancer–promoter
888 connections with CRISPR interference. *Science* 354:769–773. doi:10.1126/science.aag2445
889 Fulco CP, Nasser J, Jones TR, Munson G, Bergman DT, Subramanian V, Grossman SR, Anyoha
890 R, Doughty BR, Patwardhan TA, Nguyen TH, Kane M, Perez EM, Durand NC, Lareau CA,
891 Stamenova EK, Aiden EL, Lander ES, Engreitz JM. 2019. Activity-by-Contact model of
892 enhancer-promoter regulation from thousands of CRISPR perturbations. *Nat Genet* 51:1664–
893 1669. doi:10.1038/s41588-019-0538-0
894 Gasperini M, Hill AJ, McFaline-Figueroa JL, Martin B, Kim S, Zhang MD, Jackson D, Leith A,
895 Schreiber J, Noble WS, Trapnell C, Ahituv N, Shendure J. 2019. A Genome-wide Framework
896 for Mapping Gene Regulation via Cellular Genetic Screens. *Cell* 176:377-390.e19.
897 doi:10.1016/j.cell.2018.11.029
898 Gilbert LA, Horlbeck MA, Adamson B, Villalta JE, Chen Y, Whitehead EH, Guimaraes C,
899 Panning B, Ploegh HL, Bassik MC, Qi LS, Kampmann M, Weissman JS. 2014. Genome-
900 Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* 159:647–661.
901 doi:10.1016/j.cell.2014.09.029
902 Gilbert LA, Larson MH, Morsut L, Liu Z, Brar GA, Torres SE, Stern-Ginossar N, Brandman O,
903 Whitehead EH, Doudna JA, Lim WA, Weissman JS, Qi LS. 2013. CRISPR-Mediated
904 Modular RNA-Guided Regulation of Transcription in Eukaryotes. *Cell* 154:442–451.
905 doi:10.1016/j.cell.2013.06.044
906 Haapaniemi E, Botla S, Persson J, Schmierer B, Taipale J. 2018. CRISPR–Cas9 genome editing
907 induces a p53-mediated DNA damage response. *Nat Med* 24:927–930. doi:10.1038/s41591-
908 018-0049-z
909 Haswell JR, Mattioli K, Gerhardinger C, Maass PG, Foster DJ, Peinado P, Wang X, Medina PP,
910 Rinn JL, Slack FJ. 2021. Genome-wide CRISPR interference screen identifies long non-
911 coding RNA loci required for differentiation and pluripotency. *Plos One* 16:e0252848.
912 doi:10.1371/journal.pone.0252848
913 Hawkins JS, Silvis MR, Koo B-M, Peters JM, Osadnik H, Jost M, Hearne CC, Weissman JS,
914 Todor H, Gross CA. 2020. Mismatch-CRISPRi Reveals the Co-varying Expression-Fitness
915 Relationships of Essential Genes in Escherichia coli and Bacillus subtilis. *Cell Syst* 11:523-
916 535.e9. doi:10.1016/j.cels.2020.09.009
917 Hill AJ, McFaline-Figueroa JL, Starita LM, Gasperini MJ, Matreyek KA, Packer J, Jackson D,
918 Shendure J, Trapnell C. 2018. On the design of CRISPR-based single-cell molecular screens.
919 *Nat Methods* 15:271–274. doi:10.1038/nmeth.4604

Horlbeck MA, Gilbert LA, Villalta JE, Adamson B, Pak RA, Chen Y, Fields AP, Park CY, Corn JE, Kampmann M, Weissman JS. 2016a. Compact and highly active next-generation libraries for CRISPR-mediated gene repression and activation. *Elife* 5:e19760. doi:10.7554/elife.19760

Horlbeck MA, Witkowsky LB, Guglielmi B, Replogle JM, Gilbert LA, Villalta JE, Torigoe SE, Tjian R, Weissman JS. 2016b. Nucleosomes impede Cas9 access to DNA in vivo and in vitro. *Elife* 5:e12677. doi:10.7554/elife.12677

Horlbeck MA, Xu A, Wang M, Bennett NK, Park CY, Bogdanoff D, Adamson B, Chow ED, Kampmann M, Peterson TR, Nakamura K, Fischbach MA, Weissman JS, Gilbert LA. 2018. Mapping the Genetic Landscape of Human Cells. *Cell* 174:953-967.e22. doi:10.1016/j.cell.2018.06.010

Ihry RJ, Worringer KA, Salick MR, Frias E, Ho D, Theriault K, Kommineni S, Chen J, Sondey M, Ye C, Randhawa R, Kulkarni T, Yang Z, McAllister G, Russ C, Reece-Hoyes J, Forrester W, Hoffman GR, Dolmetsch R, Kaykas A. 2018. p53 inhibits CRISPR–Cas9 engineering in human pluripotent stem cells. *Nat Med* 24:939–946. doi:10.1038/s41591-018-0050-6

Jost M, Chen Y, Gilbert LA, Horlbeck MA, Krenning L, Menchon G, Rai A, Cho MY, Stern JJ, Prota AE, Kampmann M, Akhmanova A, Steinmetz MO, Tanenbaum ME, Weissman JS. 2017. Combined CRISPRi/a-Based Chemical Genetic Screens Reveal that Rigosertib Is a Microtubule-Destabilizing Agent. *Mol Cell* 68:210-223.e6. doi:10.1016/j.molcel.2017.09.012

Jost M, Santos DA, Saunders RA, Horlbeck MA, Hawkins JS, Scaria SM, Norman TM, Hussmann JA, Liem CR, Gross CA, Weissman JS. 2020. Titrating gene expression using libraries of systematically attenuated CRISPR guide RNAs. *Nat Biotechnol* 38:355–364. doi:10.1038/s41587-019-0387-5

Kearns NA, Pham H, Tabak B, Genga RM, Silverstein NJ, Garber M, Maehr R. 2015. Functional annotation of native enhancers with a Cas9 -histone demethylase fusion. *Nat Methods* 12:401–403. doi:10.1038/nmeth.3325

Klann TS, Black JB, Chellappan M, Safi A, Song L, Hilton IB, Crawford GE, Reddy TE, Gersbach CA. 2017. CRISPR–Cas9 epigenome editing enables high-throughput screening for functional regulatory elements in the human genome. *Nat Biotechnol* 35:561–568. doi:10.1038/nbt.3853

Knapp DJHF, Michaels YS, Jamilly M, Ferry QRV, Barbosa H, Milne TA, Fulga TA. 2019. Decoupling tRNA promoter and processing activities enables specific Pol-II Cas9 guide RNA expression. *Nat Commun* 10:1490. doi:10.1038/s41467-019-09148-3

Kosicki M, Tomberg K, Bradley A. 2018. Repair of double-strand breaks induced by CRISPR–Cas9 leads to large deletions and complex rearrangements. *Nat Biotechnol* 36:765–771. doi:10.1038/nbt.4192

Leng K, Rose IVL, Kim H, Xia W, Romero-Fernandez W, Rooney B, Koontz M, Li E, Ao Y, Wang S, Krawczyk M, TCW J, Goate A, Zhang Y, Ullian EM, Sofroniew MV, Fancy SPJ, Schrag MS, Lippmann ES, Kampmann M. 2022. CRISPRi screens in human astrocytes elucidate regulators of distinct inflammatory reactive states. Biorxiv 2021.08.23.457400. doi:10.1101/2021.08.23.457400

