Higher-order combinatorial chromatin perturbations by engineered CRISPR-Cas12a for functional genomics

Hsiung CC[1,2,4,9,10], Wilson CM[2,3,4,9,10], Sambold NA[9], Dai R[2,4,5,9,10], Chen Q[7], Misiukiewicz S[2,4,5], Arab A[10], Teyssier N[6], O'Loughlin T[2,4,9], Cofsky JC[8], Shi J[7], Gilbert LA[#,2,4,9,10]

Author affiliations

[1]Dept. of Pathology, Stanford University School of Medicine, Stanford, CA 94305, USA
[2]Dept. of Urology, University of California, San Francisco, CA 94158, USA
[3]Tetrad Graduate Program, University of California, San Francisco, CA 94158, USA
[4]Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, San Francisco, CA 94158
[5]Biomedical Sciences Graduate Program, University of California, San Francisco, San Francisco, CA 94158, USA
[6]Biological and Medical Informatics Graduate Program, University of California, San Francisco, San Francisco, CA 94158, USA
[7]Department of Cancer Biology, University of Pennsylvania, Philadelphia, PA, 19104, USA
[8]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA, 02115, USA
[9]Innovative Genomics Institute, University of California, San Francisco, San Francisco, CA 94158, USA
[10]Arc Institute, Palo Alto, CA 94304, USA

[#]Correspondence to luke@arcinstitute.org

Abstract

Multiplexed genetic perturbations are critical for testing functional interactions among coding or non-coding genetic elements. Compared to DNA cutting, repressive chromatin formation using CRISPR interference (CRISPRi) avoids genotoxicity and is more effective for perturbing non-coding regulatory elements in pooled assays. However, current CRISPRi pooled screening approaches are generally limited to targeting 1-3 genomic sites per cell. To develop a tool for higher-order (≥3) combinatorial targeting of genomic sites with CRISPRi in functional genomics screens, we engineered an *Acidaminococcus* Cas12a variant -- referred to as multiplexed transcriptional interference AsCas12a (multiAsCas12a). multiAsCas12a significantly outperforms state-of-the-art Cas12a variants in combinatorial CRISPRi targeting using high-order multiplexed arrays of CRISPR RNAs (crRNA) delivered by lentiviral transduction,

including in high-throughput pooled screens using 6-plex crRNA array libraries. Using multiAsCas12a CRISPRi, we discover new enhancer elements and dissect the combinatorial function of cis-regulatory elements. These results demonstrate that multiAsCas12a enables group testing strategies to efficiently survey potentially numerous combinations of chromatin perturbations for biological discovery and engineering.

Introduction

Functional interactions among combinations of genetic elements underlie many natural and engineered phenotypes (Wong et al. 2016; Costanzo et al. 2019; Domingo et al. 2019). Such interactions can often involve higher-order (3 or more) combinations of genetic elements. A notable gain-of-function example includes the discovery of 4 factors as the successful combination that can achieve reprogramming to pluripotency (Takahashi and Yamanaka 2006). In loss-of-function experiments, discovering the functions of paralogous genes often requires combined perturbations to overcome functional redundancy (Parrish et al. 2021; Dede et al. 2020; Gonatopoulos-Pournatzis et al. 2020; Ewen-Campen et al. 2017). In the non-coding space, multiple cis-regulatory elements often co-regulate the transcriptional state of a given genomic locus with varying degrees of functional overlap, often requiring higher-order combinatorial perturbations to test their functional logic (Osterwalder et al. 2018; Xie et al. 2017; Kvon et al. 2021; Blayney et al. 2022; Blobel et al. 2021). Thus, combined perturbation of $\geq$3 coding or non-coding genetic elements can be critical for discovering and engineering biological properties, otherwise unattainable by lower-order perturbations.

Despite the value of higher-order genetic perturbations across biological contexts, such perturbations have been generally difficult to achieve in a scalable manner, with prior systematic analyses primarily achieved in yeast (Kuzmin et al. 2018; Domingo et al. 2018; Taylor and Ehrenreich 2014, 2015; Celaj et al. 2020). In mammalian functional genomics, pooled CRISPR screens are currently the most widely adopted and scalable approaches (Doench 2018; Przybyla and Gilbert 2022). Such screens use pooled oligo synthesis for one-pot cloning of a complex guide RNA library, which is delivered to a population of cells such that each cell generally receives a unique guide RNA construct. The phenotypes induced by each guide RNA construct are measured in a massively parallel fashion across the cell population by deep sequencing in a single biological sample (Doench 2018; Przybyla and Gilbert 2022). However, CRISPR/Cas9-based pooled screens using sequencing readouts have thus far been limited in multiplexing capability, with only a few studies targeting 3 genomic sites per cell (Zhou et al. 2020; Wong et al. 2015; Adamson et al. 2016). Further multiplexing using Cas9-based pooled screening is challenging due to 1) increasingly complex iterative cloning schemes for

larger constructs encoding for multiple sgRNAs each expressed from a separate promoter ((Zhou et al. 2020; Wong et al. 2015; Adamson et al. 2016)), and 2) length-dependent high frequencies of recombination in sgRNA libraries that are typically delivered as lentiviral constructs (Adamson et al. 2018; Sack et al. 2016; Basu et al. 2008). Moreover, conceptually, it remains unclear how to tractably survey the potentially vast combinatorial spaces for ≥3-plex perturbations.

Cas12a, a member of the type V CRISPR-Cas family, has been proposed as an alternative to (d)Cas9 for genetic perturbations due to enhanced multiplexing capabilities. Cas12a harbors RNase activity, separable from its DNase activity, that can process a compact primary transcript expressed from a single promoter into multiple CRISPR RNAs (crRNA), without the need for tracrRNA (Fonfara et al. 2016; Zetsche et al. 2015). The compactness of the Cas12a crRNA, individually consisting of a 19nt direct repeat and a 19-23nt spacer, enables deterministic encoding of multiple crRNAs on a given chemically synthesized oligo for single-step cloning into the plasmid vector, expressed from a single promoter (Zetsche et al. 2017; Breinig et al. 2019; Campa et al. 2019; DeWeirdt et al. 2021; Gier et al. 2020). Cas12a has been engineered for mammalian cell applications using its DNase activity to disrupt coding gene function using single or multiplexed crRNA constructs in individual well-based assays (Breinig et al. 2019; DeWeirdt et al. 2022; Campa et al. 2019; Zetsche et al. 2017; Kleinstiver et al. 2019; Gonatopoulos-Pournatzis et al. 2020; Zetsche et al. 2015) and in pooled sequencing screens (Gier et al. 2020; DeWeirdt et al. 2021; Chow et al. 2019; Liu et al. 2019; Dede et al. 2020; Gonatopoulos-Pournatzis et al. 2020; Esmaeili Anvar et al. 2023). However, extended multiplexing with DNase-competent Cas12a is expected to be constrained by increasing genotoxicity in many biological contexts. Even cancer cell lines can show cumulative genotoxicity with multi-site double-stranded DNA breaks (DeWeirdt et al. 2021; Meyers et al. 2017; Aguirre et al. 2016), and non-transformed cell types can be sensitive to genotoxicity from nuclease targeting of as low as one or two sites per cell (Chen et al. 2021; Ihry et al. 2018; Bowden et al. 2020; Haapaniemi et al. 2018). In principle, perturbations of genetic elements can avoid genotoxicity by using DNase-dead Cas enzymes fused to effector domains that alter the chromatin state of the targeted site, as has been successfully achieved by synthetic transcriptional repression (CRISPRi) and activation (CRISPRa) by DNase-dead Cas9 (dCas9) fusion proteins in pooled sequencing screens (Gilbert et al. 2013, 2014; Konermann et al. 2015). Another advantage of CRISPRi over DNase-based screens is that CRISPRi is more efficient at perturbing enhancers in pooled screens (Ren et al. 2021; Tycko et al. 2019), likely due to its larger genomic window of activity via formation of repressive chromatin (Ecco et al. 2017). Thus, a DNase-dead Cas12a (dCas12a) functional genomics platform capable of co-targeting 3 or more genomic sites per cell for CRISPRi

would be highly desirable as a general-purpose tool for testing the combinatorial functions of coding and non-coding genetic elements.

Despite the success of Cas12a as a tool for DNA cleavage, no dCas12a-based pooled CRISPRi screening platforms have been reported thus far. Several studies have used dCas12a for CRISPRi in human cells in individual well-based assays, reporting either successful (Campa et al. 2019; Guo et al. 2022; Liu et al. 2017; Nuñez et al. 2021) or unsuccessful (O'Geen et al. 2017) repression of target genes. These dCas12a transcriptional repression studies delivered crRNA plasmids exclusively by transient transfection, which introduces high copy number and expression of synthetic components, but is limited in assay throughput and cell type compatibility. Outcomes from transient plasmid transfection experiments are often poor surrogates for success in pooled screens using sequencing readouts (Stegmeier et al. 2005; DeWeirdt et al. 2021), likely because transient transfections can often result in 10 to 1000-fold higher intracellular concentrations of the delivered components than in pooled screens. Pooled screens require single-copy integration of crRNA constructs by lentiviral transduction at low multiplicity of infection (MOI) to ensure that cellular phenotypes can be attributed to unique crRNA constructs (Doench 2018; Przybyla and Gilbert 2022), imposing a stringent requirement on the functional potency of each ribonucleoprotein molecule. Prior published studies reporting use of dCas12a-based for transcriptional repression (Campa et al. 2019; Guo et al. 2022; Liu et al. 2017; Nuñez et al. 2021) did not use lentiviral delivery of crRNA constructs, raising the question of whether published dCas12a CRISPRi constructs are sufficiently potent for high-throughput pooled screens using sequencing readouts.

In this study, we show that existing dCas12a CRISPRi fusion constructs function poorly when used with limiting doses of lentivirally delivered components, thus precluding their application in pooled functional genomics studies. This deficiency is generally exacerbated when using multiplexed crRNA arrays, which effectively further dilute the concentration of Cas12a fusion protein molecules available to bind each individual crRNA encoded in the array. Given the stark discrepancy in activity between dCas12a CRISPRi and Cas12a DNA cutting applications, we reasoned that DNase-inactivation in dCas12a may render fusion proteins ineffective for transcriptional repression, likely by destabilizing its chromatin occupancy. To overcome this, we engineered a new Cas12a variant -- referred to as multiAsCas12a (multiplexed transcriptional interference AsCas12a) -- that incorporates a key mutation, R1226A. Prior in vitro experiments demonstrated that this mutation slows DNA cutting to a regime that favors predominantly nicked DNA products and results in a more stable ribonucleoprotein:DNA complex (Cofsky et al. 2020), but has not been tested in any context related to transcriptional control in prior literature. We show that in human cells, R1226A

significantly enhances CRISPRi activity in the setting of lentivirally delivered crRNA constructs and enables multiplexed transcriptional repression using up to 10-plex arrays of crRNAs. We demonstrate that multiAsCas12a critically enables high-throughput pooled CRISPRi screens, including using 6-plex crRNA arrays. We use multiAsCas12a CRISPRi to efficiently discover new enhancer elements and to test higher-order combinatorial perturbations of cis-regulatory elements that co-regulate a locus. We propose a generalizable group testing framework using multiAsCas12a to realize exponential efficiency gains in searching through potentially large combinatorial spaces of chromatin perturbations.

Results

CRISPRi using state-of-the-art dAsCas12a fusion proteins is dose-limited and hypoactive in the setting of lentivirally delivered components

While several orthologs of Cas12a have been used for mammalian cell applications, here we focus on building a CRISPRi functional genomics platform using *Acidaminococcus* Cas12a (AsCas12a) as it is the only ortholog with demonstrated success in the published literature in DNA cutting-based pooled sequencing screens in mammalian cells (Gier et al. 2020; DeWeirdt et al. 2021; Chow et al. 2019; Liu et al. 2019; Dede et al. 2020; Gonatopoulos-Pournatzis et al. 2020; Esmaeili Anvar et al. 2023). A previous study reported using dAsCas12a for CRISPRi by plasmid transient transfection delivery of dAsCas12a-KRABx3 protein (harboring the E993A DNase-dead mutation) and crRNA in HEK 293T cells (Campa et al. 2019). To test this construct in the setting of lentivirally delivered crRNA, we introduced dAsCas12a-KRABx3 by piggyBac transposition in K562 cells and sorted for a pool of cells stably expressing the construct, as monitored by a P2A-BFP marker. We designed single crRNA constructs targeting TTTV protospacer adjacent motifs (PAM) proximal to transcription start sites (TSS) of endogenous genes encoding cell surface proteins, whose knockdown by dCas9-KRAB has been previously successful and is fitness-neutral (Replogle et al. 2022a). Throughout this study we encoded crRNAs in a previously optimized CROP-seq (Datlinger et al. 2017) style lentiviral vector containing a U6 promoter that transcribes a pre-crRNA with a 3' direct repeat that enhances crRNA activity (Gier et al. 2020). After lentiviral transduction of single-plex crRNA constructs into K562 cells stably expressing dAsCas12a-KRABx3 (Fig. 1A), we measured target gene knockdown by flow cytometry using cell surface antibody staining (Fig. 1B) in cells gated for successful crRNA transduction (Fig. S1). We observed no expression change in any of the individually targeted genes (CD55, CD81, B2M and KIT in Fig. 1B and Fig. S2). We confirmed expression of dAsCas12a-KRABx3 by western blot (Fig. S3) and by routinely monitoring expression of the in-frame P2A-BFP transgene marker by flow cytometry.

This lack of CRISPRi activity when using lentivirally transduced crRNAs is not limited to K562 cells, as the absence was similarly observed in C4-2B cells (prostate cancer cell line) stably engineered with dAsCas12a-KRABx3 by piggyBac transposition (Fig. S4). In contrast, transient co-transfection of dAsCas12a-KRABx3 and CD55-targeting crRNA plasmids shows modest CRISPRi knockdown in HEK 293T cells (Fig. S5), consistent with prior work (Campa et al. 2019). These findings indicate that the requirements for successful transcriptional repression using dAsCas12a-KRABx3 in the setting of lentivirally delivered crRNA constructs are distinct from those of plasmid transient transfection delivery in HEK 293T cells used in prior studies reporting successful Cas12a CRISPRi at a few loci (Campa et al. 2019).

In an attempt to overcome this lack of CRISPRi activity, we tested combinations of several Cas12a mutations representing state-of-the-art optimizations of Cas12a from the literature. These include: 1) E174R/S542R/K548R (enhanced AsCas12a, enAsCas12a), which are expected to contact PAM proximal DNA and have been shown to improve DNA cutting in human cells (Kleinstiver et al. 2019); 2) M537R/F870L (AsCas12a ultra), which interact with the PAM (M537R) and the crRNA stem loop (F870L), and improve DNA cutting in human cells (Liyang Zhang et al. 2021); 3) W382A, a mutation that reduces R-loop dissociation in vitro for an orthologous enzyme (LbCas12a W355A), but has not yet been tested in cells (Naqvi et al. 2022).

We generated six dAsCas12a variants that each harbor the DNase-inactivating D908A mutation, plus a select combination of the mutations described above. We cloned these variants into the same fusion protein architecture (Fig. 1C) consisting of an N-terminal 6x Myc-NLS (Gier et al. 2020) and C-terminal XTEN80-KRAB-P2A-BFP (Replogle et al. 2022) in a lentiviral expression vector. This allows us to monitor fluorescence from the P2A-BFP reporter by flow cytometry as a quantitative proxy of MOI and transgenic expression. We tested these constructs for CRISPRi activity by stable lentiviral expression of each dCas12a variant fusion construct and a crRNA construct targeting the TSS of either CD55 and CD81 in K562 cells (Fig. 1C and Fig. S6). Among this panel, denAsCas12a-KRAB (E174R/S542R/K548R, plus D908A DNase-dead mutation) performed the best and demonstrated strong repression of CD55. However, even for this best construct we observed weak repression of CD81, indicating inconsistent performance across crRNAs (Fig. 1C and Fig. S6).

Dose-response and construct potency are key considerations for multiplexed applications, as increased multiplexing effectively reduces the concentration of Cas protein available to bind each individual crRNA. Focusing on denAsCas12a-KRAB as the top variant, we tested the effect of separately altering the dosage of Cas12a protein and crRNAs delivered. We found that increasing the MOI of the denAsCas12a-KRAB

construct from ~1 to ~5 can improve CRISPRi knockdown of CD81 for a single crRNA targeting CD81 (crCD81-1) and when encoded in the context of a 3-plex crRNA array in the 3' position (crCD55-4_crCD151-3_crCD81-1), but still at a suboptimal level (~60% median expression knockdown relative to non-targeting control, Fig. 1D and Fig. S7A). Even at high protein construct MOI of ~5, we found that CRISPRi activity of denAsCas12a-KRAB is significantly lost when the crRNA MOI is reduced to <1 to mimic that required to ensure single-copy integrations in pooled screens using sequencing readouts (Fig. 1E and Fig. S7B). Even more problematically, across all protein (Fig. 1D and Fig. S7A) and crRNA (Fig. 1E and Fig. S7B) doses tested, a 3-plex crRNA in the reverse orientation (CD81-1_CD151-3_CD55-4) shows extremely weak CD81 knockdown (~0%-25% median expression knockdown relative to non-targeting control), indicating that denAsCas12a-KRAB CRISPRi activity can be sensitive to the specific arrangement of crRNAs encoded in a multiplexed array.

Given the inconsistent and deficient performance of denAsCas12a-KRAB, we tested an alternative CRISPRi approach without mutating the RuvC DNase domain. In the setting of transient plasmid transfection delivery in HEK 293T cells, wild-type AsCas12a protein without any engineered RuvC DNase domain mutations has been used for transcriptional control with truncated (15nt) crRNA spacers, which enable DNA binding but not cleavage (Campa et al. 2019; Breinig et al. 2019). We tested this approach by fusing KRAB or KRABx3 in different N- and C-terminal arrangements to opAsCas12a (containing 6x MycNLS and DNA affinity-enhancing mutations E174R/S542R), a DNase-active Cas12a optimized for pooled screens (Gier et al. 2020). Confirming previous findings, we showed that the 15nt spacers did not support DNA cleavage, while the 23nt spacers did (Fig. S8). However, using 15nt spacers, we observed weak or no cell fitness phenotype as a proxy of CRISPRi activity when targeting the transcriptional start site of a common essential gene, Rpa3, in two cell lines engineered with a panel of opAsCas12a fused to KRAB or 3xKRAB (Fig. S8). In total, we have tested 3 separate approaches that abolish the DNase activity of AsCas12a: 1) E993A in dAsCas12a-KRABx3, 2) D908A in denAsCas12a-KRAB, and 3) use of truncated spacers with opAsCas12a fused to KRAB or 3xKRAB. All of these CRISPRi approaches, despite incorporating state-of-the-art optimizations, perform poorly when used with lentivirally transduced crRNA constructs. These results collectively suggest additional optimization of construct potency is required for the goal of developing a potent and predictable Cas12-based CRISPRi functional genomics platform.

multiAsCas12a-KRAB (R1226A/E174R/S542R/K548R), a variant that favors a nicked DNA intermediate, substantially improves lentivirally delivered CRISPRi

The mediocre performance of dAsCas12a for CRISPRi surprised us given the successful application of Cas12a in DNA-cutting pooled screens. We wondered whether full inactivation of DNA cutting in dCas12a may render transcriptional repression ineffective by adversely impacting other aspects of protein function important for CRISPRi activity, specifically chromatin occupancy. Previous studies indicate that the interaction between Cas12a and a DNA target can be strengthened by DNA cleavage (Singh et al. 2018; Knott et al. 2019; Cofsky et al. 2020). In the Cas12a DNA cleavage process, a single DNase active site first cuts the non-target strand, followed by cleavage of the target strand (Swarts and Jinek 2019). While double-strand breaks are undesired for CRISPRi applications, we wondered whether favoring the intermediate nicked DNA state might reduce the R-loop dissociation rate (Fig. 2A, see Discussion). In support of this possibility, in vitro binding assays showed that dCas12a:DNA complexes are 20-fold more stable when the non-target strand is pre-cleaved (Cofsky et al. 2020), and single-molecule Förster resonance energy transfer studies suggest that non-target strand nicking biases Cas12a:DNA complexes away from dissociation-prone conformations (Zhang et al. 2019; Jeon et al. 2018).

To engineer nicking-induced stabilization of Cas12a binding to DNA for CRISPRi applications, we incorporated R1226A, a mutation that has not been tested in the context of transcriptional control. Relative to WT AsCas12a, the AsCas12a R1226A mutant protein, described as a nickase in its original characterization (Yamano et al. 2016), is ~100-1,000 fold slower in cleaving the non-target DNA strand and ~10,000-fold slower in cleaving the target DNA strand in vitro (Cofsky et al. 2020). Consistent with nicking-induced stabilization, AsCas12a R1226A indeed binds DNA more strongly in vitro than the fully DNase-inactivated D908A variant (Cofsky et al. 2020). We expect the R1226A mutation to both disfavor R-loop reversal and slow progression to double-stranded breaks (Fig 2A; see Discussion). We hypothesized that, by trapping the complex in a nicked-DNA intermediate state, the R1226A mutation would prolong chromatin occupancy and thus the time available for the KRAB domain to recruit transcriptional repressive complexes.

To test the impact of R1226A on CRISPRi activity, we replaced the DNase-inactivating D908A in denAsCas12a-KRAB with R1226A, and hereafter refer to this Cas12a variant as multiAsCas12a (multiplexed transcriptional interference, i.e. R1226A/E174R/S542R/K548R). To test their CRISPRi performance at different protein doses, we lentivirally transduced denAsCas12a-KRAB and multiAsCas12a-KRAB constructs at MOI = ~1 vs. MOI = ~5 in K562 cells and sorted for stably expressing BFP-positive cell population. Using lentivirally delivered crRNAs targeting CD81 and CD55 as single crRNAs and as part of 3-plex crRNAs, we compared the CRISPRi performance of multiAsCas12a-KRAB vs. denAsCas12a-KRAB across different

combinations of protein MOI and crRNA MOIs. Across a panel of single and 3-plex crRNA constructs, we found that multiAsCas12a-KRAB consistently achieves robust CRISPRi with less sensitivity to low protein MOI, and with especially large improvements over denAsCas12a-KRAB in the setting of low crRNA MOI (Fig. 2B and Fig. S9) and especially for 3-plex crRNAs (Fig. 2C and Fig. S10). As a notable example, the 3-plex crRNA (crCD81-1_crCD151-3_crCD55-4) that is virtually inactive for CD81 knockdown by denAsCas12a-KRAB even in the setting of high protein MOI ~5 shows >95% median CD81 expression knockdown by multiAsCas12a-KRAB (Fig. 2C). This same 3-plex crRNA construct shows double knockdown of CD55 and CD81 in only 14.3% of single cells for denAsCas12a-KRAB vs. 76.7% for multiAsCas12a-KRAB (Fig. 2D). Similarly, multiAsCas12a-KRAB is able to rescue the CRISPRi activity of single and 3-plex crRNA constructs targeting CD151 that are otherwise completely inactive when used with denAsCas12a-KRAB (Fig. 2E). multiAsCas12a-KRAB CRISPRi activity shows generally minimal or no off-target effects on the transcriptome as evaluated by bulk RNA-seq (Fig. S11).