Li W, Xu H, Xiao T, Cong L, Love MI, Zhang F, Irizarry RA, Liu JS, Brown M, Liu XS. 2014. MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol* 15:554. doi:10.1186/s13059-014-0554-4

Liang JR, Lingeman E, Luong T, Ahmed S, Muhar M, Nguyen T, Olzmann JA, Corn JE. 2020. A Genome-wide ER-phagy Screen Highlights Key Roles of Mitochondrial Metabolism and ER-Resident UFMylation. *Cell* 180:1160-1177.e20. doi:10.1016/j.cell.2020.02.017

966 Liu SJ, Horlbeck MA, Cho SW, Birk HS, Malatesta M, He D, Attenello FJ, Villalta JE, Cho MY,
967 Chen Y, Mandegar MA, Olvera MP, Gilbert LA, Conklin BR, Chang HY, Weissman JS, Lim
968 DA. 2017. CRISPRi-based genome-scale identification of functional long noncoding RNA
969 loci in human cells. *Science* 355. doi:10.1126/science.aah7111
970 Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for
971 RNA-seq data with DESeq2. *Genome Biol* 15:550. doi:10.1186/s13059-014-0550-8
972 Luteijn RD, Zaver SA, Gowen BG, Wyman S, Garelis N, Onia L, McWhirter SM, Katibah GE,
973 Corn JE, Woodward JJ, Raulet DH. 2019. SLC19A1 transports immunoreactive cyclic
974 dinucleotides. *Nature* 573:434–438. doi:10.1038/s41586-019-1553-0
975 Mandegar MA, Huebsch N, Frolov EB, Shin E, Truong A, Olvera MP, Chan AH, Miyaoka Y,
976 Holmes K, Spencer CI, Judge LM, Gordon DE, Eskildsen TV, Villalta JE, Horlbeck MA,
977 Gilbert LA, Krogan NJ, Sheikh SP, Weissman JS, Qi LS, So P-L, Conklin BR. 2016.
978 CRISPR Interference Efficiently Induces Specific and Reversible Gene Silencing in Human
979 iPSCs. *Cell Stem Cell* 18:541–553. doi:10.1016/j.stem.2016.01.022
980 Meyers RM, Bryan JG, McFarland JM, Weir BA, Sizemore AE, Xu H, Dharia NV, Montgomery
981 PG, Cowley GS, Pantel S, Goodale A, Lee Y, Ali LD, Jiang G, Lubonja R, Harrington WF,
982 Strickland M, Wu T, Hawes DC, Zhivich VA, Wyatt MR, Kalani Z, Chang JJ, Okamoto M,
983 Stegmaier K, Golub TR, Boehm JS, Vazquez F, Root DE, Hahn WC, Tsherniak A. 2017.
984 Computational correction of copy number effect improves specificity of CRISPR–Cas9
985 essentiality screens in cancer cells. *Nat Genet* 49:1779–1784. doi:10.1038/ng.3984
986 Michlits G, Hubmann M, Wu S-H, Vainorius G, Budusan E, Zhuk S, Burkard TR, Novatchkova
987 M, Aichinger M, Lu Y, Reece-Hoyes J, Nitsch R, Schramek D, Hoepfner D, Elling U. 2017.
988 CRISPR-UMI: single-cell lineage tracing of pooled CRISPR–Cas9 screens. *Nat Methods*
989 14:1191–1197. doi:10.1038/nmeth.4466
990 Morgens DW, Chan C, Kane AJ, Weir NR, Li A, Dubreuil MM, Tsui CK, Hess GT, Lavertu A,
991 Han K, Polyakov N, Zhou J, Handy EL, Alabi P, Dombroski A, Yao D, Altman RB, Sello JK,
992 Denic V, Bassik MC. 2019. Retro-2 protects cells from ricin toxicity by inhibiting ASNA1-
993 mediated ER targeting and insertion of tail-anchored proteins. *Elife* 8:e48434.
994 doi:10.7554/elife.48434
995 Nuñez JK, Chen J, Pommier GC, Cogan JZ, Replogle JM, Adriaens C, Ramadoss GN, Shi Q,
996 Hung KL, Samelson AJ, Pogson AN, Kim JYS, Chung A, Leonetti MD, Chang HY,
997 Kampmann M, Bernstein BE, Hovestadt V, Gilbert LA, Weissman JS. 2021. Genome-wide
998 programmable transcriptional memory by CRISPR-based epigenome editing. *Cell* 184:2503-
999 2519.e17. doi:10.1016/j.cell.2021.03.025
1000 Przybyla L, Gilbert LA. 2021. A new era in functional genomics screens. *Nat Rev Genet* 1–15.
1001 doi:10.1038/s41576-021-00409-w
1002 Raffeiner P, Hart JR, García-Caballero D, Bar-Peled L, Weinberg MS, Vogt PK. 2020. An
1003 MXD1-derived repressor peptide identifies noncoding mediators of MYC-driven cell
1004 proliferation. *Proc National Acad Sci* 117:6571–6579. doi:10.1073/pnas.1921786117
1005 Replogle JM, Norman TM, Xu A, Hussmann JA, Chen J, Cogan JZ, Meer EJ, Terry JM, Riordan
1006 DP, Srinivas N, Fiddes IT, Arthur JG, Alvarado LJ, Pfeiffer KA, Mikkelsen TS, Weissman
1007 JS, Adamson B. 2020. Combinatorial single-cell CRISPR screens by direct guide RNA
1008 capture and targeted sequencing. *Nat Biotechnol* 38:954–961. doi:10.1038/s41587-020-0470-
1009 y
1010 Replogle JM, Saunders RA, Pogson AN, Hussmann JA, Lenail A, Guna A, Mascibroda L,
1011 Wagner EJ, Adelman K, Lithwick-Yanai G, Iremadze N, Oberstrass F, Lipson D, Bonnar JL,

1012 Jost M, Norman TM, Weissman JS. 2022. Mapping information-rich genotype-phenotype
1013 landscapes with genome-scale Perturb-seq. *Cell*. doi:10.1016/j.cell.2022.05.013
1014 Sage C le, Lawo S, Panicker P, Scales TME, Rahman SA, Little AS, McCarthy NJ, Moore JD,
1015 Cross BCS. 2017. Dual direction CRISPR transcriptional regulation screening uncovers gene
1016 networks driving drug resistance. *Sci Rep-uk* 7:17693. doi:10.1038/s41598-017-18172-6
1017 Sanson KR, Hanna RE, Hegde M, Donovan KF, Strand C, Sullender ME, Vaimberg EW,
1018 Goodale A, Root DE, Piccioni F, Doench JG. 2018. Optimized libraries for CRISPR-Cas9
1019 genetic screens with multiple modalities. *Nat Commun* 9:5416. doi:10.1038/s41467-018-
1020 07901-8
1021 Schellenberger V, Wang C, Geething NC, Spink BJ, Campbell A, To W, Scholle MD, Yin Y,
1022 Yao Y, Bogin O, Cleland JL, Silverman J, Stemmer WPC. 2009. A recombinant polypeptide
1023 extends the in vivo half-life of peptides and proteins in a tunable manner. *Nat Biotechnol*
1024 27:1186–1190. doi:10.1038/nbt.1588
1025 Schmidt R, Steinhart Z, Layeghi M, Freimer JW, Bueno R, Nguyen VQ, Blaeschke F, Ye CJ,
1026 Marson A. 2022. CRISPR activation and interference screens decode stimulation responses in
1027 primary human T cells. *Sci New York N Y* 375:eabj4008. doi:10.1126/science.abj4008
1028 Semesta KM, Tian R, Kampmann M, Zastrow M von, Tsvetanova NG. 2020. A high-throughput
1029 CRISPR interference screen for dissecting functional regulators of GPCR/cAMP signaling.
1030 *Plos Genet* 16:e1009103. doi:10.1371/journal.pgen.1009103
1031 Smits AH, Ziebell F, Joberty G, Zinn N, Mueller WF, Clauder-Münster S, Eberhard D, Savitski
1032 MF, Grandi P, Jakob P, Michon A-M, Sun H, Tessmer K, Bürckstümmer T, Bantscheff M,
1033 Steinmetz LM, Drewes G, Huber W. 2019. Biological plasticity rescues target activity in
1034 CRISPR knock outs. *Nat Methods* 16:1087–1093. doi:10.1038/s41592-019-0614-5
1035 Thakore PI, D'Ippolito AM, Song L, Safi A, Shivakumar NK, Kabadi AM, Reddy TE, Crawford
1036 GE, Gersbach CA. 2015. Highly specific epigenome editing by CRISPR-Cas9 repressors for
1037 silencing of distal regulatory elements. *Nat Methods* 12:1143–1149. doi:10.1038/nmeth.3630
1038 Tian R, Abarientos A, Hong J, Hashemi SH, Yan R, Dräger N, Leng K, Nalls MA, Singleton
1039 AB, Xu K, Faghri F, Kampmann M. 2021. Genome-wide CRISPRi/a screens in human
1040 neurons link lysosomal failure to ferroptosis. *Nat Neurosci* 24:1020–1034.
1041 doi:10.1038/s41593-021-00862-0
1042 Tian R, Gachechiladze MA, Ludwig CH, Laurie MT, Hong JY, Nathaniel D, Prabhu AV,
1043 Fernandopulle MS, Patel R, Abshari M, Ward ME, Kampmann M. 2019. CRISPR
1044 Interference-Based Platform for Multimodal Genetic Screens in Human iPSC-Derived
1045 Neurons. *Neuron* 104:239-255.e12. doi:10.1016/j.neuron.2019.07.014
1046 Tsherniak A, Vazquez F, Montgomery PG, Weir BA, Kryukov G, Cowley GS, Gill S,
1047 Harrington WF, Pantel S, Krill-Burger JM, Meyers RM, Ali L, Goodale A, Lee Y, Jiang G,
1048 Hsiao J, Gerath WFJ, Howell S, Merkel E, Ghandi M, Garraway LA, Root DE, Golub TR,
1049 Boehm JS, Hahn WC. 2017. Defining a Cancer Dependency Map. *Cell* 170:564-576.e16.
1050 doi:10.1016/j.cell.2017.06.010
1051 Xie S, Cooley A, Armendariz D, Zhou P, Hon GC. 2018. Frequent sgRNA-barcode
1052 recombination in single-cell perturbation assays. *Plos One* 13:e0198635.
1053 doi:10.1371/journal.pone.0198635
1054 Yeo NC, Chavez A, Lance-Byrne A, Chan Y, Menn D, Milanova D, Kuo C-C, Guo X, Sharma
1055 S, Tung A, Cecchi RJ, Tuttle M, Pradhan S, Lim ET, Davidsohn N, Ebrahimkhani MR,
1056 Collins JJ, Lewis NE, Kiani S, Church GM. 2018. An enhanced CRISPR repressor for

1057 targeted mammalian gene regulation. *Nat Methods* 15:611–616. doi:10.1038/s41592-018-
1058 0048-5
1059 Yin JA, Frick L, Scheidmann MC, Trevisan C, Dhingra A, Spinelli A, Wu Y, Yao L, Vena DL,
1060 De Cecco E, Ging K, Liu T, Täger J, Rodriguez S, Guo J, Berry S, Losa M, Hornemann S,
1061 Kampmann M, Pelkmans L, Hoepfner D, Heutink P, Aguzzi A. 2022. Robust and Versatile
1062 Arrayed Libraries for Human Genome-Wide CRISPR Activation, Deletion and Silencing.
1063 *bioRxiv*. doi:10.1101/2022.05.25.493370
1064 Zhu S, Cao Z, Liu Z, He Y, Wang Y, Yuan P, Li W, Tian F, Bao Y, Wei W. 2019. Guide RNAs
1065 with embedded barcodes boost CRISPR-pooled screens. *Genome Biol* 20:20.
1066 doi:10.1186/s13059-019-1628-0
1067
1068
1069
1070