We characterized the impact on DNA sequence of long-term constitutive targeting of multiAsCas12a-KRAB to genomic sites in K562 cells. Using K562 cells lentivirally engineered with multiAsCas12a-KRAB or denAsCas12a-KRAB at protein MOI = ~5, we tested for the DNA sequence alterations at CRISPRi target sites near the CD55 and CD81 TSS's after 20 days of constitutive ribonucleoprotein expression. Any DNA sequence alterations are expected to accumulate over this duration, which is representative of the upper end of duration for typical CRISPR functional genomics experiments. We reasoned that testing a highly effective CRISPRi crRNA (crCD55-4) would represent an experimental condition of high target occupancy over a long time period, enabling us to estimate the upper end of indel formation. In this experiment, we observed indel frequencies of 7.9% by multiAsCas12a-KRAB, 0.1% by denAsCas12a-KRAB, and 97.9% by a fully DNase-active construct, opAsCas12a (Gier et al. 2020) (Fig. 2F). For crCD81-1, a crRNA with an intermediate level of CRISPRi activity more representative of most crRNAs we have tested, we observed indel frequencies of 2.48% by multiAsCas12a-KRAB, 1.25% by denAsCas12a, and 71.5% by opAsCas12a (Fig. 2F). If indels were the sole driver of target gene knockdown, and further assuming an indel of any size generates a complete null CD81 allele, in this triploid region of the K562 genome we would expect ~89.5% of cells to harbor zero indels and retain full expression level, at most ~10.1% of cells to harbor a deletion of any size in one DNA copy and thus retain ~67% of CD81 expression level, ~0.38% of cells to inactivate two DNA copies, and virtually no cells to inactivate all 3 DNA copies (see Methods). The expected knockdown in median CD81 expression in the cell population under this adversarial assumption of abolishing expression in *cis* by any sized deletion would be 3.6% (see Methods). This expectation is wholly inconsistent

with the magnitudes of median CD81 expression knockdown by multiAsCas12a-KRAB, including up to ~95% knockdown in excess of denAsCas12a-KRAB for crCD81-1_crCD151-3_crCD55-4 (Fig. 2C, far right subpanel). This is the first of multiple lines of evidence, further addressed in subsequent sections, demonstrating that the magnitude of target gene knockdown by multiAsCas12a-KRAB is far from being accounted for by DNA sequence alterations alone.

multiAsCas12a-KRAB enables multi-gene CRISPRi perturbations using higher-order arrayed crRNA lentiviral constructs

To test the performance of multiAsCas12a-KRAB in targeting >3 genomic sites per cell for CRISPRi, we designed a lentiviral system for expressing higher-order multiplexed crRNA arrays, keeping the overall U6 promoter and CROP-seq vector design with a 3' direct repeat (Gier et al. 2020). To minimize the possibility of lentiviral recombination, this system (Fig. 3A) uses a unique direct repeat variant at each position of the array, selected from a set of previously engineered direct repeat variants (DeWeirdt et al. 2020). Using this lentiviral expression system, we assembled a panel of 13 distinct crRNA constructs (7 single-plex, two 3-plex, two 4-plex, two 5-plex, and two 6-plex), with the higher-order crRNA arrays assembled from individually active spacers targeting the TSS's of CD55, CD81, B2M and KIT (Fig. 3B-C and Fig. S12). For this panel of 13 crRNA constructs, we compared the CRISPRi activities of denAsCas12a-KRAB, multiAsCas12a-KRAB, and multiAsCas12a (no KRAB) as a negative control for the impact of the KRAB domain. For a subset of crRNA constructs we also added enAsCas12a-KRAB (DNase fully active) to test the effect of fully active DNA cutting in conjunction with the KRAB domain in target gene knockdown. Unless otherwise specified, for these experiments and the remainder of this study we use piggyBac transposition to constitutively express all fusion protein constructs, which yields results similar to that obtained from high MOI (~5) lentiviral delivery of protein constructs and avoids day-to-day variations in lentiviral titers. Each piggyBac-delivered construct is expressed in K562 cells at very similar protein levels as measured by western blot (Fig. S3) and routine flow cytometry monitoring of the P2A-BFP fluorescence signal (Fig. S13).

To summarize results of the full panel of crRNA constructs, multiAsCas12a-KRAB substantially outperforms denAsCas12a-KRAB in CRISPRi activity for 7 out of 7 constructs tested for CD81 knockdown (Fig. 3B); 4 out of 6 constructs tested for B2M knockdown (Fig. S12); and 6 out of 6 constructs tested for KIT knockdown (Fig. 3C). For CD55 (Fig. S12), multiCas12a-KRAB substantially outperforms denAsCas12a-KRAB for the single-plex crCD55-5 (weaker spacer), and performs either the same as or marginally better than denAsCas12a-KRAB for all 7 constructs containing crCD55-4

(strongest spacer). Similarly superior CRISPRi performance by multiAsCas12a-KRAB over denAsCas12a-KRAB was also observed when using up to 6-plex crRNA arrays in a different cell type, C4-2B prostate cancer cells (Fig. S4).

For all crRNA constructs tested, multiAsCas12a alone shows much lower impact on target gene expression than multiAsCas12a-KRAB (e.g. Fig. 3B for CD81), demonstrating that the large improvements in gene knockdown by multiAsCas12a-KRAB depends on the KRAB domain. For some target genes, such as KIT, partial knockdown can be observed for multiAsCas12a alone (Fig. 3C). Such gene knockdown may be due to 1) direct obstruction of the transcriptional machinery, or 2) alteration of DNA sequences crucial for transcription via double-stranded break formation and repair. To distinguish these possibilities, we quantified indels generated by the panel of fusion proteins using a 6-plex crRNA array (crCD55-4_crB2M-1_crB2M-3_crKIT-2_crKIT-3_crCD81-1) containing two crRNAs targeting opposite strands at sites 95bp apart near the KIT TSS (Fig. 3D and Fig. S14A). This genomic distance between two crRNA binding sites is known to optimally facilitate deletions of the intervening region by DNA cutting Cas proteins (Joberty et al. 2020), thus representing an upper estimate of the frequencies of deleting intervening regions of multiple target sites in *cis*. The maximum indel frequencies measured for any given base were only 3%-5.4% for multiAsCas12a and 1.2%-3.7% for multiAsCas12a-KRAB at each individual crRNA binding site (Fig. 3D, region A and region C). In the intervening region between the two crRNA binding sites, both multiAsCas12a and multiAsCas12a-KRAB generated only a maximum of 0.2% indels (Fig. 3D, region B). In comparison, fully DNase-active enAsCas12a-KRAB generates up to ~94.5% indels at each individual crRNA binding site and up to 7.6% in the intervening region (Fig. 3D). Based on these measured indel frequencies and known triploidy of this K562 genomic region (Zhou et al. 2019), we simulated the expected single-cell gene expression distributions due to effects purely arising from the maximal observed frequencies of deletions of any size in the PCR amplicon (Fig. S14B). Under this purely deletional assumption, we calculated an upper estimate of expected 1.8% median KIT expression knockdown by multiAsCas12a-KRAB, far less than the observed 90.4% median expression knockdown, which is 44.4% in excess of the observed for denAsCas12a-KRAB (Fig. S14B). These results demonstrate that target gene knockdown by multiAsCas12a-KRAB is largely attributable to non-genetic perturbation of transcription. For multiAsCas12a (without KRAB), the upper estimate of expected knockdown under the purely deletional assumption is 2.5%, vs. 67.7% observed (Fig. S14B). Given the low observed indel frequency and the absence of the KRAB domain, this observed knockdown for multiAsCas12a likely reflects direct obstruction of the transcriptional machinery, especially by the crKIT-2 target site downstream of TSS. This is consistent with previously reported transcriptional interference via direction

obstruction by dCas9 (Gilbert et al. 2013, 2014; Qi et al. 2013). Similar trends were obtained for single-site targeting using crKIT-2 encoded within a 4-plex crRNA array (Fig. S15). We conclude that marginal effects of indels on target gene knockdown by multiAsCas12a-KRAB are insufficient to affect interpretations in most functional genomics applications.

At the single cell level, multiAsCas12a-KRAB consistently outperforms denAsCas12a-KRAB in the fraction of single cells with successful double knockdown (Fig. 3E-F and Fig. S16) and triple knockdown (Fig. 3G and Fig. S16) of target genes using higher-order crRNA arrays. To test the upper limit of multiplexing, we constructed 8-plex and 10-plex constructs assembled using individually active spacers. In these 8-plex and 10-plex arrays, spacers encoded in various positions within the array maintain robust CRISPRi activity (i.e. for CD55, KIT and B2M, Fig 3H). However, for crCD81-1 encoded at the 3' most position shows progressive diminishment in CRISPRi activity with further multiplexing at 8-plex and 10-plex (Fig. 3H). These results indicate that 8-plex and 10-plex crRNA arrays at least partially support robust CRISPRi activity for most spacers within these arrays. The patterns are most consistent with an intrinsic deficiency of crCD81-1 that is unmasked by a reduction of effective multiAsCas12a-KRAB concentration by further multiplexing, likely related to its observed dose sensitivity (Fig. 2B-C). We also observed that a specific 6-plex crRNA construct (crCD81-1_crB2M-1_crB2M-3_crKIT-2_crKIT-3_crCD55-4, 6-plex #2 in Fig. 3H) fails to knockdown B2M while achieving robust CRISPRi of the other target genes. However, the same combination of spacers in a slightly different 6-plex arrangement (crCD55-4_crB2M-1_crB2M-3_crKIT-2_crKIT-3_crCD81-1) and also in 8-plex and 10-plex embodiments achieve decent B2M knockdown (Fig. 3H). These results indicate the existence of still unpredictable pre-crRNA sequence context influences on CRISPRi activity of specific spacers within crRNA arrays that is unrelated to distance from the U6 promoter, and perhaps related to pre-crRNA folding (Creutzburg et al. 2020).

multiAsCas12a-KRAB outperforms denCas12a-KRAB and performs similarly to dCas9-KRAB in pooled single-guide CRISPRi screens

Given the success of multiAsCas12a-KRAB in individual well-based assays using lentivirally delivered crRNAs, we next evaluated its performance in the context of high-throughput pooled screens using sequencing to quantify the activities of each crRNA construct within a pooled library. We designed a library, referred to as Library 1 (summarized in Fig. S17), aimed at extracting patterns for Cas12a CRISPRi activity with respect to genomic position relative to the TSS using cell fitness as a readout. Library 1 contains 77,387 single crRNA lentiviral constructs tiling all predicted canonical TTTV PAM sites and non-canonical PAM's (recognizable by enAsCas12a (Kleinstiver et al.

2019)) in the -50bp to +300bp region around the TSS's of 559 common essential genes with K562 cell fitness defects in prior genome-wide dCas9-KRAB screens (Horlbeck et al. 2016a). The library also includes two types of negative controls: 1) 524 crRNAs targeting intergenic regions away from predicted regulatory elements across all human cell lines based on ENCODE chromatin accessibility data, and 2) 445 non-targeting crRNAs that do not map to the human genome.

Using K562 cells piggyBac-engineered to constitutively express multiAsCas12a-KRAB or denAsCas12a-KRAB, we transduced cells with this TSS tiling crRNA library (MOI = 0.15), collected a sample at the start of the screen and then carried out the cell fitness screen for ~8 total cell population doublings in replicate. Genomic DNA was extracted from each sample and the relative abundance of each crRNA in the cell population was measured by sequencing. In this assay, cell fitness defects due to CRISPRi knockdown of target essential genes is reflected in depletions of crRNA sequencing read relative abundances over the duration of the screen. Using relative read abundances normalized to the medians of negative control crRNAs, we calculate a cell fitness score for each crRNA construct that quantifies the fractional fitness defect per cell population doubling (defined as γ in (Kampmann et al. 2013)). Concordance between cell fitness scores of screen replicates is high for multiAsCas12a-KRAB (R = 0.71) and much lower for denAsCas12a-KRAB (R = 0.32), the latter due to much lower signal-to-background ratio (Fig. S18). The cell fitness score distributions are virtually indistinguishable between the intergenic targeting negative controls and the non-targeting negative controls (Fig. S19), indicating no appreciable non-specific genotoxicity from multiAsCas12a-KRAB single-site targeting. Among the 3,326 crRNA's targeting canonical TTTV PAM's, 24.5% vs. 17.5% showed a fitness defect in multiAsCas12a-KRAB vs. denAsCas12a-KRAB, respectively (using the 5th percentile of intergenic negative controls as a threshold), with the magnitude of effect for each crRNA overall stronger for multiAsCas12a-KRAB (Fig. 4B). In contrast, crRNAs targeting non-canonical PAMs show fitness scores less clearly distinguishable from the negative control distribution (Fig. 4C). We compared our observed cell fitness scores to on-target activity predictions by CRISPick, the state-of-the-art crRNA activity prediction algorithm trained on enAsCas12a DNase screening data (DeWeirdt et al. 2021; Kim et al. 2018). Because CRISPick on-target predictions are already tightly associated with whether a crRNA targets a TTTV or non-canonical PAM, we analyzed the predictive power of CRISPick within each of these two PAM categories separately (Fig. 4C). Within each PAM category, the CRISPick on-target prediction score weakly correlates with the observed cell fitness score (R = -0.18 for TTTV PAM and R = -0.1 for non-canonical PAMs), indicating significant sources of crRNA CRISPRi activity variation beyond what is modeled by CRISPick.

Previous studies using dCas9-KRAB have identified a strong association between CRISPRi activity and genomic proximity of the crRNA binding site to the TSS (Gilbert et al. 2014; Nuñez et al. 2021). To facilitate direct comparison of our current multiAsCas12a-KRAB and denCas12a-KRAB TSS tiling analysis with prior dCas9-KRAB data, we focused on crRNAs targeting TTTV PAM's near the TSS's of 240 essential genes for which dCas9-KRAB tiling data is available (Nuñez et al. 2021). The average cell fitness scores of these crRNAs at each genomic position relative to the TSS reveal a remarkably similar bimodal pattern in CRISPRi activity as that obtained by dCas9-KRAB sgRNA's targeting NGG PAM's in the same genomic windows (Fig. 4D). The weakened activity centered around +125bp to +150bp region is consistent with hindrance by a well-positioned nucleosome (Nuñez et al. 2021; Horlbeck et al. 2016b). multiAsCas12a-KRAB shows similar magnitudes of averaged cell fitness scores as dCas9-KRAB in this meta-TSS analysis, whereas denCas12a-KRAB is substantially weaker than both at all positions relative to the TSS (Fig. 4D). Using the average CRISPRi activity of the top 3 best performing crRNAs/sgRNA for each TSS as a benchmark, multiAsCas12a-KRAB still outperforms denAsCas12a-KRAB and is similar to dCas9-KRAB for (Fig. 4E). In this comparison, the top 3 sgRNA's per TSS for dCas9-KRAB are taken from a prior genome-wide screen that used 10 sgRNAs per TSS, which were pre-selected based on bioinformatic prediction of strong sgRNA activity (Horlbeck et al. 2016a). In summary, by testing CRISPRi activity across numerous single crRNAs and genes in these large-scale pooled experiments, we demonstrate that multiAsCas12a-KRAB systematically outperforms denCas12a-KRAB (Fig. 4B, Fig. 4D-E) and that multiAsCas12a-KRAB performs similarly to dCas9-KRAB (Fig. 4D-E).

multiAsCas12a-KRAB enables pooled sequencing screens using 6-plex crRNA arrays

To evaluate the performance of multiAsCas12a-KRAB in pooled sequencing screens using multiplexed crRNA constructs, we constructed a library consisting of 6-plex crRNAs, as this length is supported by the upper end of commercially available pooled oligo synthesis. We refer to this 6-plex library as Library 2 (summarized in Fig. S17), which includes Sublibrary A (described in this section) and Sublibrary B (described in the next section). Sublibrary A was designed to contain 84,275 6-plex constructs for evaluating CRISPRi activity at each of the 6 positions in the array in a K562 cell fitness screen (Fig. 4F). Each 6-plex construct has one of the 6 positions designated as the "test" position, which can encode either 1) a spacer targeting one of the top 50 essential gene TSS's (ranked based on prior dCas9-KRAB screen data (Nuñez et al. 2021)), or 2) an intergenic negative control (Fig. 4F). The remaining 5 positions in the array are designated as "context" positions that encode negative control spacers drawn from a

separate set of 30 negative control spacers (Fig. 4F). The motivation for this library design was to enable sampling multiple sets of context spacers for a given test position.

The entirety of Library 2 was used in a cell fitness screen conducted as described in the previous section, but over ~13.5 total cell population doublings in K562 cells piggyBac-engineered to stably express multiAsCas12a-KRAB. As Library 2 was designed and cloned prior to the completion of the Library 1 screen, the majority of Library 2 contains constructs encoding for spacers in the test position that in hindsight do not produce strong phenotypes as single crRNAs in the Library 1 screen. Thus, we focused our analysis on 1) 2,987 6-plex crRNA arrays that encode in the test position one of 123 spacers with empirically strong cell fitness scores as single crRNAs in the Library 1 screen, and 2) 12,029 6-plex crRNA arrays that encode in the test position one of 506 negative control spacers. We calculated the average cell fitness scores from the top 3 context constructs ranked by cell fitness score for a given test position spacer, applying this calculation equally to essential TSS-targeting test position spacers and negative control test position spacers (Fig. 4F). The top 3 context-averaged cell fitness scores for essential TSS-targeting spacer are clearly distinguishable from the negative control distributions at each test position in the 6-plex array (Fig. 4G), albeit with weaker magnitudes than for the same spacer encoded as individual single crRNA in the Library 1 screen (Fig. S20). We used the 5th percentile of the intergenic negative control cell fitness score distribution in each test position as a threshold for calling whether the test position spacer shows successful CRISPRi activity. Using this threshold, we quantified the % recall of empirically active single crRNA spacers from the Library 1 screen by the 6-plex crRNA constructs in this Library 2 Sublibrary A screen, observing a range of 59%-85% recall across positions (Fig. 4G). As the % recall is not a monotonically decreasing function of distance from the U6 promoter, the pattern is inconsistent with any potential abortive RNA Pol III transcription of the pre-crRNA being the sole driver of these biases. As each position in the array is assigned a unique direct repeat variant that is held constant across all constructs in this analysis, it is possible these apparent positional effects may reflect contributions from unknown properties intrinsic to the direct repeat variant sequences. These results suggest that redundantly sampling the same combination of spacers encoded in different orders within arrays can reduce false negative results. We conclude that multiAsCas12a-KRAB enables pooled sequencing screens using 6-plex crRNA arrays.

multiAsCas12a-KRAB enables discovery and higher-order combinatorial perturbations of cis-regulatory elements

The human genome contains ~500,000 predicted enhancers (ENCODE Project Consortium et al. 2020), but only a small minority have been functionally tested by

perturbations. Previous studies have shown that dCas9 CRISPRi can outperform Cas9 DNA cutting in perturbing enhancer function in pooled screens (Ren et al. 2021; Tycko et al. 2019), likely due to the broader genomic window DNA that is perturbed by formation of repressive chromatin versus indels generated by individual guide RNAs. To our knowledge, no study has reported enhancer perturbation by CRISPRi using Cas12a. We confirmed that multiAsCas12a-KRAB targeting using single crRNAs can effectively perturb a known enhancer of the HBG gene, HS2 (Fig. 5A), with knockdown of HBG mRNA expression in K562 cells comparable to the effect of dCas9-KRAB targeting HS2 (Li et al. 2020).

We next aimed to use multiAsCas12a-KRAB to discover previously uncharacterized enhancers using the CD55 locus in K562 cells as a myeloid cell model. CD55 encodes for decay-accelerating factor, a cell surface protein that inhibits the activation of complement and is expressed in most human cell types (Dho et al. 2018). CD55 function in the myeloid lineage is particularly relevant in multiple disease states, including paroxysmal nocturnal hemoglobinuria (Hillmen et al. 2004) and malaria (Egan et al. 2015; Shakya et al. 2021). To our knowledge, no known enhancers in myeloid cells have been identified for CD55. In K562 cells, several DNase hypersensitive sites (DHS) bearing histone 3 lysine 27 (H3K27Ac), a modification associated with active enhancers, reside near CD55 (Fig. 5B). The activity-by-contact (ABC) enhancer prediction algorithm (Fulco et al. 2019) predicts 4 of these DHSs (R1-R4) as candidate enhancers (Fig. 5B). While R1-R3 reside in a region between 3kb-11kb upstream of the CD55 promoter, R4 sits in an intronic region of the C1orf116 gene, ~297kb away from the CD55 promoter (Fig. 5B). To conduct a focused screen of the DHSs within this general region for enhancers that regulate CD55, we designed a total of 21 4-plex crRNAs (encompassing 88 individual spacers) targeting 11 regions bearing varying levels of DNase hypersensitivity and H3K27Ac (R1-R4 predicted by ABC; R5-R11 picked manually), plus a negative control region (R12) devoid of DHS and H3K27Ac. The regions are each independently targeted by two 4-plex crRNAs (except R10 and R12 are each targeted by one 4-plex crRNA). Each 4-plex crRNA was lentivirally transduced into K562 cells piggyBac-engineered to constitutively express multiAsCas12a-KRAB, followed by flow cytometry readout of CD55 expression. We found that the ABC-predicted R1-R4 show ~50%-75% reduction in CD55 expression upon multiAsCas12a CRISPRi targeting, whereas no decrease in CD55 expression is observed for R5-R12. For each of the functionally validated R1-R4 enhancers, the two 4-plex crRNA arrays that target each enhancer show quantitatively similar levels of CD55 knockdown (Fig. 5B), indicating each array contains an active 4-plex or lower-order active combination of spacers. This consistency in the magnitude of CD55 expression knockdown likely reflects the magnitude of true enhancer impact on gene transcription, rather than technical peculiarities of individual spacer activities, which

might be more unpredictably variable and labor-intensive to test if encoded as single-plex perturbations. In contrast to multiAsCas12a-KRAB, using opAsCas12a to target R1-R4 for DNA cutting using the same 4-plex crRNAs elicits very little or no CD55 expression knockdown, despite potent knockdown by a positive control crRNA targeting a coding exon (Fig. 5C). This demonstrates a key advantage of multiAsCas12a-KRAB over state-of-the-art Cas12a DNA cutting tools for perturbing enhancer function, even in the setting of multiple crRNA target sites within the same enhancer. These findings are consistent with prior studies reporting superior performance of dCas9-KRAB over Cas9 for enhancer discovery in pooled screens (Ren et al. 2021; Tycko et al. 2019). To our knowledge, R1-R4 are the first functionally demonstrated enhancers for CD55 in a myeloid cell type, in addition to another CD55 enhancer that was recently reported in a B-cell model (Cheng et al. 2022).