Figures and Figure Legends

Figure 1

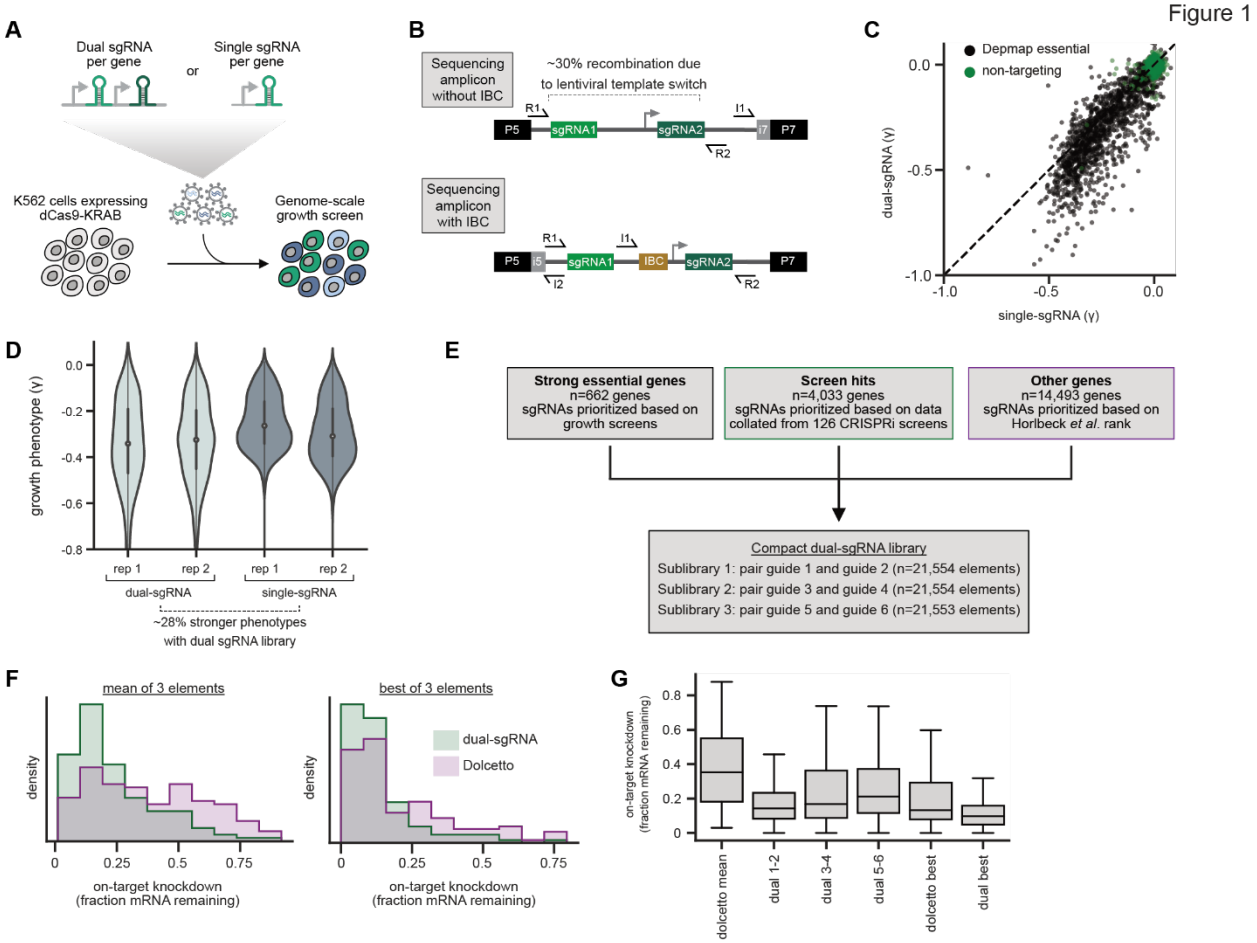


Figure 1. Design and validation of ultra-compact dual-sgRNA CRISPRi libraries.

a. Schematic of growth screen used to compare single- and dual-sgRNA libraries.

b. Schematic of dual-sgRNA library sequencing strategies.

c. Comparison of growth phenotypes for DepMap essential genes between single- and dual-sgRNA libraries. Sequencing libraries were prepared using the strategy labeled "Sequencing amplicon without IBC" in panel b. Growth phenotypes are reported as γ ($\log_2$ fold-enrichment of T_{final} over T_0, per doubling) and well-correlated between libraries ($r = 0.91$).

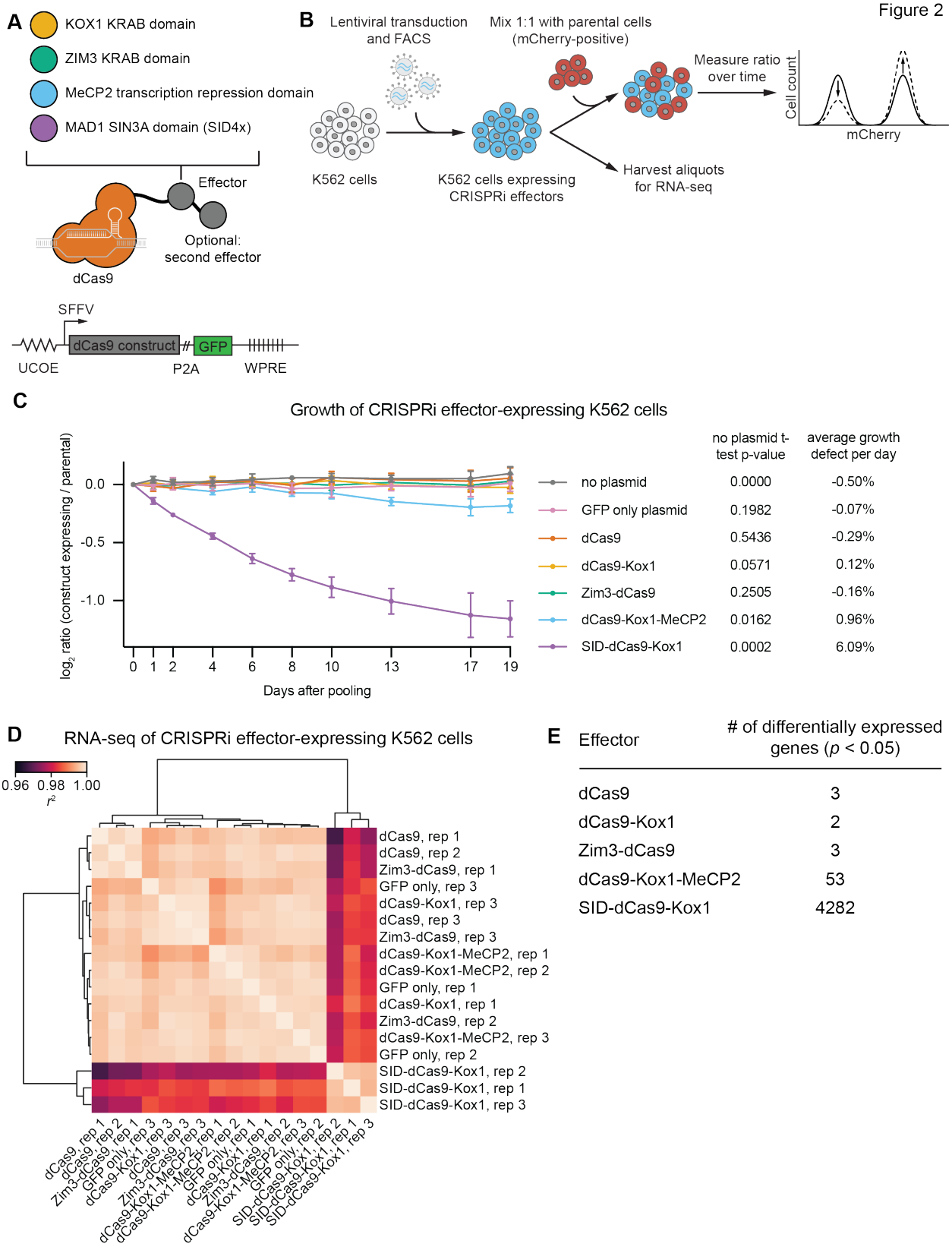
d. Comparison of growth phenotypes for DepMap essential genes between single- and dual-sgRNA libraries.

e. Design of final dual-sgRNA library.

f. Comparison of target gene knockdown by dual-sgRNA library versus Dolcetto. Target gene knockdown was measured by single-cell RNA-sequencing (Perturb-seq). For each library, the

1085 "mean of 3 elements" was calculated as the mean knockdown of all three elements targeting each
1086 gene. The "best of 3 elements" represents the element with the best knockdown per each gene.
1087 g. Comparison of target gene knockdown across elements in dual-sgRNA library versus
1088 Dolcetto.
1089 See also Figure S1.
1090

Figure 2



E

Effector	# of differentially expressed genes ($p < 0.05$)
dCas9	3
dCas9-Kox1	2
Zim3-dCas9	3
dCas9-Kox1-MeCP2	53
SID-dCas9-Kox1	4282

Figure 2. CRISPRi effectors containing SID or MeCP2 domains have non-specific effects

on cell viability and gene expression.

a. Schematics of CRISPRi transcription repressor domains and general lentiviral expression construct used for all CRISPRi effectors.

b. Experimental design to test effects of stable expression of each CRISPRi effector on growth and transcription in K562 cells.

c. Growth defects of effector-expressing cells, measured as the $\log_2$ ratio of mCherry-negative (effector-expressing) to mCherry-positive (not effector-expressing) cells in each well. mCherry levels were measured for 19 days after pooling cells. Data represent mean ± SD from three independent transductions of expression constructs. *p*-values are from an unpaired two-tailed t-test comparing D19 values for each sample to the D19 value for the ”no plasmid” sample. Average percent growth defect per day is the $\log_2$ D19 value divided by the number of days, multiplied by 100 for a percent value.

d. Clustered heatmap of correlation of transcript counts from K562 cells expressing indicated CRISPRi effectors or a GFP control. Correlations across samples were calculated using normalized counts (reads per million) for all genes with mean normalized count >1 and then clustered using the Ward variance minimization algorithm implemented in scipy. r^2 is squared Pearson correlation. Data represent three independent transductions of expression constructs.

e. Number of differentially expressed genes ($p < 0.05$) for cells expressing each effector versus cells expressing GFP only. *p*-values were calculated using a Wald test and corrected for multiple hypothesis testing as implemented in DeSeq2.

See also Figure S2.

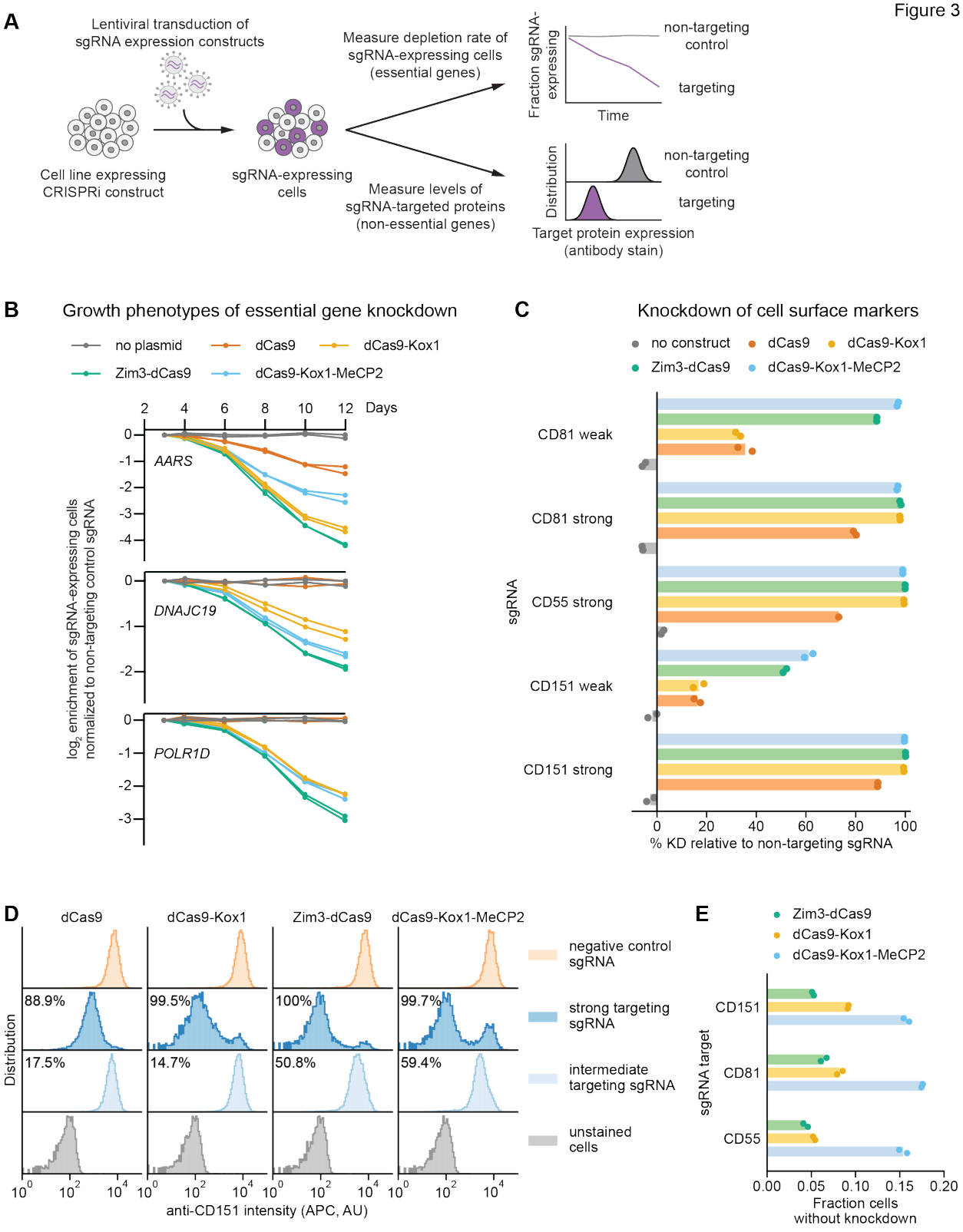


Figure 3. Zim3-dCas9 and dCas9-Kox1-MeCP2 mediate strongest knockdown.

a. Experimental design to measure knockdown mediated by different CRISPRi effectors by delivering sgRNAs targeting either essential genes or cell surface markers.
b. Depletion of K562 cells expressing essential gene-targeting sgRNAs and different CRISPRi effectors, measured as the ratio of mCherry-positive (sgRNA-expressing) to mCherry-negative (not sgRNA-expressing) cells in a given well. mCherry levels were measured for 12 days after transduction, starting on day 3. Data from two replicate transductions.
c. Percent knockdown of cell surface markers by different CRISPRi effectors in K562 cells. Cell surface marker levels were measured on day 6 post-transduction by staining with an APC-conjugated antibody. Knockdown was calculated as the ratio of median APC signal in sgRNA-expressing cells and median APC signal in cells expressing a non-targeting control sgRNA after subtraction of background APC signal. Data from two replicate transductions. Cells expressing dCas9 and a strong CD55-targeting sgRNA are represented by a single replicate.
d. Distribution of anti-CD151 signal intensity (APC) in individual cells from one representative transduction. Data from second replicate are shown in Figure S3b. Knockdown was quantified as in Figure 3c.
e. Fraction of cells without observable knockdown despite expressing a strong sgRNA, as quantified from the fluorescence distributions.
See also Figure S3.