To further test the utility of multiAsCas12a-KRAB in studies of enhancer function, we used the MYC locus as a model. MYC is an essential gene in most proliferative cells when perturbed by CRISPRi, enabling the use of cell fitness as a readout in pooled screens to identify genomic elements that regulate MYC expression. Prior studies using CRISPRi pooled screens in K562 cells have shown that MYC expression is proportional to cell fitness and is regulated by several enhancers identified by both screens using cell fitness (Fulco et al. 2016) and mRNA expression (Reilly et al. 2021) readouts. A recent study found that pairwise dCas9-KRAB perturbations of these enhancers showed stronger phenotypes than perturbing single enhancers (Lin et al. 2022). In that study, a single-step large scale pooled screen was used to test 295 x 295 = 87,025 pairs of guide RNAs targeting known MYC enhancers in a single step. To our knowledge, no study has reported the phenotypic impact of 3-plex or higher-order perturbations of regulatory elements at the MYC locus.

We used multiAsCas12a-KRAB to dissect higher-order combinatorial cis-regulation at the MYC locus. To avoid testing intractably numerous higher-order combinations of crRNA spacers that are largely uninformative due to the inclusion of weak or inactive crRNA spacers, we opted to pre-screen for a small group of active 3-plex crRNA combinations that can be assembled into higher-order combinations in a subsequent step. We used multiAsCas12a-KRAB to test four 3-plex crRNA constructs targeting combinations of MYC cis-regulatory elements (3 crRNAs for promoter and 3 crRNAs for each of 3 known enhancers, e1, e2 and e3) in a well-based cell competition assay (Fig. 6A-B). We found that these four 3-plex crRNAs induce varying degrees of cell fitness defect as a proxy of MYC expression knockdown. This pre-nomination step indicates that each construct contains some 3-plex or lower-order spacer combinations that exhibit CRISPRi activity. For comparison, we included denAsCas12a-KRAB, multiAsCas12a, enAsCas12a-KRAB and opAsCas12a as controls, which showed

relative activities in MYC knockdown phenotype that are consistent with our prior results from targeting other loci (Fig. 3C and Fig. S12). This further supports multiAsCas12a-KRAB's superior CRISPRi potency and that the cell fitness phenotypes elicited by targeting regions in the MYC locus are not attributable to its residual DNA cutting activity (Fig. 6B).

We then in silico assembled these 12 nominated spacers and 3 intergenic negative control spacers into Library 2 Sublibrary B (summarized in Fig. S17), consisting of 6,370 6-plex permutations encoded as 6-plex crRNA arrays (Fig. 6C). These 6-plex crRNA arrays each target up to 4 cis-regulatory elements (promoter + 3 enhancers) with up to 3 spacers per element. Negative control spacers fill in the remaining positions in arrays that are not fully filled by targeting spacers. This Sublibrary B was included as part of the cell fitness screen for the entirety of Library 2, as described in the previous section. Among 1,823 6-plex arrays with sufficient read coverage for analysis, we grouped them into 16 categories, based on whether it encodes at least 1 spacer targeting the promoter, and/or at least 1 spacer targeting each of the 3 enhancers (Fig. 6D). The 6-plex crRNA arrays targeting only e1, e2, or e3 alone showed a modest cell fitness defect, with e2 having the strongest effect of all 3 enhancers (Fig. 6D, left panel). Co-targeting each additional enhancer increased the magnitude of fitness defect, such that the crRNA arrays co-targeting e1/e2/e3 shows the strongest fitness defect (Fig. 6D, left panel). Co-targeting the promoter together with any combination of enhancers showed increased cell fitness defect over targeting the promoter alone while also exhibiting the cumulative effects of multi-enhancer targeting (Fig. 6D, right panel). These results suggest that when targeting subsets of cis-regulatory elements in a locus by CRISPRi components, other cis-regulatory elements can compete with KRAB domain-mediated repression to sustain partial levels of gene transcription. Such effects may reflect how cis-regulatory elements combinatorially respond to endogenous repressive cues in the natural regulation of MYC gene transcription. Having multiple cis-regulatory elements that each contributes to the control of MYC gene transcription may serve to collectively integrate a broader variety of upstream signals to fine-tune the transcription of this key regulator of cell growth in a physiologically appropriate manner. For objectives focused on maximizing CRISPRi knockdown of coding genes, co-targeting of gene TSS and enhancers may achieve better gene knockdown for certain loci than targeting the gene TSS alone.

Discussion

In this study, we engineered multiAsCas12a-KRAB as a new platform for higher-order combinatorial CRISPRi perturbations of gene transcription and enhancer function. The enhanced CRISPRi potency of multiAsCas12a-KRAB is more robust to lower effective concentrations of ribonucleoprotein (Fig. 2B-C, Fig. 2E), critically enabling high-order

multiplexing (Fig. 3) and high-throughput pooled screening applications conducted at single-copy integrations of crRNA expression (Fig. 4 and Fig. 6C-D). We propose that the improved CRISPRi activity of multiAsCas12a-KRAB emerges from prolonged chromatin occupancy due to DNA nicking (Fig. 2A). This strategy is conceptually distinct from prior protein engineering approaches to improving Cas12a function in mammalian cells, which focused on substituting for positively charged amino acid residues near the protein:DNA interface (Kleinstiver et al. 2019; Guo et al. 2022), using directed evolution to optimize DNA cleavage (Zhang et al. 2021), or optimizing transcriptional effector domain function rather than Cas12a enzymology (Griffith et al. 2023). We propose the following biophysical explanation for improved multiAsCas12a function, grounded in prior in vitro literature. In the absence of nicking, R-loop reversal occurs by invasion of the non-target strand into the crRNA:target strand duplex, displacing the crRNA in a process analogous to toehold-mediated nucleic acid strand displacement (Srinivas et al. 2013). Severing the non-target strand increases its conformational entropy and effectively destroys the toehold, decreasing the rate at which the non-target strand can invade the crRNA:target strand duplex (Srinivas et al. 2013). This mechanistic model can also explain previous observations of cutting-dependent complex stabilization (Cofsky et al. 2020; Knott et al. 2019; Singh et al. 2018) and suggests that favoring nicked DNA intermediates may be a generalizable strategy for improving the efficacy of other Cas enzymes in chromatin targeting. Another potential explanation for multiAsCas12a's enhanced CRISPRi activity may be the formation of stabilizing protein:DNA contacts after non-target-strand nicking (Naqvi et al., 2022). Separately, DNA nicking is expected to relax local supercoiling, which might affect transcription at nearby loci (Baranello et al. 2012).

The multiAsCas12a-KRAB platform enables new solutions to addressing a central challenge in combinatorial genetics. With increased crRNA multiplexing beyond $\geq$3-plex combinations, the combinatorial space rapidly explodes in size, with functionally important combinations being a rare subset of the entire combinatorial space. Exhaustive testing of all combinations to search is thus highly inefficient and often infeasible. However, testing a single higher-order N-plex combination also indirectly tests all or many of its constituent lower-order combinations, for up to a total of 2^N combinations. Thus, increases in multiplexing capability potentially yield exponential increases in search efficiency using the general concept of group testing (Dorfman 1943; Du 1993). In group testing (Fig. 7), a primary screen is conducted on grouped subjects (e.g. a multiplexed array of crRNA constructs) to reduce the costs otherwise incurred by individually testing all subjects (e.g. an individual crRNA). Our screen for CD55 enhancers instantiates this approach by testing 22 4-plex crRNA arrays targeting 12 candidate regions, therefore indirectly testing 22 x 2^4 = 352 crRNA combinations in a cost-effective experiment using flow cytometry to assay just 22 wells in a plate (Fig. 5B).

For this experimental objective, the grouped hits can be biologically interpreted without further testing (Fig. 7). For other objectives, such as the combinatorial analysis of cis-regulation at the MYC locus (Fig. 6), grouped hits can be followed by a focused secondary pooled screen or individual validation as needed (Fig. 7).

These results together demonstrate that the group testing framework can be used flexibly for individual well-based assays and/or pooled screening readouts. For pooled screens with sequencing readouts, the ability to deterministically program only specific higher-order combinations of compact Cas12a crRNAs by oligo synthesis for the initial screen is crucial for group testing. In contrast, cloning combinatorial guide libraries by a multiplicative and stochastic approach (Zhou et al. 2020) requires testing all combinations at the onset, and thus are incompatible with group testing. The application of group testing can significantly compress the size of crRNA libraries to facilitate combinatorial genetic screens, including in biological systems limited by assayable cell numbers, such as primary cells, organoids, and in vivo models. Group testing may be combined with concepts from compressed sensing (Yao et al. 2023; Cleary and Regev 2020) to further enhance the efficiency of surveying combinatorial genetic perturbations for multidimensional phenotypic readouts (Adamson et al. 2016, 2018; Dixit et al. 2016; Wessels et al. 2023; Datlinger et al. 2017; Schraivogel et al. 2020; Replogle et al. 2020; Norman et al. 2019; Feldman et al. 2019).

A key parameter in group testing is the extent of potential signal dilution relative to individual testing. For Cas12a perturbations, signal dilution can arise from 1) low doses of ribonucleoprotein due to limitations in delivery, such as crRNA expression from single-copy integrations in pooled sequencing screens, or 2) increased multiplexing, which effectively dilutes the concentration of functional Cas12a protein available to bind each individual crRNA. Despite some signal dilution in the stringent setting of pooled screens using 6-plex crRNAs expressed from single-copy integrants (Fig. 4G and Fig. S20), multiAsCas12a-KRAB demonstrates sufficient potency to yield new biological insights into combinatorial cis-regulation at the MYC locus using pooled screening of 6-plex crRNA arrays (Fig. 6C-D). While we have focused our optimizations to meet the stringent single-copy crRNA integration requirement of pooled screening formats, multiAsCas12a also significantly lowers technical barriers to higher-order combinatorial perturbations in array-based screening (Fig. 5B), which is compatible with more diverse phenotypic readouts and has recently improved significantly in throughput (Yin et al. 2022). The assay format will likely influence the deliverable dose of synthetic components and thus the absolute upper limit of multiplexing for effective CRISPRi using multiAsCas12a-KRAB, which currently remains unclear. Among the spacers examined in the largest crRNA array we tested (10-plex), 3 spacers performed the same as each does in shorter arrays, while one showed substantially diminished

CRISPRi activity (Fig. 3H). Thus, there exists some unpredictability in how the pre-crRNA length, position within the array, and other sequence contexts might together influence the activities of specific crRNAs in the array. Predicting such context-specific influences on the performance of higher-order multiplexed crRNA arrays is an uncharted frontier that will likely improve the design of crRNA arrays to function robustly even in highly multiplexed and/or dose-limited applications. Such improvements in crRNA array activity prediction will further extend the scalability of combinatorial genetic screens by group testing. Implementing group testing would enable a single 10-plex crRNA array to indirectly screen up to 2^{10} = 1,024 crRNA combinations. Another area for improvement is that, despite already incorporating mutations that were previously reported in transient transfection experiments to enable non-canonical PAM targeting (Kleinstiver et al. 2019), multiAsCas12a-KRAB CRISPRi activity at those non-canonical PAM sites is generally much weaker than for the canonical TTTV PAM (Fig. 4C). Preference for the TTTV PAM enables targeting AT-rich genomic regions, but also limits the number of active crRNAs for GC-rich TSS-proximal regions. Compared to denAsCas12a-KRAB, multiAsCas12a-KRAB boosts the number of potent crRNAs per TSS, but further engineering may improve this further.

While we have focused on CRISPRi applications using the KRAB domain in the present study, the discovery and engineering of effector domains for chromatin perturbations by CRISPR-Cas is rapidly evolving. Recent advances include new repressive effectors (Alerasool et al. 2020; Mukund et al. 2023; Replogle et al. 2022b; DelRosso et al. 2023), activation effectors (Alerasool et al. 2022; Mukund et al. 2023; DelRosso et al. 2023; Griffith et al. 2023), and combination effectors for epigenetic memory (Nuñez et al. 2021; Van et al. 2021; Nakamura et al. 2021; Amabile et al. 2016; O'Geen et al. 2022). We expect that multiAsCas12a can be flexibly combined with these and other effector domains to support group testing for many chromatin perturbation objectives. We envision multiAsCas12a and the group testing framework will enable elucidating and engineering combinatorial genetic processes underlying broad areas of biology at previously intractable scales.

Author Contributions

C.C.H. and L.A.G. conceived of and led the design of the overall study. C.C.H. wrote the paper incorporating input from L.A.G., C.M.W., J.C.C., J.S., and agreement from all authors. C.C.H. conceived of testing the R1226A mutation for CRISPRi applications, proposed the conceptual link to group testing, and led the design, execution, and computational analysis of all experiments, with contributions from others noted below. L.A.G. proposed the dose of lentivirally delivered CRISPR components as a key parameter for evaluation, interpreted results, and guided and obtained funding for the overall study. C.M.W. designed, executed, and analyzed quantification of indel

frequencies, and contributed to cloning and flow cytometry-based CRISPRi experiments. N.A.S. contributed to the execution of flow cytometry-based CRISPRi experiments, including all experiments in C4-2B cells, testing of higher-order crRNAs arrays, replicates of dose-response CRISPRi experiments, and enhancer CRISPRi experiments at the CD55 and MYC loci. R.D. contributed to cell line engineering and CRISPRi analysis of the DNase-dead mutant panel. Q.C. designed, executed, and analyzed all experiments testing truncated crRNAs, with guidance and funding obtained by J.S.. J.C.C. provided feedback on interpretations of the biophysics literature and proposed the entropic barrier to R-loop reversal upon severance of the non-target strand. S.M. contributed to crRNA cloning and cell culture. T.O. prepared 3' RNA-seq Illumina sequencing libraries. A.A. contributed to screen data processing and analysis for 3' RNA-seq. N.T. wrote scripts in Rust for screen read mapping and counting.

Acknowledgements

We thank Brian Yu and Khushali Patel for assistance in sequencing the indel analysis library for the KIT locus; April Pawluk for feedback on manuscript; Gavin Knott, Ashir Borah, Garrett Wong, James Nunez, Jonathan Weissman, Howard Chang, Patrick Hsu, Silvana Konermann, and the Gilbert Lab for helpful discussions.

Funding

C.C.H. is supported by the Physician Scientist Incubator, Dept. of Pathology, Stanford University School of Medicine and NIH NHGRI (K01HG012789). Q.C. and J.S. were supported by the Mark Foundation for Cancer Research. J.C.C. is a fellow of the Helen Hay Whitney Foundation. This work was funded in part by NIH (R01HG012227) and in part by GSK through an award to the UCSF Laboratory for Genomics Research to L.A.G.. L.A.G. is funded by the Arc Institute, NIH (DP2CA239597, UM1HG012660), CRUK/NIH (OT2CA278665 and CGCATF-2021/100006), a Pew-Stewart Scholars for Cancer Research award and the Goldberg-Benioff Endowed Professorship in Prostate Cancer Translational Biology. Sequencing of CRISPR screens and 3' RNA-seq libraries was performed at the UCSF CAT, supported by UCSF PBBR, RRP IMIA, and NIH 1S10OD028511-01 grants. Sequencing of indel analysis at the KIT locus was performed at the Multi-omics Technology Center at the Arc Institute.

Declaration of Interests

C.C.H., C.M.W., R.D. and L.A.G. have filed patent applications related to multiAsCas12a. J.S. is a scientific consultant for Treeline Biosciences. L.A.G has filed patents on CRISPRoff/on, CRISPR functional genomics and is a co-founder of Chroma Medicine.

Methods

Plasmid Design and Construction

A detailed table of constructs generated in this study will be provided as a Supplemental File with all sequences. Constructs will be made available on Addgene. Cloning was performed by Gibson Assembly of PCR amplified or commercially synthesized gene fragments (from Integrated DNA Technologies or Twist Bioscience) using NEBuilder Hifi Master Mix (NEB Cat# E262), and final plasmids sequence-verified by Sanger sequencing of the open reading frame and/or commercial whole-plasmid sequencing service provided by Primordium.

Protein constructs components:

The denAsCas12a open reading frame was PCR amplified from pCAG-denAsCas12a(E174R/S542R/K548R/D908A)-NLS(nuc)-3xHA-VPR (RTW776) (Addgene plasmid # 107943, from (Kleinstiver et al. 2019)). AsCas12a variants described were generated by using the denAsCas12a open reading frame as starting template and introducing the specific mutations encoded in overhangs on PCR primers that serve as junctions of Gibson assembly reactions. opAsCas12a is from (Gier et al. 2020), available as Addgene plasmid # 149723, pRG232). 6xMyc-NLS was PCR amplified from pRG232. KRAB domain sequence from KOX1 was previously reported in (Gilbert et al. 2013). The lentiviral backbone for expressing Cas12a fusion protein constructs are expressed from an SFFV promoter adjacent to UCOE and is a gift from Marco Jost and Jonathan Weissman, derived from a plasmid available as Addgene 188765. XTEN80 linker sequence was taken from (Nuñez et al. 2021) and was originally from (Schellenberger et al. 2009). For constructs used in piggyBac transposition, the open reading frame was cloned into a piggyBac vector backbone (Addgene #133568) and expressed from a CAG promoter. Super PiggyBac Transposase (PB210PA-1) was purchased from System Biosciences.

dAsCas12a-KRABx3 open reading frame sequence is from (Campa et al. 2019), in that study encoded within a construct referred to as SiT-ddCas12a-[Repr]. We generated SiT-ddCas12a-[Repr] by introducing the DNase-inactivating E993A by PCR-based mutagenesis using SiT-Cas12a-[Repr] (Addgene #133568) as template. Using Gibson Assembly of PCR products, we inserted the resulting ddCas12a-[Repr] open reading frame in-frame with P2A-BFP in a piggyBac vector (Addgene #133568) to enable direct comparison with other fusion protein constructs cloned in the same vector backbone (crRNA's are encoded on separate plasmids as described below).

Fusion protein constructs described in Fig. S8 were assembled by subcloning the protein-coding sequences of AsCas12a and KRAB into a lentiviral expression vector

using the In-Fusion HD Cloning system (TBUSA). AsCas12a mutants were cloned by mutagenesis PCR on the complete wildtype AsCas12a vector to generate the final lentiviral expression vector.

crRNA expression constructs:

Unless otherwise specified, individual single and 3-plex crRNA constructs were cloned into the human U6 promoter-driven expression vector pRG212 (Addgene 149722, originally from (Gier et al. 2020)). Library1, Library2, some 3-plex and all 4-plex, 5-plex, and 6-plex *As.* crRNA constructs were cloned into pCH67, which is derived from pRG212 by replacing the 3' DR with the variant DR8 (DeWeirdt et al. 2021). For constructs cloned into pCH67, the specific *As.* DR variants were assigned to each position of the array as follows, in 5' to 3' order:
3-plex: WT DR, DR1, DR3, DR8
4-plex: WT DR, DR1, DR10, DR3, DR8
5-plex: WT DR, DR1, DR16, DR10, DR3, DR8
6-plex: WT DR, DR1, DR16, DR18, DR10, DR3, DR8
8-plex: WT DR, DR1, DR16, DR_NS1, DR17, DR18, DR10, DR3, DR8
10-plex: WT DR, DR1, DR16, DR_NS1, DR4, DR_NS2, DR17, DR18, DR10, DR3, DR8

Where the sequences of DR_NS1 and 2 were based on combining hits from the variant DR screen from DeWeirdt et al., 2020. The sequences are DR_NS1: aattcctcctcttggaggt, and DR_NS2: aattcctcctataggaggt.

1-plex,3-plex, 8-plex, and 10-plex crRNA constructs were cloned by annealing complementary oligos, phosphorylation by T4 polynucleotide kinase (NEB M0201S), and ligated with T4 DNA ligase (NEB M0202) into BsmbI site of vector backbones. 4-plex, 5-plex and 6-plex crRNA arrays were ordered as double-stranded gene fragments and cloned into the BsmbI site of vector backbones by Gibson Assembly.

Design of individual crRNAs

For cloning individual crRNA constructs targeting TSS's, CRISPick was used in the enAsCas12a CRISPRi mode to design spacers targeting PAM's located within -50bp to +300bp region around the targeted TSS. We manually selected spacers from the CRISPick output by picking TTTV PAM-targeting spacers (except for crCD151-3, which targets a non-canonical GTTC PAM) with the highest On-Target Efficacy Scores and generally excluded any spacers with high off-target predictions. The same non-targeting spacer was used throughout the individual well-based experiments and was randomly generated and checked for absence of alignment to the human genome by BLAT (Kent 2002).

The hg19 genomic coordinates for MYC enhancers are: e1 chr8:128910869-128911521, e2 chr8:128972341-128973219, and e3 chr8:129057272-129057795. DNA sequences from those regions were downloaded from the UCSC Genome Browser and submitted to CRISPick. The top 3 spacers targeting TTTV PAM's for each enhancer were picked based on CRISPick On-target Efficacy Score, having no Tier I or Tier II Bin I predicted off-target sites, and proximity to the zenith of the ENCODE DNase hypersensitivity signal in K562 cells.

Cell culture, lentiviral production, lentiviral transduction

All cell lines were cultured at 37deg. C with 5% CO2 in tissue culture incubators. K562 and C4-2B cells were maintained in RPMI-1640 (Gibco cat# 22400121) containing 25 mM HEPES, 2mM L-glutamine, and supplemented with 10% FBS (VWR), 100 units/mL streptomycin, and 100 mg/mL penicillin. For pooled screens using K562 cells cultured in flasks in a shaking incubator, the culture media was supplemented with 0.1% Pluronic F-127 (Thermo Fisher P6866).