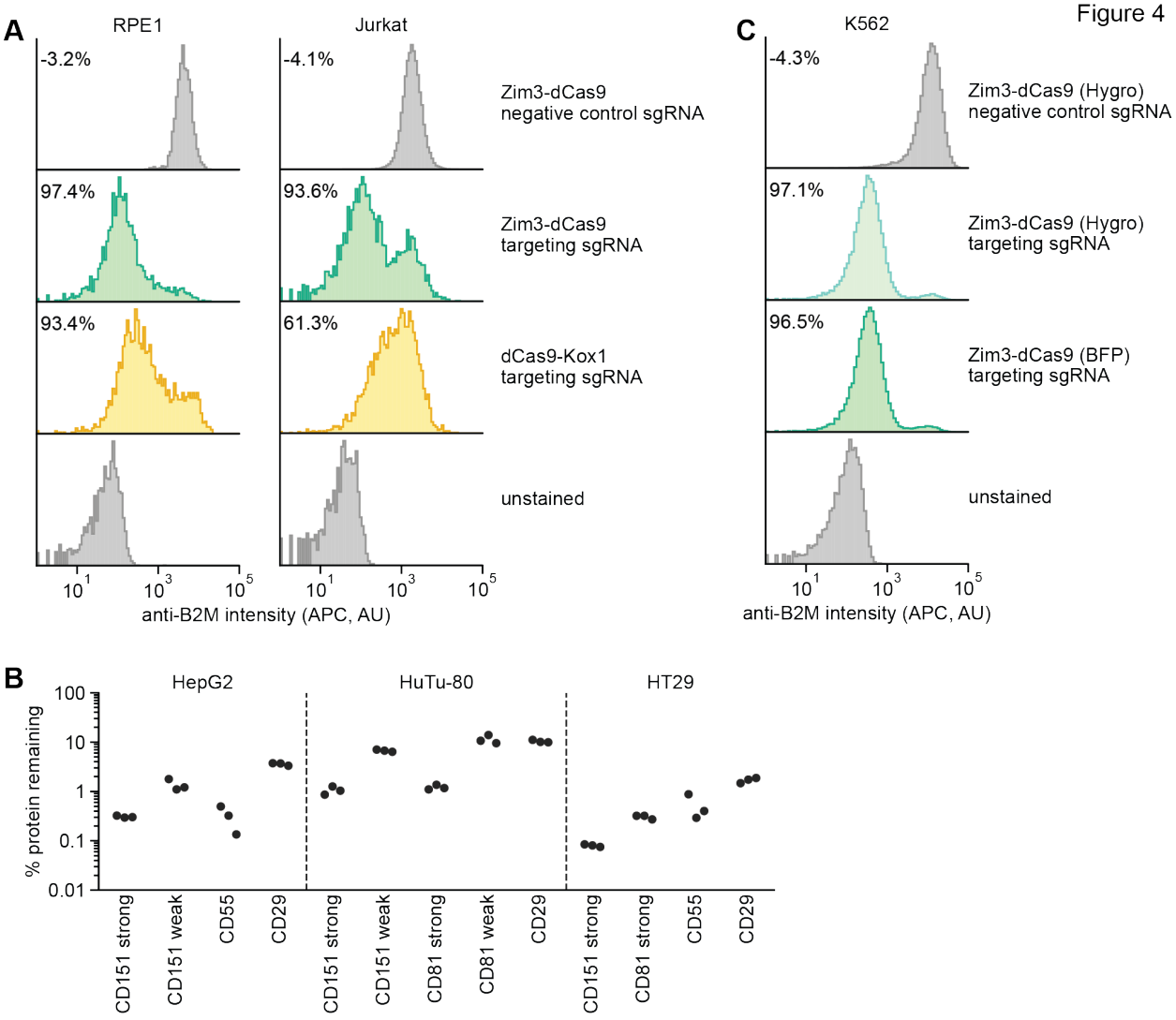


Figure 4. Validation of a suite of optimized Zim3-dCas9 cell lines.

a. Distribution of anti-B2M signal intensity (APC) in individual RPE1 (left) and Jurkat (right) cells expressing indicated CRISPRi effectors and sgRNAs. Knockdown was calculated as the ratio of median APC signal in transduced (sgRNA-expressing) cells and median APC signal in non-transduced cells in the same well, after subtraction of background APC signal.

b. Depletion of indicated cell surface markers in HepG2 (left), HuTu-80 (middle), and HT29 (right) cells expressing Zim3-dCas9. Cell surface marker levels were measured 6-14 days post-transduction by staining with APC-conjugated antibodies. Knockdown was calculated as the ratio of median APC signal in sgRNA-expressing cells and median APC signal in cells expressing a non-targeting control sgRNA after subtraction of background APC signal.

c. Distribution of anti-B2M signal intensity (APC) in individual K562 cells expressing indicated CRISPRi effectors and sgRNAs. The Zim3-dCas9 (Hygro) cell line was generated by transduction followed by hygromycin selection and does not express a fluorescent protein. Knockdown was calculated as in Figure 4a.
See also Figure S4.

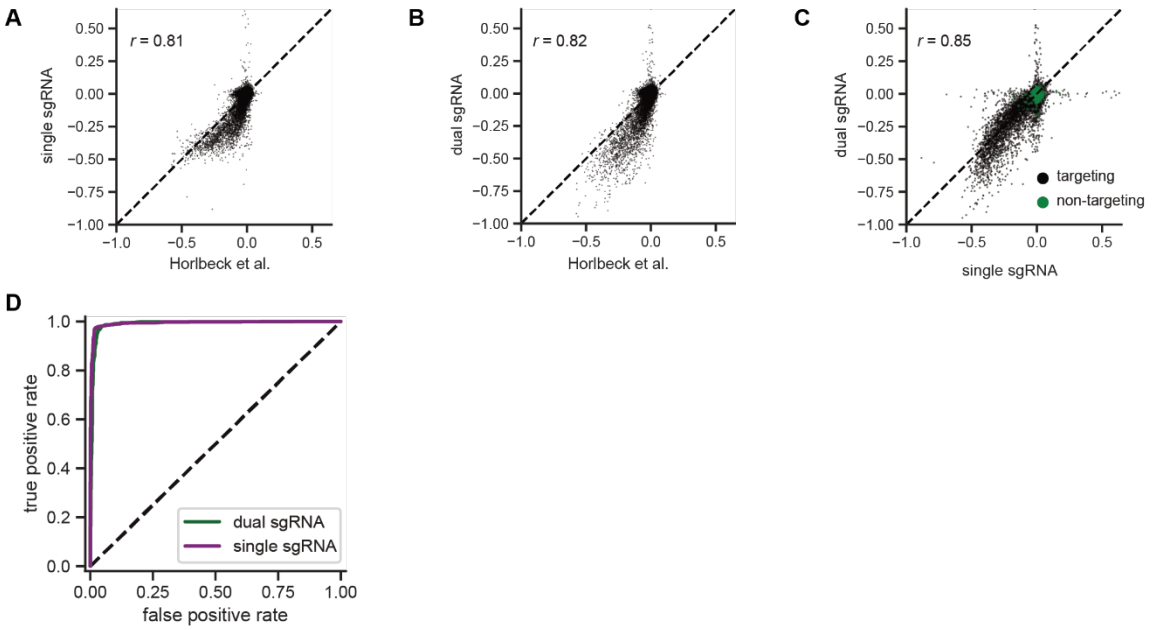


Figure S1. Additional comparisons of pilot single- and dual-sgRNA library screens.

a. Comparison of growth phenotypes for all elements between our pilot single-sgRNA library and Horlbeck et. al data, merged by gene name (n=20036 elements). Growth phenotypes are reported as γ ($\log_2$ fold-enrichment of T_{final} over T_0, per doubling) and correlated between experiments (r= 0.81).

b. Comparison of growth phenotypes for all elements between our pilot dual-sgRNA library and Horlbeck et. al data, merged by gene name (n=20036 elements). Growth phenotypes are reported as γ and correlated between experiments (r= 0.82).

c. Comparison of growth phenotypes for all elements between our pilot single- and dual-sgRNA libraries, merged by gene name (n=21049 with 20036 targeting elements and 1013 non-targeting elements). Growth phenotypes are reported as γ and correlated between experiments (r= 0.85).

d. Comparison of true and false positive rates in single element screens. "Positives" (n=1239 elements) were defined as genes with a K562 CRISPRi growth screen p-value < 0.001 and γ < -0.05 (Horlbeck et al., 2016a), and "negatives" were defined as non-targeting control sgRNA pairs (n=1013 elements).