HEK 293T cells were cultured in media consisting of DMEM, high glucose (Gibco 11965084, containing 4.5g/mL glucose and 4mM L-glutamine) supplemented with 10% FBS (VWR) and 100units/mL streptomycin, 100mg/mL penicillin. Adherent cells are routinely passaged and harvested by incubation with 0.25% Trypsin-EDTA (Thermo Fisher 25200056) at 37deg. C for 5-10min, followed by neutralization with media containing 10% FBS.

Unless otherwise specified below, lentiviral particles were produced by transfecting standard packaging vectors into HEK293T using TransIT-LT1 Transfection Reagent (Mirus, MIR2306). At <24 hours post-transfection culture media with exchanged with fresh media supplemented with ViralBoost (Alstem Bio, cat# VB100) at 1:500 dilution. Viral supernatants were harvested ~48-72 hours after transfection and filtered through a 0.45 mm PVDF syringe filter and either stored in 4deg. C for use within <2 weeks or stored in -80deg. C until use. Lentiviral infections included polybrene (8 mg/ml).

For experiments described in Fig. S8, lentivirus was produced by transfecting HEK293T cells with lentiviral vector, VSVG and psPAX2 helper plasmids using polyethylenimine. Media was changed ~6–8 h post transfection. Viral supernatant was collected every 12 h for 5 times and passed through 0.45 μm PVDF filters. Lentivirus was added to target cell lines with 8 μg/mL Polybrene and centrifuged at 650 × g for 25 min at room temperature. Media was replaced 15 h post infection. Antibiotics (1 μg/mL puromycin) was added 48 h post infection.

Antibody staining and flow cytometry

Antibodies used: CD55-APC (Biolegend 311312), CD81-PE (Biolegend 349506), B2M-APC (Biolegend 316311), KIT-PE (Biolegend 313204), FOLH1-APC (Biolegend 342508), CD56(NCAM1)-APC (Invitrogen 17-0567-42). Cells were stained with antibodies in 96-well plates, using 500g 5min at 4deg. C for centrifugation steps and decanting in between each step. Cells were washed once with 200ul with FACS Buffer (PBS with 1% BSA), then resuspended in 50ul of antibodies diluted at 1:100 in FACS Buffer for 30min at 4deg. C. Then 150ul of FACS Buffer was added, followed by centrifugation and supernatant, then washed one more time with 200ul FACS buffer, followed by final resuspension in 200ul FACS Buffer for flow cytometry. For CRISPRi experiments, all data points shown in figures are events first gated for single cells based on FSC/SSC, then gated on GFP-positivity as a marker for cells successfully transduced with crRNA construct. Flow cytometry was performed on the Attune NxT instrument unless otherwise specified.

For cell fitness competition assays, the percentage of cells expressing the GFP marker encoded on the crRNA expression vector is quantified by flow cytometry. log2 fold change of % GFP-positive cells was calculated relative to day 2 (for experiments targeting the Rpa3 locus in Fig. S8) or day 6 (for experiments targeting the MYC locus in Fig. 6B). For experiments targeting the Rpa3 locus, flow cytometry was performed on the Guava Easycyte 10 HT instrument.

Indel analysis

200K cells were collected on day 14 after crRNA transduction and genomic DNA was isolated using NucleoSpin Blood (Macherey-Nagel, Catalog no. 740951.50). Briefly, PCRs for loci of interest were run using Amplicon-EZ (Genewiz) partial IlluminaÒ adapters and amplicons were processed using NucleoSpin Gel and PCR Clean-up Kit (Macherey-Nagel, Catalog no. 740609.250). Paired end (2 x 250 bp) sequencing was completed at GENEWIZ (Azenta Life Sciences). Raw fastq files were obtained from GENEWIZ and aligned to reference sequences using *CRISPResso2 (Clement et al. 2019)* with the following modifications:

--quantification_window_size 12
--quantification_window_center -3
CRISPResso --fastq_r1 R1.fastq.gz --fastq_r2 R2.fastq.gz --amplicon_seq acccgtcttgtttgtcccacccttggtgacgcagagccccagcccagaccccgcccaaagcactcatttaactggtattgcg gagccacgaggcttctgcttactgcaactcgctccggccgctgggcgtagctgcgactcggcggagtcccggcggcgcg tccttgttctaacccggcgcgccatgaccgtcgcgcggccgagcgtgcccgcggcgctgcccctcctcggggagctgccc cggctgctgctgctggtgctgttgtgcctgccggccgtgtggggtgagtaggggcccggcggccggggaagcccctggg ctgggtgggaggtccaagtcggtctctgaga -g actggtattgcggagccacgagg -wc -3 -w 12

For crRNA constructs in which the PAM is found on the opposite strand with respect to the amplicon sequence (in this case, CD81) the following modifications were included:
--quantification_window_size 20
--quantification_window_center -18
CRISPResso --fastq_r1 pCH45H-CD81-array_S5_L001_R1_001.fastq.gz --fastq_r2 pCH45H-CD81-array_S5_L001_R2_001.fastq.gz --amplicon_seq ctgcttcgcggggacgaggggggggctcgcgggcgggactcctggcgccccgcccccatgagctcatcaagagccgc cgcccctggatggtggggcgggggcgcacactttgccggaggttgggggcgatccgcctcactctttccccagcccagct cactctccaatctgcggtcaccacccgagaccttcctgggggtcgcgcctaaaaggagcgcagactcccgccgggatgg cccagaagctggggtgcgcgcaccctggccgtccctgcctgggagccgatctccctctcctcacccagacacgttccagc ggaggcctcctcccagaagggctctggaggcctcgcaggagtggggatcccgcggttctgagttgg -p 3 -g gagaccttcctgggggtcgcgcc -wc -18 -w 20

Quantification diagrams were generated in R.

For analysis of dual cutting at the *KIT* TSS, briefly, DNA was isolated using QuickExtract DNA Solution (Lucigen) and amplicons were generated using 15 cycles of PCR to introduce Illumina sequencing primer binding sites and 0-8 staggered bases to ensure library diversity. After reaction clean-up using ExoSAP-IT kit (Thermo Fisher 78201), an additional 15 cycles of PCR was used to introduce unique dual indices and Illumina P5 and P7 adaptors. Libraries were pooled and purified by SPRIselect magnetic beads before paired-end sequencing using an Illumina MiSeq. Sequencing primer binding sites, unique dual indices (from Illumina TruSeq kits), P5 and P7 adaptor sequences are from Illumina Adaptor Sequences Document # 1000000002694 v16.

Reads were analyzed using CRISPRessoBatch from CRISPResso2 (Clement et al. 2019) with the following modifications: wc -4 -w 15
CRISPRessoBatch --batch_settings batch2.batch --amplicon_seq aagagcaggggccagacgCCGCCGGGAAGAAGCGAGACCCGGGCGGGCGCGAGGGAGG GGAGGCGAGGAGGGGCGTGGCCGGCGCGCAGAGGGAGGGCGCTGGGAGGAGG GGCTGCTGCTCGCCGCTCGCGGCTCTGGGGGCTCGGCTTTGCCGCGCTCGCTGCA CTTGGGCGAGAGCTGGAACGTGGACCAGAGCTCGGATCCCATCGCAGCTACCGCG ATGAGAGGCGCTCGCGGCGCCTGGGATTTTCTCTGCGTTCTGCTCCTACTGCTTCG CGTCCAGACAGGTGGGACACCGCGGCTGGCACCCCGACCGTGcgactactcggcgaagcc tgtg -p 3 -g TCTGCGTTCTGCTCCTACTGCTT -wc -4 -w 15

For dual gRNA cutting, both guides were included in the batch analysis.The total number of insertions and deletions at each amplicon position were calculated and displayed using the effect_vector_combined.txt output.

Frequencies of nucleotide substitutions with multiAsCas12a-KRAB targeting are negligible and indistinguishable from sequencing error (≤4.5%) observed in unmodified K562 cells.

Simulations of indel impacts on gene expression

The fraction of reads containing indels within each specified region was subjected to technical noise background subtraction by the fraction of reads containing indels observed in K562 parental cells. This background-subtracted indel allelic frequency was used to calculate per-copy deletion probabilities in Fig. S14B and Fig. S15B. For denAsCas12a-KRAB this background-subtraction can result in a small negative value and in those cases 0% is reported as per-copy deletion probability.

We simulated the impact of indels on gene expression under the assumption that gene expression changes are entirely driven by indels generated by the given AsCas12a fusion protein at the crRNA target site. A prior study reported at one Cas9 target site that the frequency of larger (>250bp) indels is ~20% relative to the smaller (<250bp) indels. This 20:80 ratio of unobserved-to-observed indels is very likely a high overestimate in our case because our PCR amplicons are 340bp-382bp and thus are expected to capture a large fraction of even the >250bp indels. Nevertheless we added an additional 20% to the observed indel frequencies to arrive at our final estimates of probability of the occurrence of any ≥1bp indels per DNA copy for a single target site. Based on this indel probability we calculated the proportion of cell population expected to harbor a given number of DNA copies with indels assuming indels occur independently among DNA copies and using previously measured DNA copy number for each genomic locus (Zhou et al. 2019). For dual targeting of the KIT locus by crKIT-2 and crKIT-3 we assume that the occurrence of a large (>250bp) unobserved deletion at one target site precludes a deletion at the other target site. Starting from the single-cell expression distributions obtained by flow cytometry from the non-targeting crRNA control, we simulated the expected change in single-cell expression distribution under the assumption that any ≥1bp indel in the PCR amplicon would generate a null allele completely abolishing expression of the target gene in cis (i.e. indel in 1 out of 3 copies would reduce expression by 33%). We refer to this as the "expected null" expression change, which is expected to be a high overestimate of impact on expression. To better more accurately estimate the true hypomorphic effects of indels we calculated a

"hypomorphic coefficient" defined as the ratio of the observed median expression change vs. the expected null expression change for opAsCas12a. We multiply the expected null expression change for all other fusion proteins by this hypomorphic coefficient to derive an "expected hypomorph" expression change for each fusion protein and crRNA construct combination.

Pooled crRNA library design

For all crRNAs in Library 1 and Library 2: we excluded in the analysis spacers with the following off-target prediction criteria using CRISPick run in the CRISPRi setting: 1) off-target match = 'MAX' for any tier or bin, or 2) # Off-Target Tier I Match Bin I Matches > 1). The only crRNAs for which this filter was not applied are the non-targeting negative control spacers, which do not have an associated CRISPick output. All crRNA sequences were also filtered to exclude BsmbI sites used for cloning and >3 consecutive T's, which mimic RNA Pol III termination signal.

Library 1 (single crRNA's)

To design crRNA spacers targeting gene TSS's for Library 1, we used the -50bp to +300bp regions of TSS annotations derived from capped analysis of gene expression data and can include multiple TSS's per gene (Horlbeck et al. 2016a). We targeted the TSS's of 559 common essential genes from DepMap with the strongest cell fitness defects in K562 cells based on prior dCas9-KRAB CRISPRi screen (Horlbeck et al. 2016a). We used CRISPick with enAsCas12a settings to target all possible PAM's (TTTV and non-canonical) in these TSS-proximal regions. Except for the criteria mentioned in the previous paragraph, no other exclusion criteria were applied. For the TSS-level analyses shown in Fig. 4D-E, each gene was assigned to a single TSS targeted by the crRNA with the strongest fitness score for that gene.

Negative controls in Library 1 fall into two categories: 1) intergenic negative controls, and 2) non-targeting negative controls. Target sites for intergenic negative controls were picked by removing all regions in the hg19 genome that are within 10kb of annotated ensembl genes (retrieved from biomaRt from https://grch37.ensembl.org) or within 3kb of any ENCODE DNase hypersensitive site (wgEncodeRegDnaseClusteredV3.bed from http://hgdownload.cse.ucsc.edu/goldenpath/hg19/encodeDCC/wgEncodeRegDnaseClustered/). The remaining regions were divided into 1kb fragments. 90 such 1kb fragments were sampled from each chromosome. Fragments containing >=20 consecutive N's were removed. The remaining sequences were submitted to CRISPick run under CRISPRi settings. The CRISPick output was further filtered for spacers that meet these criteria: 1) off-target prediction criteria described in the beginning of this

section, and 2) On-target Efficacy Score >= 0.5 (the rationale is to maximize representation by likely active crRNAs to bias for revealing any potential cell fitness effects from non-specific genotoxicity due to residual DNA cutting by multiCas12a-KRAB), 3) mapping uniquely to the hg19 genome by Bowtie (Langmead et al. 2009) using '-m 1' and otherwise default parameters, 3) filtered once more against those whose uniquely mapped site falls within 10kb of annotated ensembl genes or any ENCODE DNase hypersensitive site.

Non-targeting negative control spacers were generated by combining 1) non-targeting negative controls in the Humagne C and D libraries, 2) taking 20nt non-targeting spacers from the dCas9-KRAB CRISPRi_v2 genome-wide library (Horlbeck et al. 2016a), removing the G in the 1st position, and appending random 4-mers to the 3' end. This set of spacers were then filtered for those that do not map to the hg19 genome using Bowtie with default settings.

Library 2 (6-plex cRNA's)

Sublibrary A (84,275 constructs): Test position spacers were encoded at each position of the 6-plex array, with remaining positions referred to as context positions and filled with negative control spacers. Test positions encodes one of 506 intergenic negative control spacers and 2,303 essential TSS-targeting spacers. The essential TSS-targeting spacers were selected from among all spacers targeting PAM's within -50bp to +300bp TSS-proximal regions of 50 common essential genes with the strongest K562 cell fitness defect in prior dCas9-KRAB CRISPRi screen (Horlbeck et al. 2016a), and must have $\geq$0.7 CRISPick On-target Efficacy Score. Negative control context spacers consist of 5 6-plex combinations, 3 of these combinations consist entirely of non-targeting negative controls and 2 of the combinations consist entirely of intergenic negative controls.

Sublibrary B (6,370 constructs): crRNA combinations targeting cis-regulatory elements at the MYC locus were assembled from a subset of combinations possible from 15 starting spacers (3 targeting MYC TSS, 3 targeting each of 3 enhancers, and 3 intergenic negative control spacers). The 3 enhancer elements are described in the subsection "Design of individual crRNAs." These 15 starting spacers were grouped into 5 3-plex combinations, each 3-plex combination exclusively targeting one of the 4 cis-regulatory elements, or consisting entirely of intergenic negative controls. Each 3-plex was then encoded in positions 1-3 of 6-plex arrays, and positions 4-6 were filled with all possible 3-plex combinations chosen from the starting 15 spacers. All 6-plex combinations were also encoded in the reverse order in the array.

All-negative control constructs (2000 constructs): 1500 6-plex combinations were randomly sampled from the intergenic negative control spacers described for Library 1. 500 6-plex combinations were randomly sampled from non-targeting negative control spacers described for Library 1.

Intergenic negative controls and non-targeting negative controls are defined the same as in Library 1.

crRNA library construction

For Library 1, ~140 fmol of pooled oligo libraries from Twist were subjected to 10 cycles of PCR amplification using primers specific to adaptor sequences flanking the oligos and containing BsmbI sites. The PCR amplicons were cloned into a crRNA expression backbone (pCH67) by Golden Gate Assembly with ~1:1 insert:backbone ratio using ~500 fmol each. Golden Gate Assembly reaction was carried out in a 100ul reaction containing 2.5U Esp3I (Thermo ER0452) and 1000U T4 DNA Ligase in T4 DNA Ligase reaction buffer (NEB M0202L). The reaction mix was incubated for 31 cycles alternating between 37deg. C and 16deg. C for 20min at each temperature, then heat-inactivated at 65deg. C for 5min. Assembly reactions were column purified with Zymo DNA clean and concentrator-5 (Zymo D4004), eluted in 12ul of water and <7ul added to 70ul of MegaX DH10B T1R Electrocomp Cells (C640003) for electroporation using BioRad Gene Pulser Xcell Electroporator with settings 2.0kV, 200ohms, 25µF. Cells were recovered at 37deg. C for rotating for 1h in ~5ml recovery media from the MegaX DH10B T1R Electrocomp Cells kit and small volumes plated onto bacterial LB plates containing carbenicillin for quantification of colony forming units. The remaining recovery culture was inoculated directly into 200ml liquid LB media with carbenicillin and incubated in 37deg. C shaker for 12h-16h prior to harvesting for plasmid purification using ZymoPURE II Plasmid Midiprep kit (Zymo D4200). Based on the colony forming units from the small volumes in the bacterial plates, the estimated coverage of the library is 778x. 24 individual colonies were verified by Sanger sequencing and the library subjected to deep sequencing as described in Illumina sequencing library preparation. For Library 2, 915 fmol of pooled oligo libraries from Twist was subjected to 18 cycles of PCR amplification and agarose gel purification of the correctly sized band before proceeding similarly with the remainder of the protocol as described above. The estimated coverage of the library from colony forming units is ~60x.

Illumina sequencing library preparation

crRNA inserts were amplified from genomic DNA isolated from screens using 16 cycles of first round PCR using pooled 0-8nt staggered forward and reverse primers, treated with ExoSAP-IT (Thermo Fisher 78201.1.ML), followed by second round of PCR to

introduce Illumina unique dual indices and adaptors. Sequencing primer binding sites, unique dual indices, P5 and P7 adaptor sequences are from Illumina Adaptor Sequences Document # 1000000002694 v16. PCR amplicons were subject to size selection by magnetic beads (SPRIselect, Beckman B23318) prior to sequencing on an Illumina NovaSeq6000 using SP100 kit for Library 1 or SP500 kit for Library 2. Sequencing of plasmid libraries were performed similarly, except 7 cycles of amplification were each used for Round 1 and Round 2 PCR. The size distribution of the final library was measured on an Agilent TapeStation system. We noted that even after magnetic bead selection of Round 2 PCR-amplified Library 2 plasmid library (colonies from which were Sanger sequencing verified) and genomic DNA from screens, smaller sized fragments from non-specific PCR amplification during Illumina sequencing library preparation persisted. This might contribute to the fraction of reads that could not be mapped to our reference 6-plex array. Thus, these unmapped reads do not necessarily reflect recombination of the crRNA library constructs, though the latter could contribute as well.

Cell fitness screens

Library 1 screen: K562 cells engineered by piggyBac transposition to constitutively express denAsCas12a-KRAB or multiAsCas12a-KRAB were transduced with lentivirally packaged Library 1 constructs at MOI = ~0.15. Transduced cells were then selected using 1ug/ml puromycin for 2 days, followed by washout of puromycin. On Day 6 after transduction, initial (T0) time point was harvested, and the culture was split into 2 replicates that are separately cultured henceforth. 10 days later (T10), the final time point was harvested (8.6 total doublings for multiAsCas12a-KRAB cells, 9.15 total doublings for denasCas12a-KRAB cells). A cell coverage of >500x was maintained throughout the screen. Library 2 screen: K562 cells engineered by piggyBac transposition to constitutively express multiAsCas12a-KRAB were transduced with lentivirally packaged Library 2 constructs at MOI = ~0.15. The screen was carried out similarly as described for Library 1 screen, except the screen was carried out for 14 days (T14) or 13.5 total doublings and maintained at a cell coverage of >2000x throughout. Genomic DNA was isolated using the NucleoSpin Blood XL Maxi kit (Machery-Nagel 740950.50).

Screen data processing and analysis

Summary of library contents are in Fig. S17.

Library 1: Reads were mapped to crRNA constructs using sgcount (https://noamteyssier.github.io/sgcount/), requiring perfect match to the reference sequence.

Library 2: First, reference construct sequences were created by interspersing provided spacer and constant regions. Each construct is then given a unique construct id (CID). Each CID is then split into R1 and R2 reference sequences, which are constructed by taking the first three and last three spacer-construct pairs of the reference sequence respectively. The R2 sequence is then reverse complemented for matching against the R2 sequencing reads. Next, two hashmaps are created for the R1 and R2 spacer-construct pairs respectively, which map the R1/R2 sequences to a set of corresponding CIDs. Finally, for each R1/R2 sequencing pair, each k-mer (k = length of R1/R2 respective construct sequence) in the sequence is mapped against their respective R1/R2 hashmap. If both sequencing pairs are able to be mapped to a CID set, then the intersection of their sets is their original construct, and the total count of that CID is incremented. We implemented the above algorithm as the `casmap constructs` command in a package written in Rust, available at https://github.com/noamteyssier/casmap.

Starting from read counts, the remainder of analyses were performed using custom scripts in R. Constructs that contained <1 read per million reads (RPM) aligned to the reference library in either replicates at T0 were removed from analysis. From the constructs that meet this read coverage threshold, a pseudocount of 1 was added for each construct and the RPM re-calculated and used to obtain a fitness score (Kampmann et al. 2013):

$$\gamma = log2\left(\frac{(RPMfinal/negctrlmedianRPMfinal)}{(RPMinitial/negctrlmedianRPMinitial)}\right)/totaldoublings,$$

where RPM = read count per million reads mapped to reference (initial = at T0, final = at end of screen), negctrlmedian = median of RPM of intergenic negative control constructs, totaldoublings = total cell population doublings in the screen. For Library 1, data from a single T0 sample was used to calculate the fitness score for both replicates due to an unexpected global loss of sequencing read counts for one of two originally intended T0 replicate samples.

3' RNA-seq experiment and data analysis

Experimental procedure

3' RNA-seq was performed as part of a batch processed using a QuantSeq-Pool Sample-Barcoded 3′ mRNA-Seq Library Prep Kit for Illumina (Lexogen cat#139) in accordance with the manufacturer's instructions. Briefly, 10 ng of each purified input

RNA was used for first strand cDNA synthesis with an oligo(dT) primer containing a sample barcode and a unique molecular identifier. Subsequently, barcoded samples were pooled and used for second strand synthesis and library amplification. Amplified libraries were sequenced on an Illumina HiSeq4000 with 100 bp paired-end reads. The QuantSeq Pool data was demultiplexed and preprocessed using an implementation of pipeline originally provided by Lexogen (https://github.com/Lexogen-Tools/quantseqpool_analysis). The final outputs of this step are gene level counts for all samples (including samples from multiple projects multiplexed together).

Gene level and differential expression analysis

For generating scatter plots, normTransform function from DESeq2 (Love et al. 2014) used to normalize the raw counts and then a pseudocount of 1 added, and log2-transformed. The output was plotted in R as scatter plots.