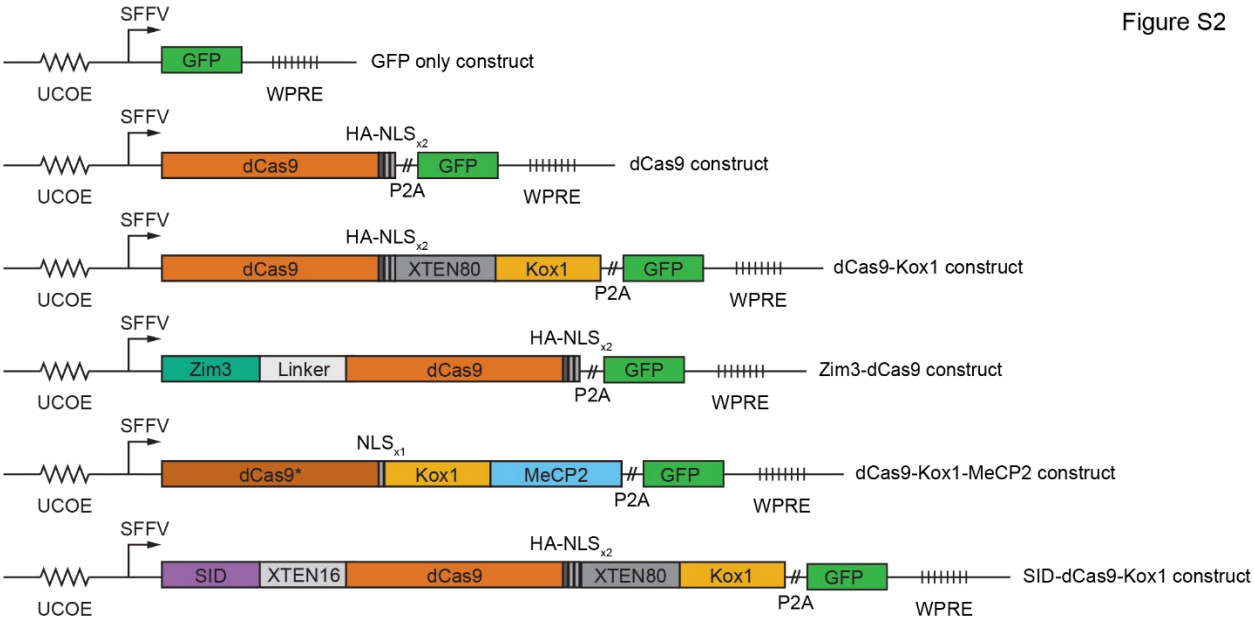


Figure S2. Design of constructs for CRISPRi effector expression.

Figure S3

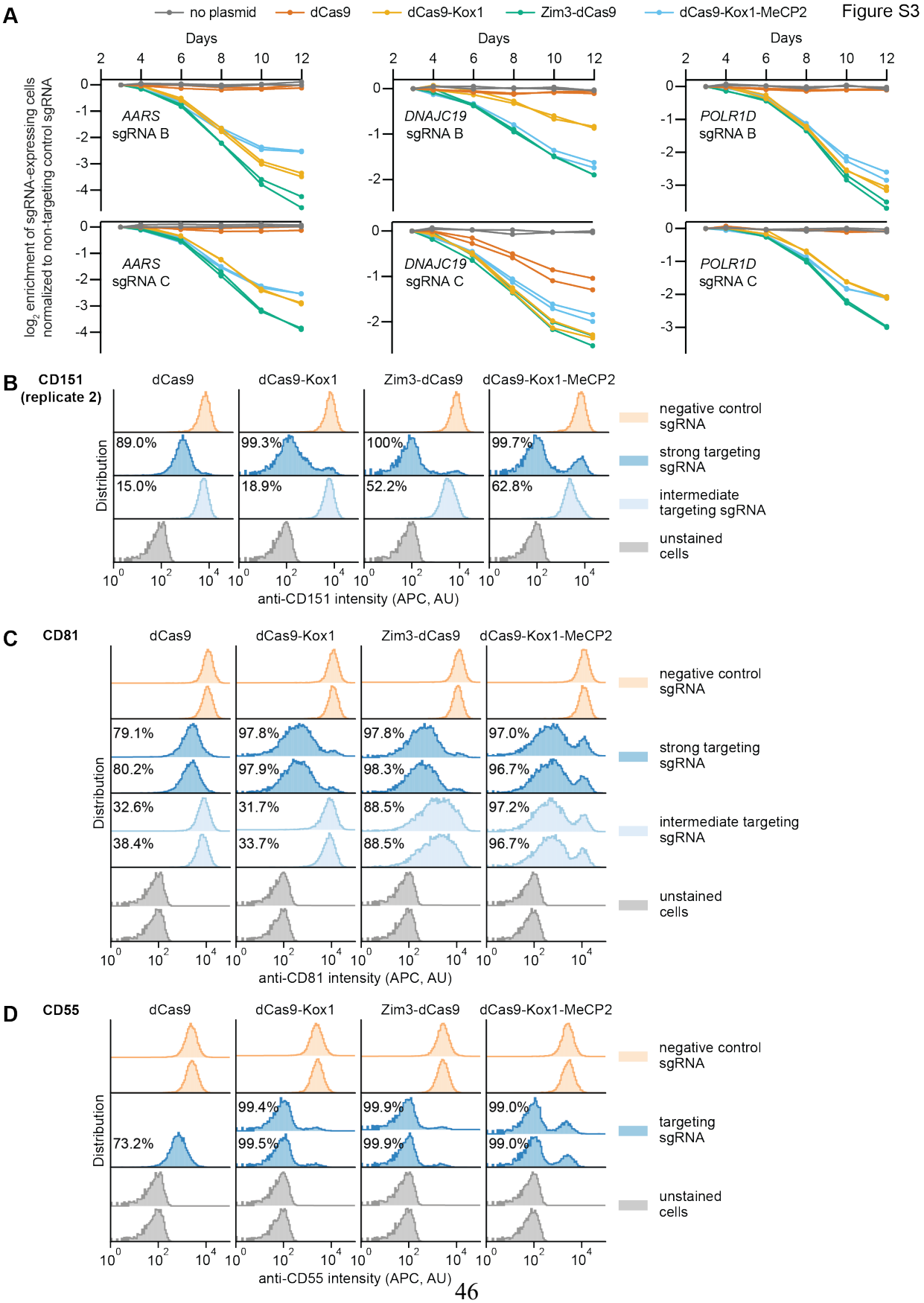


Figure S3. Additional measurements of on-target activity of CRISPRi effectors.

a. Depletion of K562 cells expressing essential gene-targeting sgRNAs and different CRISPRi effectors, measured as the ratio of mCherry-positive (sgRNA-expressing) to mCherry-negative (not sgRNA-expressing) cells in a given well, as in Figure 3a. mCherry levels were measured for 12 days after transduction, starting on day 3. Data from two replicate transductions.

b. Distribution of anti-CD151 signal intensity (APC) in K562 cells expressing indicated CRISPRi effectors from second replicate transduction. Knockdown was quantified as in Figure 3c.

c. Distribution of anti-CD81 signal intensity (APC) in K562 cells expressing indicated CRISPRi effectors from two replicate transductions. Knockdown was quantified as in Figure 3c.

d. Distribution of anti-CD55 signal intensity (APC) in K562 cells expressing indicated CRISPRi effectors from two replicate transductions. Cells expressing dCas9 and the CD55-targeting sgRNA are represented by a single replicate. Knockdown was quantified as in Figure 3c.