For differential expression analysis, DESeq2 (version 1.34) default Wald-test was used to compare each targeting construct (one replicate) with non-targeting samples (two replicates). We calculated log2 transformed TPM counts and applied the threshold of 6.5 to eliminate genes with low expression. Using ggplot2, volcano plots visualized in R are then displayed for genes with log2FoldChange above or below 2.055 and p-values smaller than 0.01. The log2FoldChange cutoff was based on visually examining the concordance between two replicates of untransduced controls and manually identifying a threshold below which the log2FoldChange are poorly correlated between the replicates of the untransduced control.

Off-target analysis of spacers

To evaluate potential off-target effect of spacers, we used the crisprVerse (version 1.0.0) (Hoberecht et al. 2022) and crisprBowtie (version 1.2.0) together with other R packages including GenomicRanges (version 1.50) (Lawrence et al. 2013) and tidyverse (version 1.3.2) (Wickham et al. 2019). First, we defined dictionary of spacers as "TCCTCCAGCATCTTCCACATTCA":"HBG-2", "TTCTTCATCCCTAGCCAGCCGCC":"HBG-3", "CTTAGAAGGTTACACAGAACCAG":"HS2-1", "TGTGTAACCTTCTAAGCAAACCT":"HS2-2", "AGGTGGAGTTTTAGTCAGGTGGT":"HS2-3", "ATTAACTGATGCGTAAGGAGATC":"NT-3". Then, `runCrisprBowtie` function used with these parameters: `crisprNuclease` as `enAsCas12a`, `n_mismatches` equal 3,

`canonical` equal FALSE, and `bowtie_index` as a path to folder including pre-indexed hg38 reference genome. Thus, the results from this step allows us to assess our previously designed spacers and annotate potential off-target loci in the human genome. To annotate results, we used reference annotation GENCODE (version 34) (Frankish et al. 2021) and we defined pam_site +/- 2500 bp for each predicted off-target to overlap them with matched transcription start sites (TSS) +/- 1000 bp of all annotated genes. Results shown as annotated tables.

RT-qPCR

For the CRISPRi experiments targeting the HBG TSS or HS2 enhancer, K562 cells engineered (by lentiviral transduction at MOI ~5) for constitutive expression of multiAsCas12a-KRAB were transduced with crRNAs and sorted, followed by resuspension of ~200k to 1 million cells in 300ul RNA Lysis Buffer from the Quick-RNA Miniprep Kit (Zymo R1055) and stored in -70deg. C. RNA isolation was performed following the kit's protocols, including on-column DNase I digestion. 500ng of RNA was used as input for cDNA synthesis primed by random hexamers using the RevertAid RT Reverse Transcription Kit (Thermo fisher K1691), as per manufacturer's instructions. cDNA was diluted 1:4 with water and 2ul used as template for qPCR using 250nM primers using the SsoFast EvaGreen Supermix (BioRad 1725200) on an Applied Biosystems ViiA 7 Real Time PCR System. Data was analyzed using the ddCT method, normalized to GAPDH and no crRNA sample as reference.

Transient transfection experiments

For co-transfection experiments, transfections were performed similar to prior study (Campa et al. 2019). Briefly, the day before transfection, 100,000 HEK293T cells were seeded into wells of a 24 well plate. The following day, we transiently transfected 0.6µg of each protein construct and 0.3µg gRNA construct per well (in duplicate) in Mirus TransIT-LT1 transfection reagent according to manufacturer's instructions. Mixtures were incubated at room temperature for 30 minutes and then added in dropwise fashion into each well. 24 hours after transfection, cells were replenished with fresh media. 48 hours after transfection, BFP and GFP positive cells (indicative of successful delivery of protein and crRNA constructs) were sorted (BD FACSAria Fusion) and carried out for subsequent flow-cytometry experiments.

Data access and availability

Tables of all sequence and read counts are included as Supplementary Files. Plasmids will be made available on Addgene. Engineered cell lines will be made available upon request.

References

Adamson B, Norman TM, Jost M, Cho MY, Nuñez JK, Chen Y, Villalta JE, Gilbert LA, Horlbeck MA, Hein MY, et al. 2016. A Multiplexed Single-Cell CRISPR Screening Platform Enables Systematic Dissection of the Unfolded Protein Response. *Cell* **167**: 1867–1882.e21.

Adamson B, Norman TM, Jost M, Weissman JS. 2018. Approaches to maximize sgRNA-barcode coupling in Perturb-seq screens. *bioRxiv* 298349. https://www.biorxiv.org/content/10.1101/298349v1 (Accessed October 7, 2019).

Aguirre AJ, Meyers RM, Weir BA, Vazquez F, Zhang C-Z, Ben-David U, Cook A, Ha G, Harrington WF, Doshi MB, et al. 2016. Genomic Copy Number Dictates a Gene-Independent Cell Response to CRISPR/Cas9 Targeting. *Cancer Discov* **6**: 914–929.

Alerasool N, Leng H, Lin Z-Y, Gingras A-C, Taipale M. 2022. Identification and functional characterization of transcriptional activators in human cells. *Mol Cell* **82**: 677–695.e7.

Alerasool N, Segal D, Lee H, Taipale M. 2020. An efficient KRAB domain for CRISPRi applications in human cells. *Nat Methods* **17**: 1093–1096.

Amabile A, Migliara A, Capasso P, Biffi M, Cittaro D, Naldini L, Lombardo A. 2016. Inheritable Silencing of Endogenous Genes by Hit-and-Run Targeted Epigenetic Editing. *Cell* **167**: 219–232.e14.

Baranello L, Levens D, Gupta A, Kouzine F. 2012. The importance of being supercoiled: how DNA mechanics regulate dynamic processes. *Biochim Biophys Acta* **1819**: 632–638.

Basu VP, Song M, Gao L, Rigby ST, Hanson MN, Bambara RA. 2008. Strand transfer events during HIV-1 reverse transcription. *Virus Res* **134**: 19–38.

Blayney J, Francis H, Camellato B, Mitchell L, Stolper R, Boeke J, Higgs D, Kassouf M. 2022. Super-enhancers require a combination of classical enhancers and novel facilitator elements to drive high levels of gene expression. *bioRxiv* 2022.06.20.496856. https://www.biorxiv.org/content/10.1101/2022.06.20.496856v1.full (Accessed June 29, 2022).

Blobel GA, Higgs DR, Mitchell JA, Notani D, Young RA. 2021. Testing the super-enhancer concept. *Nat Rev Genet* **22**: 749–755.

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**. http://dx.doi.org/10.7554/eLife.55325.

1430 Breinig M, Schweitzer AY, Herianto AM, Revia S, Schaefer L, Wendler L, Cobos Galvez
1431 A, Tschaharganeh DF. 2019. Multiplexed orthogonal genome editing and
1432 transcriptional activation by Cas12a. *Nat Methods* **16**: 51–54.

1433 Campa CC, Weisbach NR, Santinha AJ, Incarnato D, Platt RJ. 2019. Multiplexed
1434 genome engineering by Cas12a and CRISPR arrays encoded on single transcripts.
1435 *Nat Methods* **16**: 887–893.

1436 Celaj A, Gebbia M, Musa L, Cote AG, Snider J, Wong V, Ko M, Fong T, Bansal P, Mellor
1437 JC, et al. 2020. Highly Combinatorial Genetic Interaction Analysis Reveals a
1438 Multi-Drug Transporter Influence Network. *Cell Syst* **10**: 25–38.e10.

1439 Cheng J, Clayton JS, Acemel RD, Zheng Y, Taylor RL, Keleş S, Franke M, Boackle SA,
1440 Harley JB, Quail E, et al. 2022. Regulatory Architecture of the RCA Gene Cluster
1441 Captures an Intragenic TAD Boundary, CTCF-Mediated Chromatin Looping and a
1442 Long-Range Intergenic Enhancer. *Front Immunol* **13**: 901747.

1443 Chen Z, Arai E, Khan O, Zhang Z, Ngiow SF, He Y, Huang H, Manne S, Cao Z, Baxter
1444 AE, et al. 2021. In vivo CD8+ T cell CRISPR screening reveals control by Fli1 in
1445 infection and cancer. *Cell* **184**: 1262–1280.e22.

1446 Chow RD, Wang G, Ye L, Codina A, Kim HR, Shen L, Dong MB, Errami Y, Chen S.
1447 2019. In vivo profiling of metastatic double knockouts through CRISPR-Cpf1
1448 screens. *Nat Methods* **16**: 405–408.

1449 Cleary B, Regev A. 2020. The necessity and power of random, under-sampled
1450 experiments in biology. *arXiv [q-bioQM]*. http://arxiv.org/abs/2012.12961.

1451 Clement K, Rees H, Canver MC, Gehrke JM, Farouni R, Hsu JY, Cole MA, Liu DR,
1452 Joung JK, Bauer DE, et al. 2019. CRISPResso2 provides accurate and rapid
1453 genome editing sequence analysis. *Nat Biotechnol* **37**: 224–226.

1454 Cofsky JC, Karandur D, Huang CJ, Witte IP, Kuriyan J, Doudna JA. 2020.
1455 CRISPR-Cas12a exploits R-loop asymmetry to form double-strand breaks. *Elife* **9**.
1456 http://dx.doi.org/10.7554/eLife.55143.

1457 Costanzo M, Kuzmin E, van Leeuwen J, Mair B, Moffat J, Boone C, Andrews B. 2019.
1458 Global Genetic Networks and the Genotype-to-Phenotype Relationship. *Cell* **177**:
1459 85–100.

1460 Creutzburg SCA, Wu WY, Mohanraju P, Swartjes T, Alkan F, Gorodkin J, Staals RHJ,
1461 van der Oost J. 2020. Good guide, bad guide: spacer sequence-dependent
1462 cleavage efficiency of Cas12a. *Nucleic Acids Res* **48**: 3228–3243.

1463 Datlinger P, Rendeiro AF, Schmidl C, Krausgruber T, Traxler P, Klughammer J, Schuster
1464 LC, Kuchler A, Alpar D, Bock C. 2017. Pooled CRISPR screening with single-cell
1465 transcriptome readout. *Nat Methods* **14**: 297–301.

1466 Dede M, McLaughlin M, Kim E, Hart T. 2020. Multiplex enCas12a screens detect
1467 functional buffering among paralogs otherwise masked in monogenic Cas9
1468 knockout screens. *Genome Biol* **21**: 262.

1469 DelRosso N, Tycko J, Suzuki P, Andrews C, Aradhana, Mukund A, Liongson I, Ludwig
1470 C, Spees K, Fordyce P, et al. 2023. Large-scale mapping and mutagenesis of
1471 human transcriptional effector domains. *Nature* **616**: 365–372.

1472 DeWeirdt PC, McGee AV, Zheng F, Nwolah I, Hegde M, Doench JG. 2022. Accounting
1473 for small variations in the tracrRNA sequence improves sgRNA activity predictions
1474 for CRISPR screening. *bioRxiv* 2022.06.27.497780.
1475 https://www.biorxiv.org/content/10.1101/2022.06.27.497780v1 (Accessed August 6,
1476 2022).

1477 DeWeirdt PC, Sanson KR, Sangree AK, Hegde M, Hanna RE, Feeley MN, Griffith AL,
1478 Teng T, Borys SM, Strand C, et al. 2021. Optimization of AsCas12a for
1479 combinatorial genetic screens in human cells. *Nat Biotechnol* **39**: 94–104.

1480 Dho SH, Lim JC, Kim LK. 2018. Beyond the Role of CD55 as a Complement
1481 Component. *Immune Netw* **18**: e11.

1482 Dixit A, Parnas O, Li B, Chen J, Fulco CP, Jerby-Arnon L, Marjanovic ND, Dionne D,
1483 Burks T, Raychowdhury R, et al. 2016. Perturb-Seq: Dissecting Molecular Circuits
1484 with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens. *Cell* **167**:
1485 1853–1866.e17.

1486 Doench JG. 2018. Am I ready for CRISPR? A user's guide to genetic screens. *Nat Rev*
1487 *Genet* **19**: 67–80.

1488 Domingo J, Baeza-Centurion P, Lehner B. 2019. The Causes and Consequences of
1489 Genetic Interactions (Epistasis). *Annu Rev Genomics Hum Genet* **20**: 433–460.

1490 Domingo J, Diss G, Lehner B. 2018. Pairwise and higher-order genetic interactions
1491 during the evolution of a tRNA. *Nature* **558**: 117–121.

1492 Dorfman R. 1943. The Detection of Defective Members of Large Populations. *aoms* **14**:
1493 436–440.

1494 Du D. 1993. *Combinatorial Group Testing and Its Applications*. World Scientific.

1495 Ecco G, Imbeault M, Trono D. 2017. KRAB zinc finger proteins. *Development* **144**:
1496 2719–2729.

1497 Egan ES, Jiang RHY, Moechtar MA, Barteneva NS, Weekes MP, Nobre LV, Gygi SP,
1498 Paulo JA, Frantzreb C, Tani Y, et al. 2015. Malaria. A forward genetic screen
1499 identifies erythrocyte CD55 as essential for Plasmodium falciparum invasion.
1500 *Science* **348**: 711–714.

ENCODE Project Consortium, Moore JE, Purcaro MJ, Pratt HE, Epstein CB, Shoresh N, Adrian J, Kawli T, Davis CA, Dobin A, et al. 2020. Expanded encyclopaedias of DNA elements in the human and mouse genomes. *Nature* **583**: 699–710.

Esmaeili Anvar N, Lin C, Wilson LL, Sangree AK, Ma X, Colic M, Doench JG, Hart T. 2023. Combined genome-scale fitness and paralog synthetic lethality screens with just 44k clones: the IN4MER CRISPR/Cas12a multiplex knockout platform. *bioRxiv* 2023.01.03.522655.

Ewen-Campen B, Mohr SE, Hu Y, Perrimon N. 2017. Accessing the Phenotype Gap: Enabling Systematic Investigation of Paralog Functional Complexity with CRISPR. *Dev Cell* **43**: 6–9.

Feldman D, Singh A, Schmid-Burgk JL, Carlson RJ, Mezger A, Garrity AJ, Zhang F, Blainey PC. 2019. Optical Pooled Screens in Human Cells. *Cell* **179**: 787–799.e17.

Fonfara I, Richter H, Bratovič M, Le Rhun A, Charpentier E. 2016. The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. *Nature* **532**: 517–521.

Frankish A, Diekhans M, Jungreis I, Lagarde J, Loveland JE, Mudge JM, Sisu C, Wright JC, Armstrong J, Barnes I, et al. 2021. GENCODE 2021. *Nucleic Acids Res* **49**: D916–D923.

Fulco CP, Munschauer M, Anyoha R, Munson G, Grossman SR, Perez EM, Kane M, Cleary B, Lander ES, Engreitz JM. 2016. Systematic mapping of functional enhancer-promoter connections with CRISPR interference. *Science* **354**: 769–773.

Fulco CP, Nasser J, Jones TR, Munson G, Bergman DT, Subramanian V, Grossman SR, Anyoha R, Doughty BR, Patwardhan TA, et al. 2019. Activity-by-contact model of enhancer-promoter regulation from thousands of CRISPR perturbations. *Nat Genet* **51**: 1664–1669.

Gier RA, Budinich KA, Evitt NH, Cao Z, Freilich ES, Chen Q, Qi J, Lan Y, Kohli RM, Shi J. 2020. High-performance CRISPR-Cas12a genome editing for combinatorial genetic screening. *Nat Commun* **11**: 3455.

Gilbert LA, Horlbeck MA, Adamson B, Villalta JE, Chen Y, Whitehead EH, Guimaraes C, Panning B, Ploegh HL, Bassik MC, et al. 2014. Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* **159**: 647–661.

Gilbert LA, Larson MH, Morsut L, Liu Z, Brar GA, Torres SE, Stern-Ginossar N, Brandman O, Whitehead EH, Doudna JA, et al. 2013. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154**: 442–451.

Gonatopoulos-Pournatzis T, Aregger M, Brown KR, Farhangmehr S, Braunschweig U, Ward HN, Ha KCH, Weiss A, Billmann M, Durbic T, et al. 2020. Genetic interaction mapping and exon-resolution functional genomics with a hybrid Cas9-Cas12a

platform. *Nat Biotechnol* **38**: 638–648.

Griffith AL, Zheng F, McGee AV, Miller NW, Szegletes ZM, Reint G, Gademann F, Nwolah I, Hegde M, Liu YV, et al. 2023. Optimization of Cas12a for multiplexed genome-scale transcriptional activation. *Cell Genomics* **0**. http://www.cell.com.stanford.idm.oclc.org/article/S2666979X23001842/abstract (Accessed September 7, 2023).

Guo LY, Bian J, Davis AE, Liu P, Kempton HR, Zhang X, Chemparathy A, Gu B, Lin X, Rane DA, et al. 2022. Multiplexed genome regulation in vivo with hyper-efficient Cas12a. *Nat Cell Biol* **24**: 590–600.

Haapaniemi E, Botla S, Persson J, Schmierer B, Taipale J. 2018. CRISPR-Cas9 genome editing induces a p53-mediated DNA damage response. *Nat Med* **24**: 927–930.

Hillmen P, Hall C, Marsh JCW, Elebute M, Bombara MP, Petro BE, Cullen MJ, Richards SJ, Rollins SA, Mojcik CF, et al. 2004. Effect of eculizumab on hemolysis and transfusion requirements in patients with paroxysmal nocturnal hemoglobinuria. *N Engl J Med* **350**: 552–559.

Hoberecht L, Perampalam P, Lun A, Fortin J-P. 2022. A comprehensive Bioconductor ecosystem for the design of CRISPR guide RNAs across nucleases and technologies. *Nat Commun* **13**: 6568.

Horlbeck MA, Gilbert LA, Villalta JE, Adamson B, Pak RA, Chen Y, Fields AP, Park CY, Corn JE, Kampmann M, et al. 2016a. Compact and highly active next-generation libraries for CRISPR-mediated gene repression and activation. *Elife* **5**. http://dx.doi.org/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**. http://dx.doi.org/10.7554/eLife.12677.

Ihry RJ, Worringer KA, Salick MR, Frias E, Ho D, Theriault K, Kommineni S, Chen J, Sondey M, Ye C, et al. 2018. p53 inhibits CRISPR-Cas9 engineering in human pluripotent stem cells. *Nat Med* **24**: 939–946.

Jeon Y, Choi YH, Jang Y, Yu J, Goo J, Lee G, Jeong YK, Lee SH, Kim I-S, Kim J-S, et al. 2018. Direct observation of DNA target searching and cleavage by CRISPR-Cas12a. *Nat Commun* **9**: 2777.

Joberty G, Fälth-Savitski M, Paulmann M, Bösche M, Doce C, Cheng AT, Drewes G, Grandi P. 2020. A Tandem Guide RNA-Based Strategy for Efficient CRISPR Gene Editing of Cell Populations with Low Heterogeneity of Edited Alleles. *CRISPR J* **3**: 123–134.

Kampmann M, Bassik MC, Weissman JS. 2013. Integrated platform for genome-wide

screening and construction of high-density genetic interaction maps in mammalian cells. *Proc Natl Acad Sci U S A* **110**: E2317–26.

Kent WJ. 2002. BLAT--the BLAST-like alignment tool. *Genome Res* **12**: 656–664.

Kim HK, Min S, Song M, Jung S, Choi JW, Kim Y, Lee S, Yoon S, Kim HH. 2018. Deep learning improves prediction of CRISPR-Cpf1 guide RNA activity. *Nat Biotechnol* **36**: 239–241.

Kleinstiver BP, Sousa AA, Walton RT, Tak YE, Hsu JY, Clement K, Welch MM, Horng JE, Malagon-Lopez J, Scarfò I, et al. 2019. Engineered CRISPR-Cas12a variants with increased activities and improved targeting ranges for gene, epigenetic and base editing. *Nat Biotechnol* **37**: 276–282.

Knott GJ, Thornton BW, Lobba MJ, Liu J-J, Al-Shayeb B, Watters KE, Doudna JA. 2019. Broad-spectrum enzymatic inhibition of CRISPR-Cas12a. *Nat Struct Mol Biol* **26**: 315–321.

Konermann S, Brigham MD, Trevino AE, Joung J, Abudayyeh OO, Barcena C, Hsu PD, Habib N, Gootenberg JS, Nishimasu H, et al. 2015. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* **517**: 583–588.

Kuzmin E, VanderSluis B, Wang W, Tan G, Deshpande R, Chen Y, Usaj M, Balint A, Mattiazzi Usaj M, van Leeuwen J, et al. 2018. Systematic analysis of complex genetic interactions. *Science* **360**. http://dx.doi.org/10.1126/science.aao1729.

Kvon EZ, Waymack R, Gad M, Wunderlich Z. 2021. Enhancer redundancy in development and disease. *Nat Rev Genet* **22**: 324–336.

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**: R25.

Lawrence M, Huber W, Pagès H, Aboyoun P, Carlson M, Gentleman R, Morgan MT, Carey VJ. 2013. Software for computing and annotating genomic ranges. *PLoS Comput Biol* **9**: e1003118.

Li K, Liu Y, Cao H, Zhang Y, Gu Z, Liu X, Yu A, Kaphle P, Dickerson KE, Ni M, et al. 2020. Interrogation of enhancer function by enhancer-targeting CRISPR epigenetic editing. *Nat Commun* **11**: 485.

Lin X, Liu Y, Liu S, Zhu X, Wu L, Zhu Y, Zhao D, Xu X, Chemparathy A, Wang H, et al. 2022. Nested epistasis enhancer networks for robust genome regulation. *Science* **377**: 1077–1085.

Liu J, Srinivasan S, Li C-Y, Ho I-L, Rose J, Shaheen M, Wang G, Yao W, Deem A, Bristow C, et al. 2019. Pooled library screening with multiplexed Cpf1 library. *Nat Commun* **10**: 3144.

Liu Y, Han J, Chen Z, Wu H, Dong H, Nie G. 2017. Engineering cell signaling using tunable CRISPR-Cpf1-based transcription factors. *Nat Commun* **8**: 2095.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 550.

Meyers RM, Bryan JG, McFarland JM, Weir BA, Sizemore AE, Xu H, Dharia NV, Montgomery PG, Cowley GS, Pantel S, et al. 2017. Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells. *Nat Genet* **49**: 1779–1784.