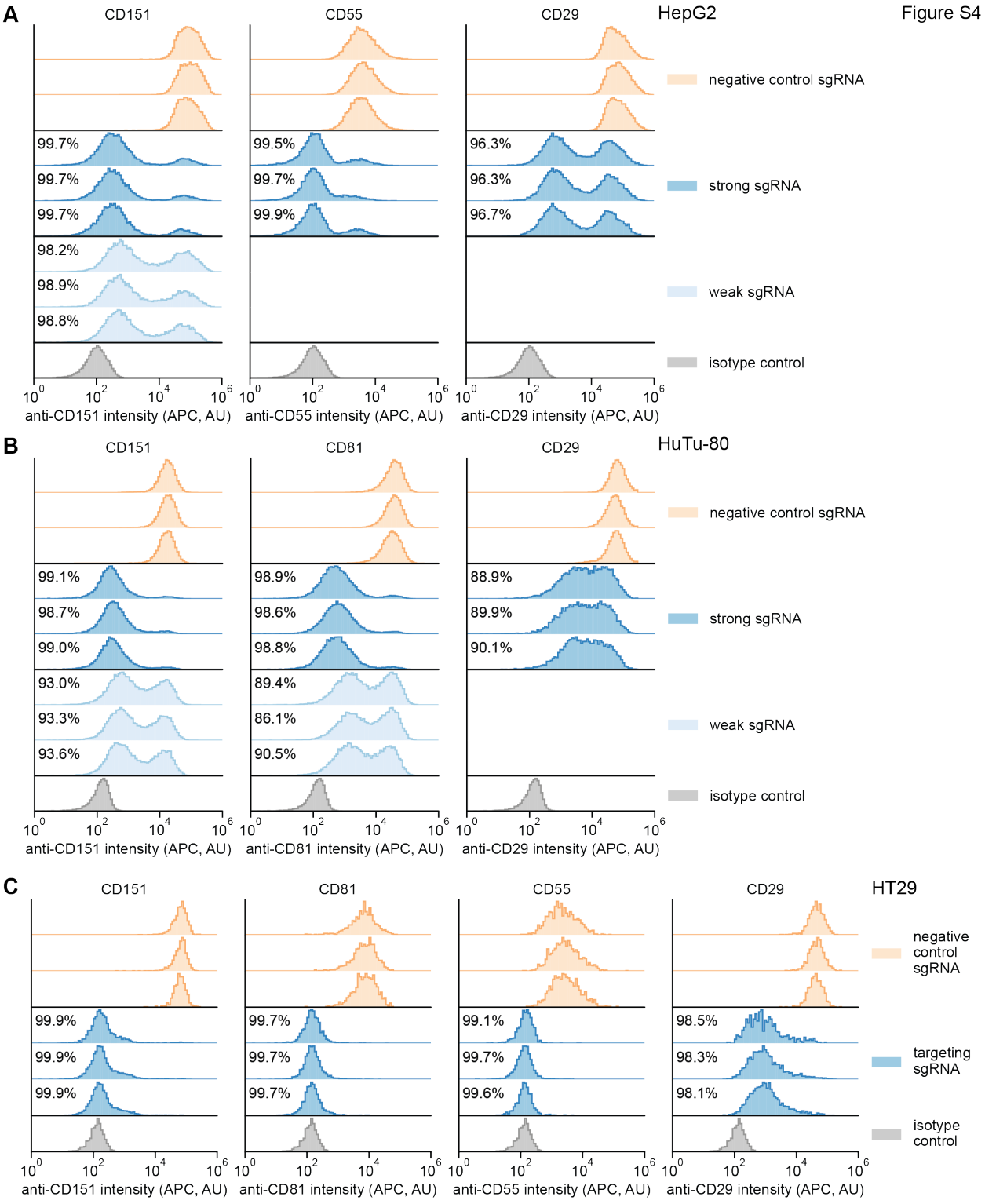


Figure S4. Single-cell distributions of knockdown in different Zim3-dCas9 cell lines.

a. Distribution of anti-CD151, anti-CD55, and anti-CD29 signal intensities (APC) in HepG2 cells expressing Zim3-dCas9. Data from 3 independent transductions are shown. A weak

targeting sgRNA was only included for CD151. For the isotype control, cells expressing the negative control sgRNA were stained with an APC-conjugated isotype control antibody. A single replicate is shown for the isotype control. Knockdown was calculated as in Figure 4b.

b. Distribution of anti-CD151, anti-CD81, and anti-CD29 signal intensities (APC) in HuTu-80 cells expressing Zim3-dCas9. Data from 3 independent transductions are shown. A weak targeting sgRNA was only included for CD151 and CD81. For the isotype control, cells expressing the negative control sgRNA were stained with an APC-conjugated isotype control antibody. A single replicate is shown for the isotype control. Knockdown was calculated as in Figure 4b.

c. Distribution of anti-CD151, anti-CD81, anti-CD55, and anti-CD29 signal intensities (APC) in HT29 cells expressing Zim3-dCas9. Data from 3 independent transductions are shown. Only strong targeting sgRNAs were included. For the isotype control, cells expressing the negative control sgRNA were stained with an APC-conjugated isotype control antibody. A single replicate is shown for the isotype control. Knockdown was calculated as in Figure 4b.

Supplementary Tables

Table S1. Dual- and single-sgRNA libraries used for preliminary comparison.

Table S2. Aggregated CRISPRi sgRNA performance across screens.

Table S3. Finalized dual-sgRNA CRISPRi libraries.

Table S4. List of integration barcodes.

Table S5. Dolcetto versus dual-sgRNA Perturb-seq comparison.

Table S6. Description of plasmids.

Table S7. Sequences of sgRNAs used for individual validation.

Table S8. Dual-sgRNA CRISPRa libraries.

Supplementary Notes

Supplementary Note 1. Protocol for cloning dual-sgRNA libraries.

Supplementary Note 2. Protocol for sample preparation and Illumina sequencing of dual-sgRNA libraries.

Supplementary Note 3. Protocol for arrayed cloning of dual-sgRNA constructs.

Supplementary Note 4. Protocol for generation of CRISPRi cell lines.