Mukund AX, Tycko J, Allen SJ, Robinson SA, Andrews C, Sinha J, Ludwig CH, Spees K, Bassik MC, Bintu L. 2023. High-throughput functional characterization of combinations of transcriptional activators and repressors. *Cell Syst* **0**. http://dx.doi.org/10.1016/j.cels.2023.07.001 (Accessed August 5, 2023).

Nakamura M, Ivec AE, Gao Y, Qi LS. 2021. Durable CRISPR-Based Epigenetic Silencing. *BioDesign Research* **2021**. https://spj.sciencemag.org/journals/bdr/2021/9815820/ (Accessed August 28, 2023).

Naqvi MM, Lee L, Montaguth OET, Diffin FM, Szczelkun MD. 2022. CRISPR-Cas12a-mediated DNA clamping triggers target-strand cleavage. *Nat Chem Biol* **18**: 1014–1022.

Norman TM, Horlbeck MA, Replogle JM, Ge AY, Xu A, Jost M, Gilbert LA, Weissman JS. 2019. Exploring genetic interaction manifolds constructed from rich single-cell phenotypes. *Science* **365**: 786–793.

Nuñez JK, Chen J, Pommier GC, Cogan JZ, Replogle JM, Adriaens C, Ramadoss GN, Shi Q, Hung KL, Samelson AJ, et al. 2021. Genome-wide programmable transcriptional memory by CRISPR-based epigenome editing. *Cell* **184**: 2503–2519.e17.

O'Geen H, Ren C, Nicolet CM, Perez AA, Halmai J, Le VM, Mackay JP, Farnham PJ, Segal DJ. 2017. dCas9-based epigenome editing suggests acquisition of histone methylation is not sufficient for target gene repression. *Nucleic Acids Res* **45**: 9901–9916.

O'Geen H, Tomkova M, Combs JA, Tilley EK, Segal DJ. 2022. Determinants of heritable gene silencing for KRAB-dCas9 + DNMT3 and Ezh2-dCas9 + DNMT3 hit-and-run epigenome editing. *Nucleic Acids Res* **50**: 3239–3253.

Osterwalder M, Barozzi I, Tissières V, Fukuda-Yuzawa Y, Mannion BJ, Afzal SY, Lee EA, Zhu Y, Plajzer-Frick I, Pickle CS, et al. 2018. Enhancer redundancy provides phenotypic robustness in mammalian development. *Nature* **554**: 239–243.

Parrish PCR, Thomas JD, Gabel AM, Kamlapurkar S, Bradley RK, Berger AH. 2021.

1647 Discovery of synthetic lethal and tumor suppressor paralog pairs in the human
1648 genome. *Cell Rep* **36**: 109597.

1649 Przybyla L, Gilbert LA. 2022. A new era in functional genomics screens. *Nat Rev Genet*
1650 **23**: 89–103.

1651 Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, Lim WA. 2013.
1652 Repurposing CRISPR as an RNA-guided platform for sequence-specific control of
1653 gene expression. *Cell* **152**: 1173–1183.

1654 Reilly SK, Gosai SJ, Gutierrez A, Mackay-Smith A, Ulirsch JC, Kanai M, Mouri K,
1655 Berenzy D, Kales S, Butler GM, et al. 2021. Direct characterization of cis-regulatory
1656 elements and functional dissection of complex genetic associations using
1657 HCR–FlowFISH. *Nat Genet* **53**: 1166–1176.

1658 Ren X, Wang M, Li B, Jamieson K, Zheng L, Jones IR, Li B, Takagi MA, Lee J,
1659 Maliskova L, et al. 2021. Parallel characterization of cis-regulatory elements for
1660 multiple genes using CRISPRpath. *Sci Adv* **7**: eabi4360.

1661 Replogle JM, Bonnar JL, Pogson AN, Liem CR, Maier NK, Ding Y, Russell BJ, Wang X,
1662 Leng K, Guna A, et al. 2022a. Maximizing CRISPRi efficacy and accessibility with
1663 dual-sgRNA libraries and optimal effectors. *bioRxiv* 2022.07.13.499814.
1664 https://www.biorxiv.org/content/10.1101/2022.07.13.499814v2 (Accessed
1665 December 7, 2022).

1666 Replogle JM, Bonnar JL, Pogson AN, Liem CR, Maier NK, Ding Y, Russell BJ, Wang X,
1667 Leng K, Guna A, et al. 2022b. Maximizing CRISPRi efficacy and accessibility with
1668 dual-sgRNA libraries and optimal effectors. *Elife* **11**.
1669 http://dx.doi.org/10.7554/eLife.81856.

1670 Replogle JM, Norman TM, Xu A, Hussmann JA, Chen J, Cogan JZ, Meer EJ, Terry JM,
1671 Riordan DP, Srinivas N, et al. 2020. Combinatorial single-cell CRISPR screens by
1672 direct guide RNA capture and targeted sequencing. *Nat Biotechnol* **38**: 954–961.

1673 Sack LM, Davoli T, Xu Q, Li MZ, Elledge SJ. 2016. Sources of Error in Mammalian
1674 Genetic Screens. *G3* **6**: 2781–2790.

1675 Schellenberger V, Wang C-W, Geething NC, Spink BJ, Campbell A, To W, Scholle MD,
1676 Yin Y, Yao Y, Bogin O, et al. 2009. A recombinant polypeptide extends the in vivo
1677 half-life of peptides and proteins in a tunable manner. *Nat Biotechnol* **27**:
1678 1186–1190.

1679 Schraivogel D, Gschwind AR, Milbank JH, Leonce DR, Jakob P, Mathur L, Korbel JO,
1680 Merten CA, Velten L, Steinmetz LM. 2020. Targeted Perturb-seq enables
1681 genome-scale genetic screens in single cells. *Nat Methods* **17**: 629–635.

1682 Shakya B, Patel SD, Tani Y, Egan ES. 2021. Erythrocyte CD55 mediates the
1683 internalization of Plasmodium falciparum parasites. *Elife* **10**.

http://dx.doi.org/10.7554/eLife.61516.

Singh D, Mallon J, Poddar A, Wang Y, Tippana R, Yang O, Bailey S, Ha T. 2018. Real-time observation of DNA target interrogation and product release by the RNA-guided endonuclease CRISPR Cpf1 (Cas12a). *Proc Natl Acad Sci U S A* **115**: 5444–5449.

Srinivas N, Ouldridge TE, Sulc P, Schaeffer JM, Yurke B, Louis AA, Doye JPK, Winfree E. 2013. On the biophysics and kinetics of toehold-mediated DNA strand displacement. *Nucleic Acids Res* **41**: 10641–10658.

Stegmeier F, Hu G, Rickles RJ, Hannon GJ, Elledge SJ. 2005. A lentiviral microRNA-based system for single-copy polymerase II-regulated RNA interference in mammalian cells. *Proc Natl Acad Sci U S A* **102**: 13212–13217.

Swarts DC, Jinek M. 2019. Mechanistic Insights into the cis- and trans-Acting DNase Activities of Cas12a. *Mol Cell* **73**: 589–600.e4.

Takahashi K, Yamanaka S. 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**: 663–676.

Taylor MB, Ehrenreich IM. 2014. Genetic interactions involving five or more genes contribute to a complex trait in yeast. *PLoS Genet* **10**: e1004324.

Taylor MB, Ehrenreich IM. 2015. Transcriptional Derepression Uncovers Cryptic Higher-Order Genetic Interactions. *PLoS Genet* **11**: e1005606.

Tycko J, Wainberg M, Marinov GK, Ursu O, Hess GT, Ego BK, Aradhana, Li A, Truong A, Trevino AE, et al. 2019. Mitigation of off-target toxicity in CRISPR-Cas9 screens for essential non-coding elements. *Nat Commun* **10**: 4063.

Van MV, Fujimori T, Bintu L. 2021. Nanobody-mediated control of gene expression and epigenetic memory. *Nat Commun* **12**: 537.

Wessels H-H, Méndez-Mancilla A, Hao Y, Papalexi E, Mauck WM 3rd, Lu L, Morris JA, Mimitou EP, Smibert P, Sanjana NE, et al. 2023. Efficient combinatorial targeting of RNA transcripts in single cells with Cas13 RNA Perturb-seq. *Nat Methods* **20**: 86–94.

Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, Grolemund G, Hayes A, Henry L, Hester J, et al. 2019. Welcome to the tidyverse. *J Open Source Softw* **4**: 1686.

Wong ASL, Choi GCG, Cheng AA, Purcell O, Lu TK. 2015. Massively parallel high-order combinatorial genetics in human cells. *Nat Biotechnol* **33**: 952–961.

Wong ASL, Choi GCG, Lu TK. 2016. Deciphering Combinatorial Genetics. *Annu Rev Genet* **50**: 515–538.

1719 Xie S, Duan J, Li B, Zhou P, Hon GC. 2017. Multiplexed Engineering and Analysis of
1720 Combinatorial Enhancer Activity in Single Cells. *Mol Cell* **66**: 285–299.e5.

1721 Yamano T, Nishimasu H, Zetsche B, Hirano H, Slaymaker IM, Li Y, Fedorova I, Nakane
1722 T, Makarova KS, Koonin EV, et al. 2016. Crystal Structure of Cpf1 in Complex with
1723 Guide RNA and Target DNA. *Cell* **165**: 949–962.

1724 Yao D, Binan L, Bezney J, Simonton B, Freedman J, Frangieh CJ, Dey K,
1725 Geiger-Schuller K, Eraslan B, Gusev A, et al. 2023. Compressed Perturb-seq:
1726 highly efficient screens for regulatory circuits using random composite
1727 perturbations. *bioRxiv* 2023.01.23.525200.

1728 Yin J-A, Frick L, Scheidmann MC, Trevisan C, Dhingra A, Spinelli A, Wu Y, Yao L, Vena
1729 DL, De Cecco E, et al. 2022. Robust and Versatile Arrayed Libraries for Human
1730 Genome-Wide CRISPR Activation, Deletion and Silencing. *bioRxiv*
1731 2022.05.25.493370. https://www.biorxiv.org/content/10.1101/2022.05.25.493370v1
1732 (Accessed July 10, 2022).

1733 Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS,
1734 Essletzbichler P, Volz SE, Joung J, van der Oost J, Regev A, et al. 2015. Cpf1 is a
1735 single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* **163**:
1736 759–771.

1737 Zetsche B, Heidenreich M, Mohanraju P, Fedorova I, Kneppers J, DeGennaro EM,
1738 Winblad N, Choudhury SR, Abudayyeh OO, Gootenberg JS, et al. 2017. Multiplex
1739 gene editing by CRISPR-Cpf1 using a single crRNA array. *Nat Biotechnol* **35**:
1740 31–34.

1741 Zhang L, Sun R, Yang M, Peng S, Cheng Y, Chen C. 2019. Conformational Dynamics
1742 and Cleavage Sites of Cas12a Are Modulated by Complementarity between crRNA
1743 and DNA. *iScience* **19**: 492–503.

1744 Zhang L, Zuris JA, Viswanathan R, Edelstein JN, Turk R, Thommandru B, Rube HT,
1745 Glenn SE, Collingwood MA, Bode NM, et al. 2021. AsCas12a ultra nuclease
1746 facilitates the rapid generation of therapeutic cell medicines. *Nat Commun* **12**:
1747 3908.

1748 Zhou B, Ho SS, Greer SU, Zhu X, Bell JM, Arthur JG, Spies N, Zhang X, Byeon S,
1749 Pattni R, et al. 2019. Comprehensive, integrated, and phased whole-genome
1750 analysis of the primary ENCODE cell line K562. *Genome Res* **29**: 472–484.

1751 Zhou P, Chan BKC, Wan YK, Yuen CTL, Choi GCG, Li X, Tong CSW, Zhong SSW, Sun
1752 J, Bao Y, et al. 2020. A Three-Way Combinatorial CRISPR Screen for Analyzing
1753 Interactions among Druggable Targets. *Cell Rep* **32**: 108020.

1754

Table of Contents for Main and Supplemental Figures

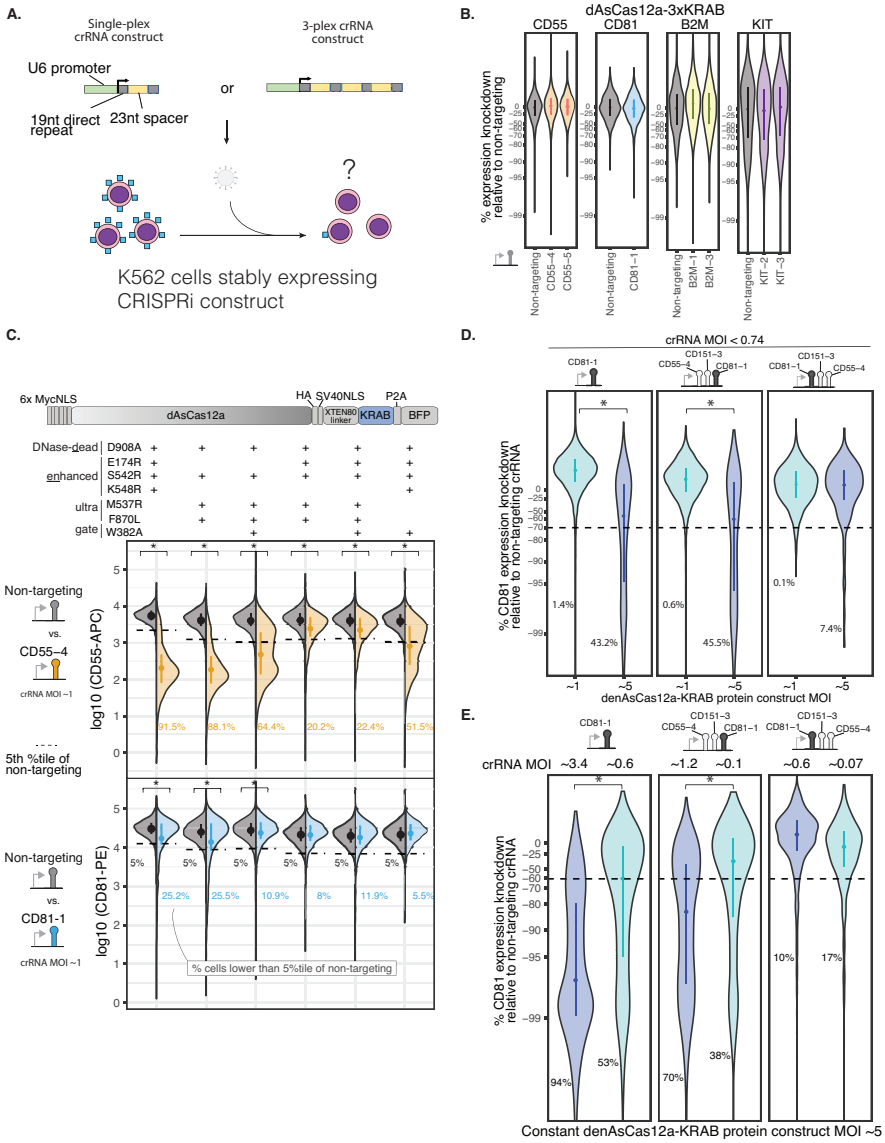


Figure 1 – dAsCas12a-KRAB variants are dose-limited and weak in CRISPRi activity when delivered lentivirally, despite incorporating state-of-the-art optimizations.

A) Schematic for assaying CRISPRi activity of Cas12a constructs using lentivirally transduced single-plex or 3-plex crRNAs targeting cell surface marker genes assayed by antibody staining and flow cytometry.

B) K562 cells constitutively expressing dAsCas12a-KRABx3 (Campa et al., 2019) were lentivirally transduced with single crRNAs targeting CD55, CD81, B2M, KIT, or a non-targeting crRNA, and assayed by flow cytometry 6 days after crRNA transduction.

C) A panel of Cas12a variants harboring combinations of mutations are tested using crCD55-4 and crCD81-1 using the fusion protein domain architecture shown. Both Cas12a fusion protein and crRNA constructs are delivered by lentiviral transduction. D908A is a mutation in the RuvC catalytic triad that renders Cas12a DNase-inactive (Yamano et al., 2016; Zetsche et al., 2015). Other mutations are described in detail in the main text. Shown are single-cell distributions of target gene expression assayed by flow cytometry 6 days after crRNA transduction for one of 3 independent replicates. One-sided Wilcoxon rank-sum test was performed comparing the single-cell distributions of the non-targeting control vs. the corresponding targeting crRNA; asterisk indicates $p<0.01$. Additional replicates and results for additional crRNA constructs (up to 3-plex crRNA constructs) are summarized in Fig. S6.

D) Analysis of CD81 knockdown in cells lentivirally transduced with denAsCas12a-KRAB protein construct at MOI ∼1 vs. MOI ∼5, while maintaining constant crRNA MOI (<0.74) for each crRNA construct. CD81 expression was assayed by flow cytometry 6 days after crRNA transduction. Shown are single-cell distributions of target gene expression knockdown as a percentage of non-targeting control for one of 3-6 biological replicates for each crRNA construct. Median and interquartile range are shown for each distribution. Percentage of cells below the 5th percentile (dashed line) of non-targeting crRNA are shown. One-sided Wilcoxon rank-sum test was performed on single-cell distributions for this replicate; asterisk indicates $p<0.01$. Summaries of all replicates shown in Fig. S7A.

E) Similar to D, but maintaining constant denAsCas12a-KRAB protein construct MOI at ∼5, while crRNA MOI is varied as indicated. CD81 expression was assayed by flow cytometry 10 days after crRNA transduction. Shown are single-cell distributions of CD81 knockdown for one of two biological replicates. One-sided Wilcoxon rank-sum test was performed on single-cell distributions for each replicate; asterisk indicates $p<0.01$. Additional replicate shown in Fig. S7B.

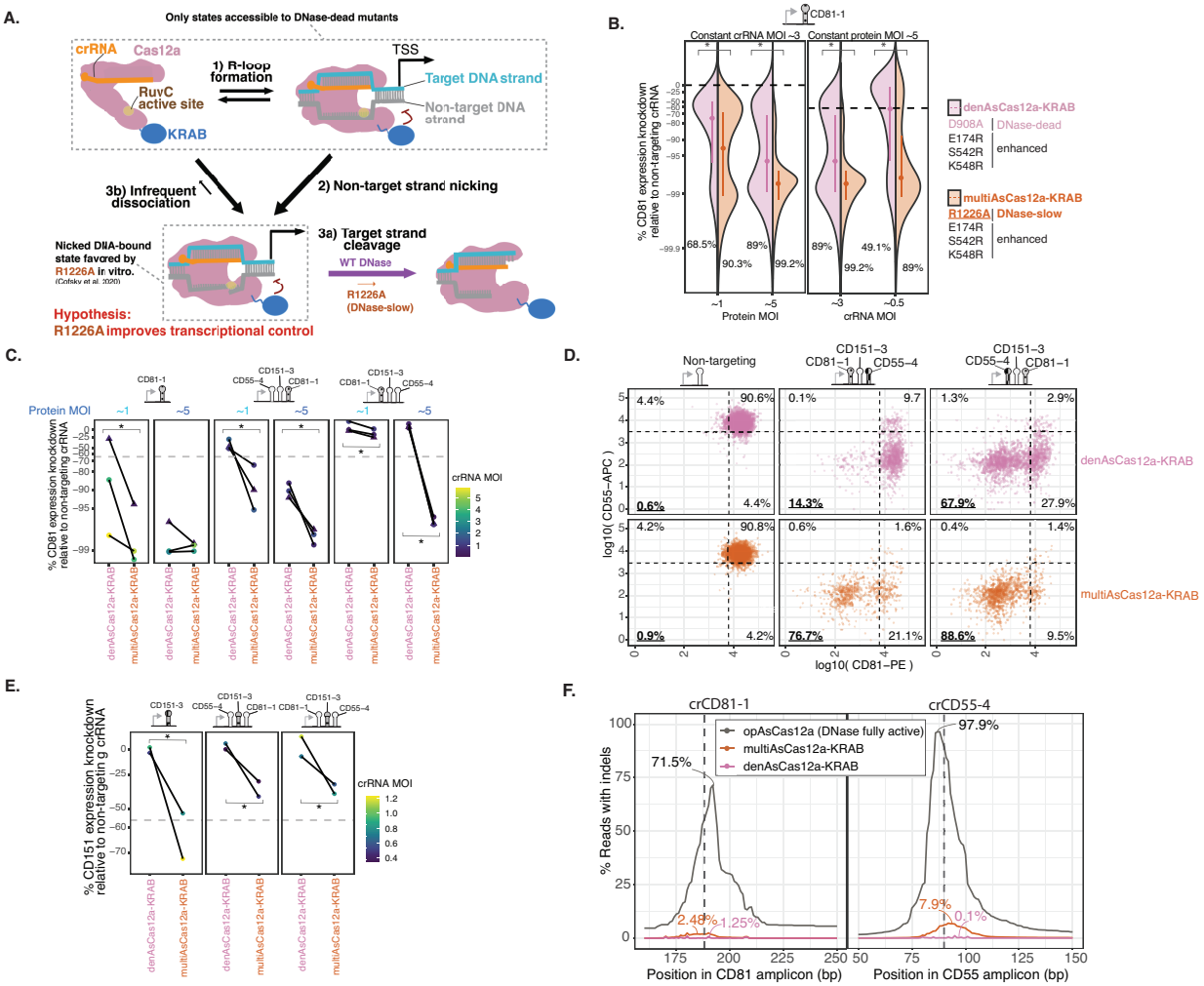


Figure 2 – multiAsCas12a-KRAB (R1226A/E174R/S542R/K548R), an engineered variant that favors a nicked DNA intermediate, substantially improves lentivirally delivered CRISPRi activity.

A) Model of Cas12a DNA binding and cleavage states for wildtype DNase vs. R1226A mutant based on prior in vitro studies as detailed in main text. Sizes of arrows qualitatively reflect reaction rates.

B) Comparison of denCas12a-KRAB (D908A/E174R/S542R/K548R) vs. multiAsCas12a-KRAB (R1226A/E174R/S542R/K548R) in CRISPRi knockdown of CD81 using crCD81-1. CD81 expression assayed by flow cytometry 10 days after crRNA transduction. Left panel: Holding crRNA MOI constant at ∼3 while testing protein MOI ∼1 vs. ∼5. Right panel: Holding protein MOI constant at ∼5 while testing crRNA MOI at ∼3 vs. ∼0.5. Asterisks indicate $p < 0.01$ for one-sided Wilcoxon rank-sum test of single-cell distributions. Percentage of cells below the 5th percentile of non-targeting crRNA control (dashed line) is shown for each condition. One biological replicate is shown for each condition; additional replicates shown in Fig. S9.

C) Comparison of CD81 knockdown by lentivirally delivered denAsCas12a-KRAB vs. multiAsCas12a-KRAB at protein MOI ∼1 vs. ∼5 across a panel of single and 3-plex crRNA constructs, while holding constant crRNA MOI for each paired fusion protein comparison for each crRNA construct. Dashed gray line indicates 5th percentile of non-targeting crRNA control. crRNA MOI indicated by color scale. Lines connect paired replicates. One-sided Wilcoxon rank-sum tests were performed on single-cell distributions for each replicate, and asterisk denotes $p<0.01$ for all paired replicates within each condition. Dots indicate flow cytometry measurement 10 days after crRNA transduction; triangles indicate flow cytometry measurement 16 days after crRNA transduction.

D) Same as C but showing scatter plot of CD55-APC and CD81-PE antibody co-staining signals on flow cytometry performed 16 days after transduction of the indicated crRNA constructs in K562 cells lentivirally transduced with denAsCas12a-KRAB vs. multiAsCas12a-KRAB at protein MOI ∼5. Quadrants drawn based on the 5th percentile of non-targeting controls and the percentage of cells in each quadrant denoted.

E) K562 cells piggyBac-engineered to constitutively express denAsCas12a-KRAB or multiAsCas12a-KRAB were transduced with the indicated crRNA constructs, followed by measurement of CD151 expression by antibody staining and flow cytometry 13 days after crRNA transduction. Median CD151 expression knockdown relative to non-targeting control is shown for each individual replicate. Dashed gray line indicates 5th percentile of non-targeting crRNA control. crRNA MOI indicated by color scale. One-sided Wilcoxon rank-sum test was performed for single-cell expression distributions comparing denAsCas12a-KRAB vs. multiAsCas12a-KRAB for each replicate; asterisk indicates $p<0.001$ for all replicate-level comparisons.

F) Indel quantification from PCR amplicons surrounding target sites of crCD81-1 and crCD55-4 in cells lentivirally transduced at protein MOI ∼5 for denAsCas12a-KRAB and multiAsCas12a-KRAB. Cells lentivirally transduced with opAsCas12a (DNase fully active) are shown for comparison. Percent of reads containing indels at each base position within the amplicon is plotted, with labels indicating maximum indel frequency observed across all bases within the amplicon.

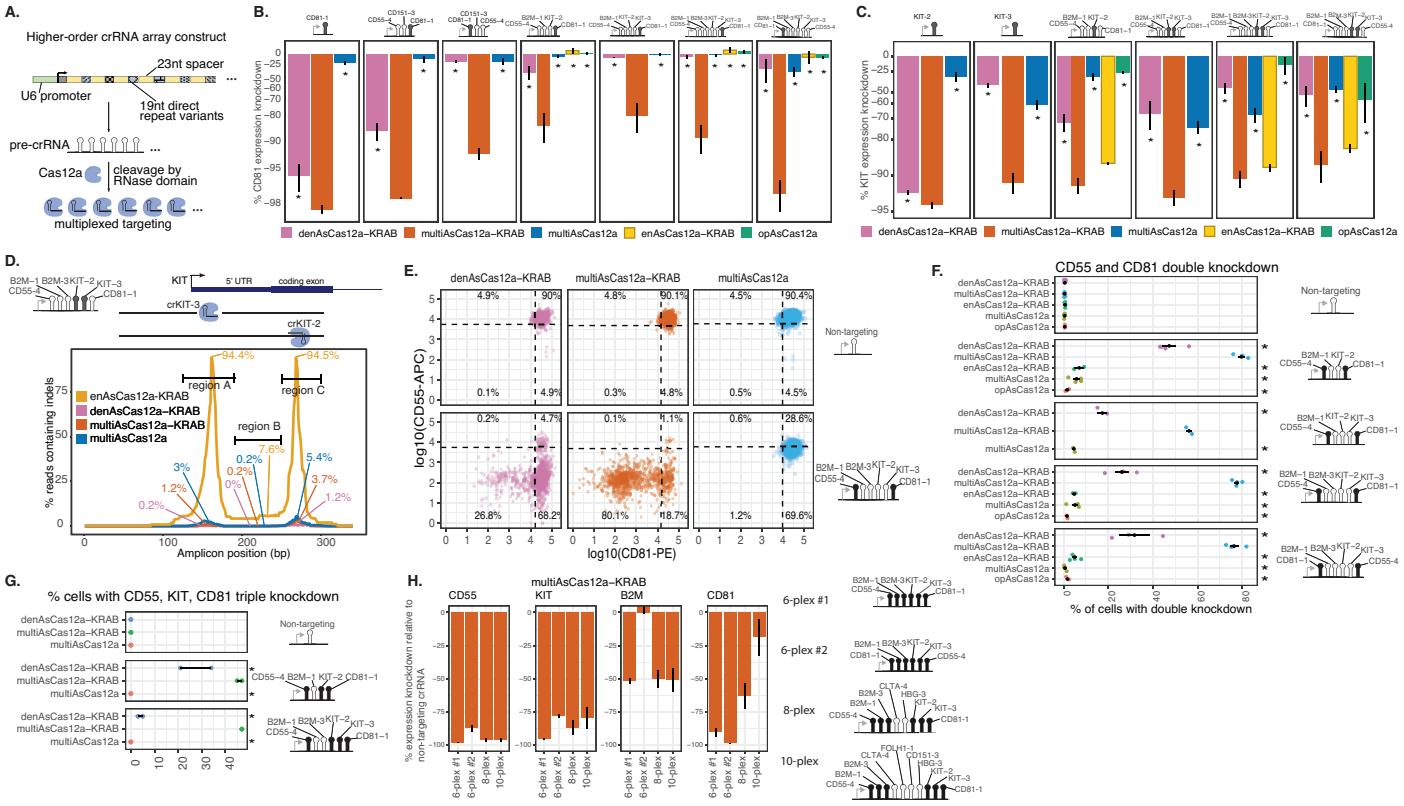


Figure 3 – multiAsCas12a-KRAB enables multi-gene CRISPRi perturbations using higher-order arrayed crRNA lentiviral constructs
A) Schematic for higher-order crRNA expression constructs. 23nt spacers are interspersed by 19nt direct repeat variants (DeWeirdt et al., 2020) uniquely assigned to each position within the array. **B)** Flow cytometry analysis of CD81 expression knockdown by antibody staining 6 days after transduction of the indicated lentiviral crRNA constructs in K562 cells engineered to constitutively express the specified fusion protein construct. Shown are averages of median single-cell expression knockdown from 2-5 biological replicates for each crRNA construct, with error bars indicating SEM. One-sided Wilcoxon rank-sum test was performed for differences in single-cell expression distributions for each fusion protein against multiAsCas12a-KRAB for each individual replicate. Asterisk indicates $p < 0.01$ for all replicates for a given pairwise comparison. **C)** Same as B, but shown for KIT expression knockdown. **D)** Indel quantification for the indicated fusion protein constructs using a 6-plex crRNA construct encoding crKIT-2 and crKIT-3 that target opposite strands at sites spaced 95bp apart near the KIT TSS. Following crRNA transduction, cells were sorted on day 3 for GFP marker on the crRNA construct, and the 340bp genomic region surrounding both crRNA binding sites was PCR amplified from cell lysates harvested 15 days after crRNA transduction. The maximum percentages of reads containing indels overlapping any base position within each of the demarcated regions (region A, region B, region C) are shown. **E)** Comparison of the indicated fusion protein constructs in dual CD55 and CD81 CRISPRi knockdown 10 days after lentiviral transduction of a 6-plex crRNA construct by flow cytometry. Shown are log10 fluorescence intensity for each antibody stain and the percentages of cells in each quadrant, defined by the 5th percentile of non-targeting crRNA for each fluorescence signal, are indicated. **F)** Summary of the same experiment in E for a larger panel of crRNA constructs, showing the percentage of cells with successful double-knockdown of CD55 and CD81 (e.g. same gating strategy as bottom left quadrant in E). 2-6 biological replicates are shown as individual data points and summarized by the mean and SEM as error bars. Two-sample chi-square test was used compare the proportion of cells with double-knockdown between multiAsCas12a-KRAB and each of the other fusion protein constructs; asterisk indicates $p<0.01$ for all replicates. **G)** Analogous to F, except triple knockdown of CD55, KIT, and CD81 was quantified by the percentage of cells that are below the 5th percentile along all 3 dimensions on day 33 after transduction of crRNA constructs. 2 biological replicates are shown as individual data points. Two-sample chi-square test was used compare the proportion of cells with double-knockdown between multiAsCas12a-KRAB and each of the other fusion protein constructs; asterisk indicates $p<0.01$ for all replicates. **H)** Gene expression knockdown by multiAsCas12a-KRAB using 6-plex, 8-plex and 10-plex crRNA array constructs was measured by flow cytometry 10-11 days after lentiviral transduction of crRNA constructs. Shown are median gene expression knockdown averaged from 2-4 biological replicates, with error bars denoting SEM. CRISPRi activities of crFOLH1-1, crCD151-3 and crHBG-3 were not assayed in this experiment, but these spacers are active when encoded as individual crRNAs as shown in Fig. S4B, Fig. 2E, and Fig. 5A, respectively. All protein constructs shown in A-H were delivered by piggyBac transposition into K562 cells and sorted for the same expression level of the P2A-BFP marker (except opAsCas12a was delivered by lentiviral transduction and selected for by puromycin-resistance marker).

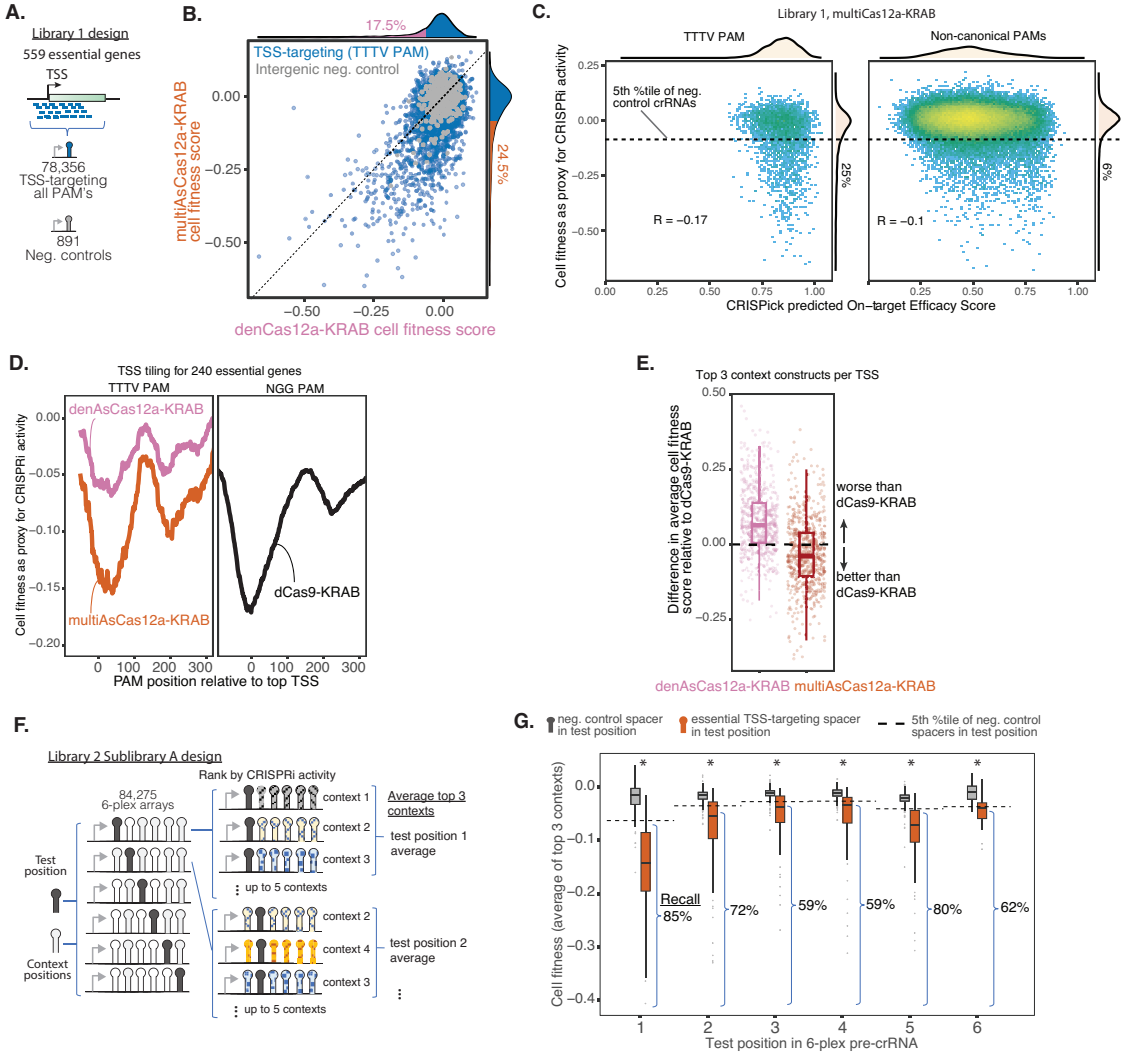


Figure 4 – multiAsCas12a-KRAB enables TSS-targeting pooled CRISPRi screens, including with 6-plex crRNA arrays.

A) Design of Library 1 consisting of single crRNAs tiling TSS-proximal regions of essential genes.

B) Library 1: Scatter plot of cell fitness scores in K562 cells for multiAsCas12a-KRAB vs. denCas12a-KRAB for 3,334 single crRNA constructs with sufficient read coverage for analysis and targeting canonical TTTV PAMs within -50bp to +300bp window of 584 essential gene TSS's. Marginal histograms show percentage of crRNA constructs with cell fitness scores exceeding the 5th percentile of negative control crRNAs. y = x line is shown.

C) 2D density plots of cell fitness scores vs. predicted crRNA on-targeting efficacy score from the CRISPick algorithm, grouped by TTTV PAM vs. non-canonical PAM's. The 5th percentiles of intergenic negative control crRNAs cell fitness scores are shown as a dashed horizontal line and the percentage of crRNAs below that threshold shown in the marginal histogram. Pearson correlation coefficients are shown.

D) Library 1: Moving average cell fitness score across all TTTV PAM-targeting crRNAs at each PAM position relative to the TSS (left), shown for the 240 essential gene TSS's for which analogous dCas9-KRAB NGG PAM tiling screen data (Nuñez et al., 2021) is available in K562 cells (right).

E) Library 1: Boxplots of average cell fitness scores of top 3 crRNAs for each essential TSS for multiAsCas12a-KRAB or denCas12a-KRAB, subtracted by the average cell fitness scores from top 3 sgRNAs for the same TSS for dCas9-KRAB (Nuñez et al., 2021). Boxplots show median, interquartile range, whiskers indicating 1.5x interquartile range, and are overlaid with individual data points.

F) Design of Library 2 Sublibrary A, aimed at evaluating CRISPRi activity each position in the 6-plex array. For each 6-plex array, a specific position is defined as the test position (which can encode either a TSS-targeting spacer or a negative control spacer), and the remaining positions are referred to as context positions encoding one of 5 sets of negative control spacers designated only for context positions.

G) Library 2 Sublibrary A: Analysis of 2,987 6-plex constructs with sufficient read coverage and encodes in the test position one of 123 spacers that scored as strong hits as single crRNAs in the Library 1 screen, compared to constructs that encode an intergenic negative control spacer in the test position. Boxplots show cell fitness scores averaged from the top 3 context constructs of each test position spacer in the 6-plex array, grouped by negative control spacers vs. essential TSS-targeting spacers in each test position. Recall is calculated as the percentage of essential TSS-targeting spacers (that were empirically active in the single crRNA Library 1 screen) recovered by the Library 2 6-plex crRNA array screen for a given test position, using the 5th percentile of constructs containing negative control spacer in the same test position as a threshold for calling hits. Boxplots display median, interquartile range, whiskers indicating 1.5x interquartile range, and outliers. One-sided Wilcoxon rank-sum test was performed for the difference in the distributions of negative control spacers vs. TSS-targeting spacers at each position, with asterisks indicating $p<0.0001$.

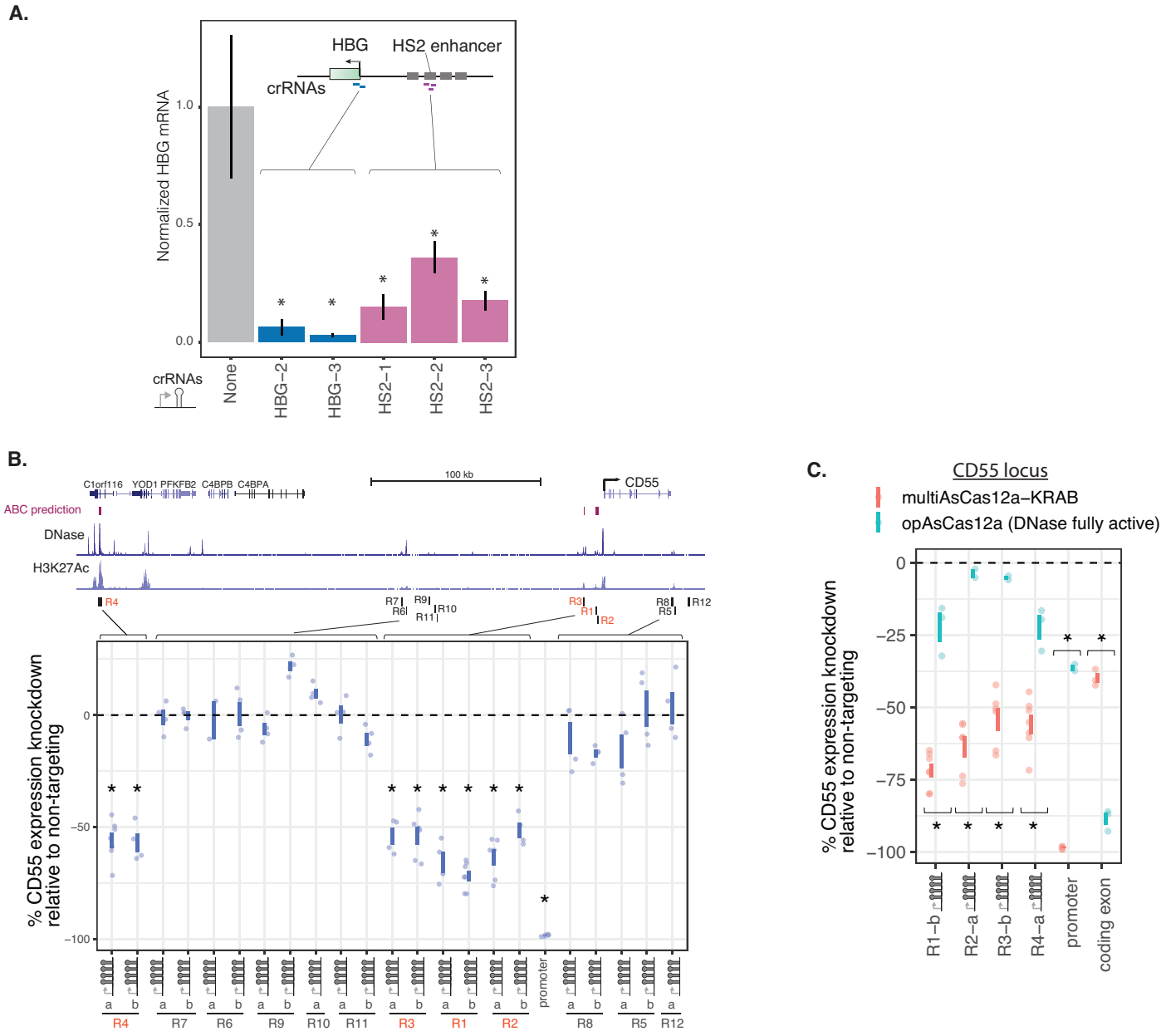


Figure 5 – multiAsCas12a-KRAB CRISPRi enables enhancer perturbation and discovery.

A) K562 cells constitutively expressing multiAsCas12a-KRAB are lentivirally transduced with single crRNAs targeting the HBG TSS or its known enhancer, HS2. Shown are HBG mRNA levels measured by RT-qPCR, normalized to GAPDH levels and averaged across 6-7 technical replicates from 2 independent experiments. Error bars denote SEM. One-sided unpaired Student's t-test was performed to compare untransduced control vs. each individual crRNA; asterisk denotes $p<0.05$.

B) Genome browser view of the CD55 locus, including predicted enhancers using the activity-by-contact model and DNase-seq and H3K27Ac ChIP-seq tracks from ENCODE. K562 cells piggyBac-engineered to constitutively express multiAsCas12a-KRAB was transduced with 4-plex crRNA constructs targeting regions (R1-R11) in the CD55 locus, and R12 as a negative control region devoid of enhancer features. Each unique 4-plex crRNA construct is labeled as "a" or "b". For comparison, targeting the CD55 promoter using a 6-plex crRNA array (crCD55-4_crB2M-1_crKIT-2_crKIT-3_crCD81-1) is included. CD55 expression was assayed by flow cytometry between 9 and 11 days after crRNA transduction. One-sided Wilcoxon rank-sum test was performed on the median expression knockdown across 4-7 biological replicates for each crRNA construct, compared to the median expression knockdown of R12 (negative control region); asterisk indicates $p < 0.01$.

C) Comparison of CRISPRi targeting in K562 cells engineered to constitutively express multiAsCas12a-KRAB vs. opAsCas12a using a subset of lentivirally transduced crRNA constructs from B, plus a crRNA construct targeting a coding exon of CD55 as a positive control for knockdown by DNA cutting. CD55 expression was assayed by flow cytometry 11 days after crRNA transduction. One-sided Wilcoxon rank-sum test was performed to compare the median expression knockdown of multiAsCas12a-KRAB vs. opAsCas12a across 2-7 biological replicates; asterisk indicates $p \leq 0.05$.

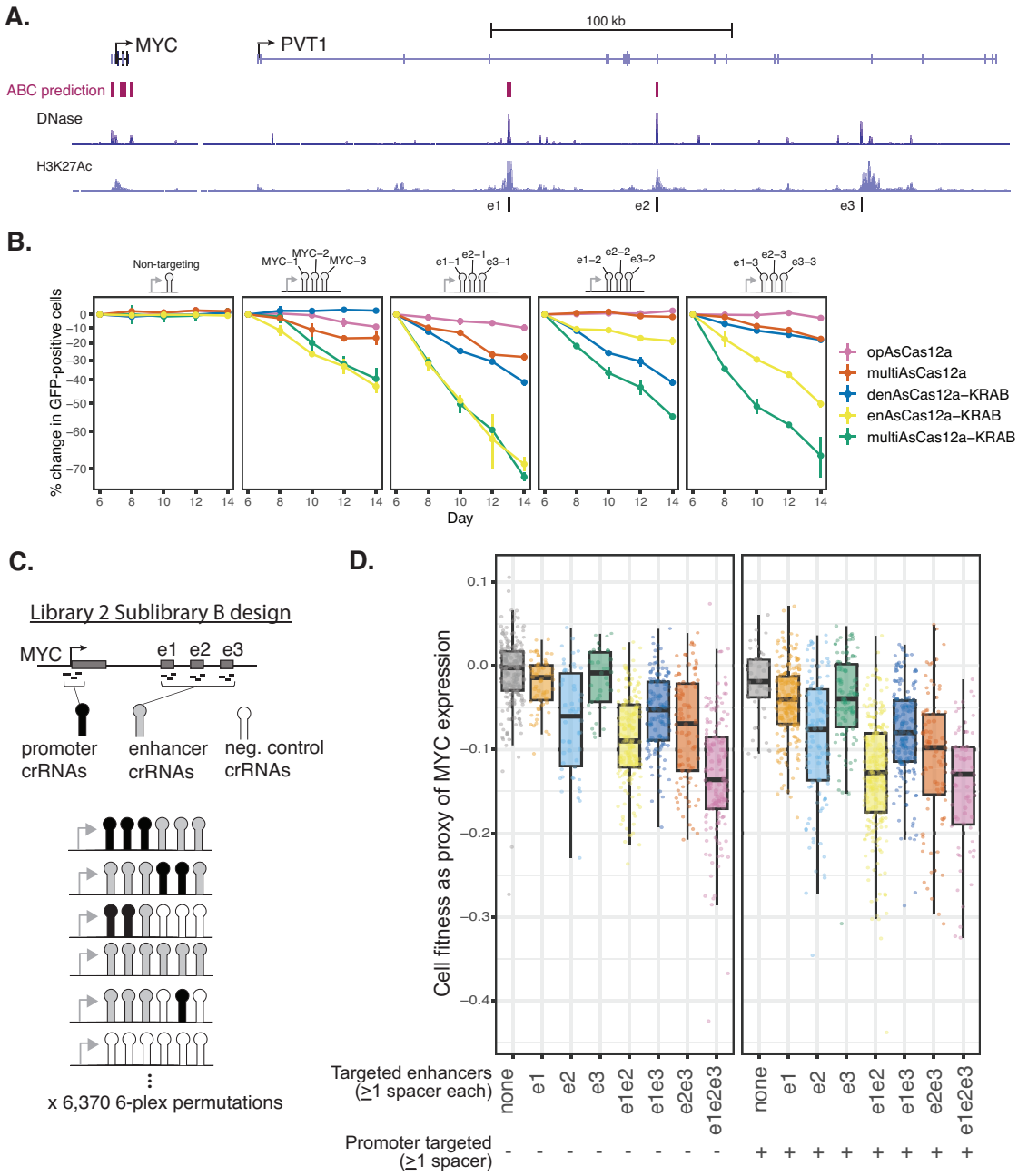


Figure 6 – multiAsCas12a-KRAB enables combinatorial targeting of cis-regulatory elements in pooled CRISPRi screens.
A) Genome browser view of the MYC locus, including activity-by-contact model predictions, and DNase-seq and H3K27Ac ChIP-seq tracks from ENCODE. 3 of the known MYC enhancers (e1, e2, e3) in the body of the non-coding RNA, PVT1, are shown.
B) K562 cells piggyBac-engineered to constitutively express the indicated panel of fusion protein constructs were transduced with one of 4 3-plex crRNA constructs targeting the MYC promoter or co-targeting the 3 enhancers using one crRNA per enhancer. Cell fitness as a proxy of MYC expression is measured as log2 fold-change in percentage of cells expressing GFP marker on the crRNA construct, relative to day 6 after crRNA transduction. Shown are the average of 2 biological replicates, with error bars denoting the range.
C) 6,370 6-plex permutations of the 12 individual spacers from B, together with 3 intergenic negative control spacers, were designed and cloned as 6-plex crRNA arrays used in the design of Library 2 Sublibrary B.
D) Library 2 Sublibrary B: Analysis of 1,823 constructs with sufficient read coverage, categorized based on whether each contains at least one of 3 crRNAs that target the MYC promoter, and/or at least one crRNA that targets each of the MYC enhancers. Boxplots summarize cell fitness score distributions (as proxy of MYC expression) of all constructs that fall in each category. Boxplots show median, interquartile range, whiskers indicating 1.5x interquartile range, and are overlaid with individual data points each representing a 6-plex construct.

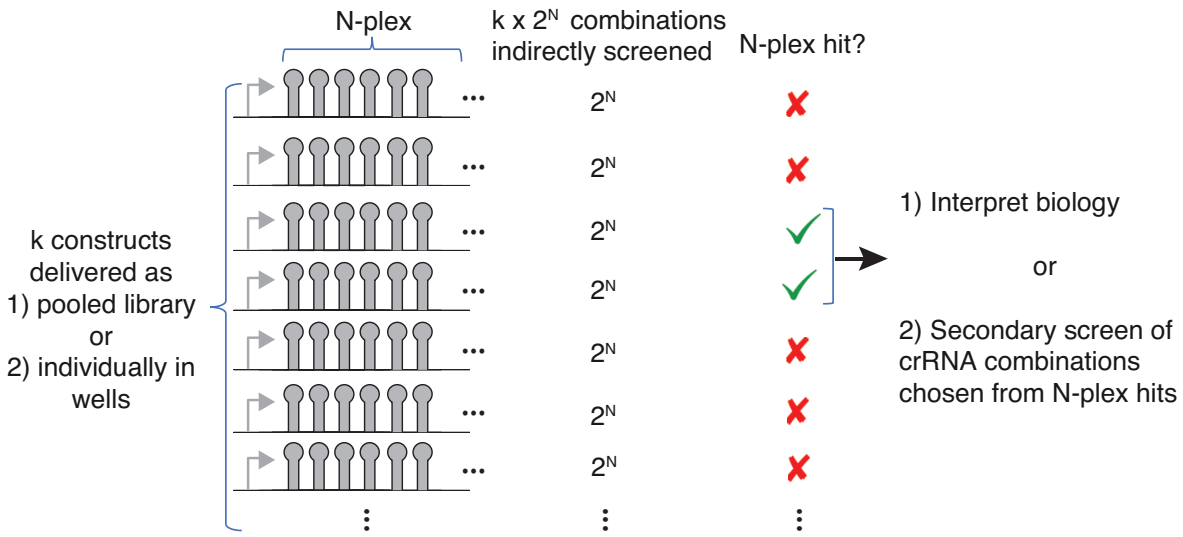


Figure 7 – Group testing framework for efficient exploration of combinatorial CRISPR perturbations.

Supplemental Figures

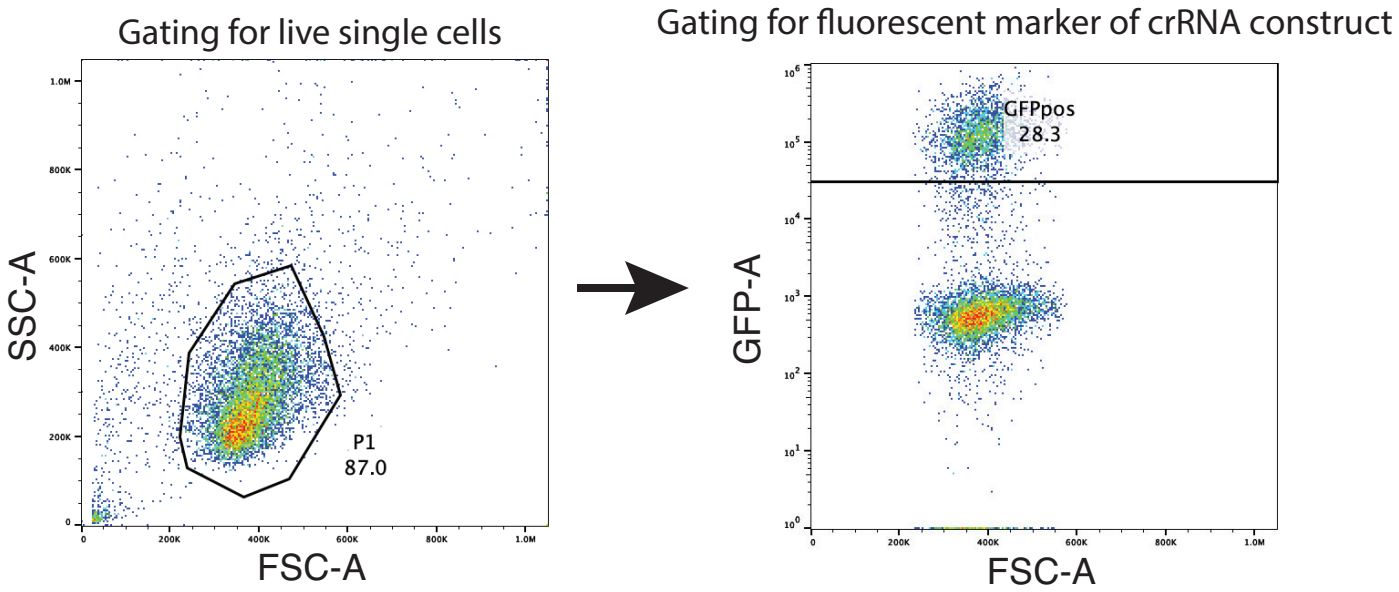


Figure S1 – Example of flow cytometry gating strategy for CRISPRi experiments. Example of the general gating strategy for CRISPRi experiments using flow cytometry readouts, shown for K562 cells as an example. Single live cells are gated by FSC vs. SSC, followed by gating for the fluorescent marker on the crRNA construct (typically GFP), which are subsequently analyzed for target gene expression in the respective fluorescence channels.

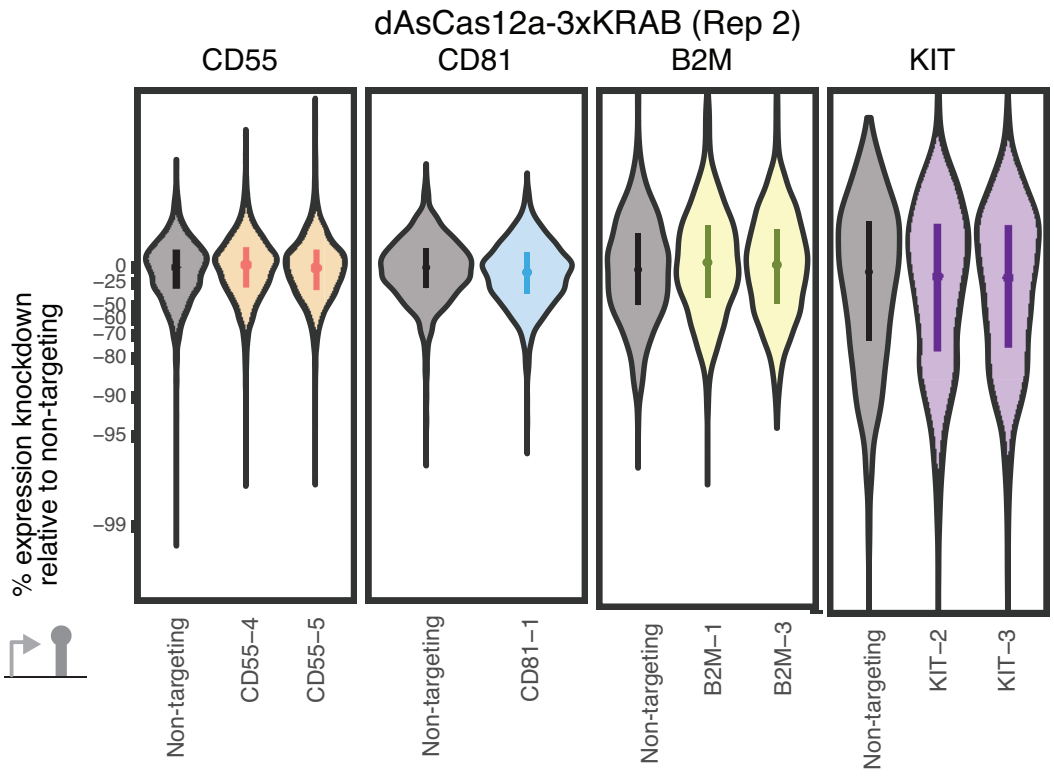
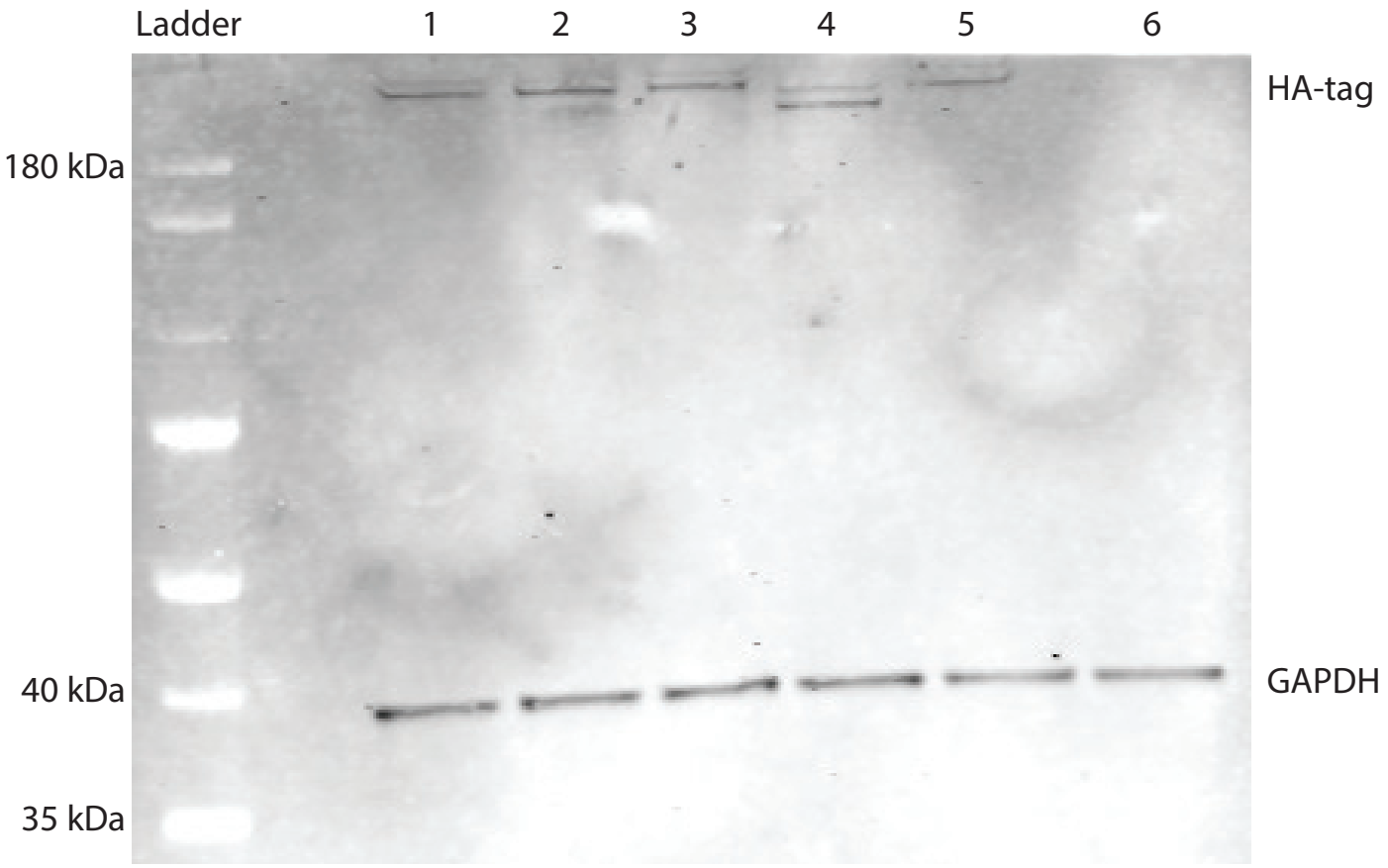


Figure S2 – Additional replicate testing dAsCas12a-3xKRAB CRISPRi activity.
Additional replicate for Fig. 1B testing CRISPRi activity of K562 cells piggyBac-engineered to constitutively express dAsCas12a-3xKRAB and using lentivirally delivered crRNA constructs. See Fig. 1B for details.



Well	Protein construct	Plasmid ID
1	multiAsCas12a-HA-XTEN80-KRAB-P2A-BFP	pCH61
2	denAsCas12a-HA-XTEN80-KRAB-P2A-BFP	pCH62
3	enAsCas12a-HA-XTEN80-KRAB-P2A-BFP	pNS49
4	multiAsCas12a-HA-P2A-BFP	pNS35
5	dAsCas12a-3xKRAB-3xHA-P2A-BFP	pCH68
6	No construct	No construct

Figure S3 – Western blot of fusion proteins. Western blot of whole-cell lysates prepared from K562 cells piggyBac engineered to constitutively express each of the fusion proteins in the panel. anti-HA tag was used for detection of the fusion protein and anti-GAPDH for detection of GAPDH as loading control.

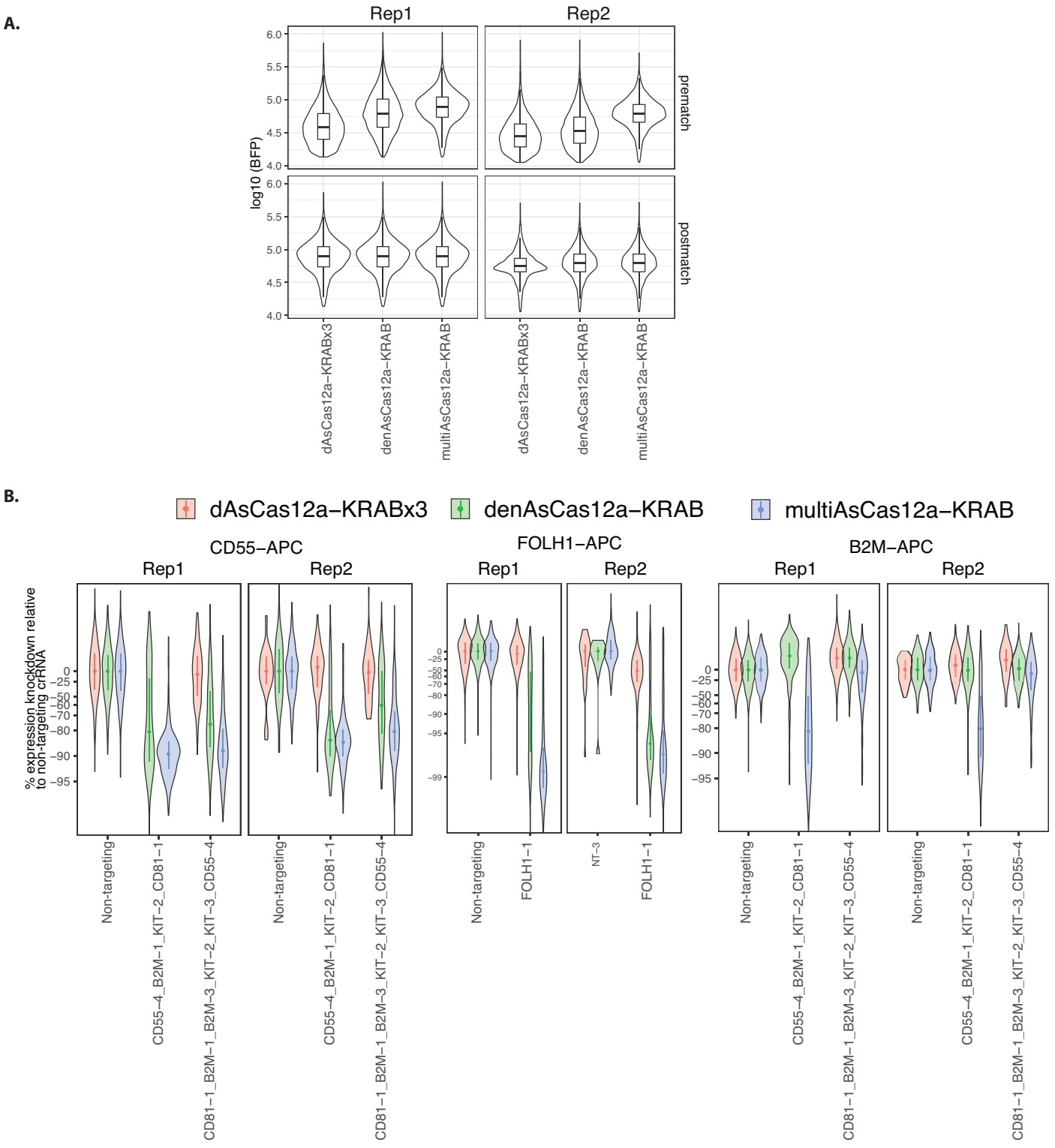


Figure S4 – CRISPRi activity of multiCas12a-KRAB, denAsCas12a-KRAB, and dAsCas12a-KRABx3 in C42B cells.

A) C4-2B cells piggyBac-engineered to constitutively express each of the fusion protein constructs are lentivirally transduced with the indicated crRNA constructs. Cells were sorted based on P2A-BFP marker signal. Because some of these cell lines showed slightly different levels of BFP signal as a proxy of fusion protein expression, to account for fusion protein expression we performed propensity score matching to subset for populations of cells for each fusion protein construct with the same distributions in BFP signals after data acquisition for flow cytometry in CRISPRi experiments. The BFP signals before and after matching are shown as violin blots overlaid with boxplots showing median and interquartile range.

B) Target gene expression are measured by cells surface antibody staining and flow cytometry 13-14 days after crRNA transduction. Single-cell distributions of expression knockdown relative to non-targeting crRNA are shown with mean and interquartile range indicated using the cells after propensity score matching for BFP levels as described in A. CRISPRi knockdown results are indistinguishable with and without propensity score matching (not shown).

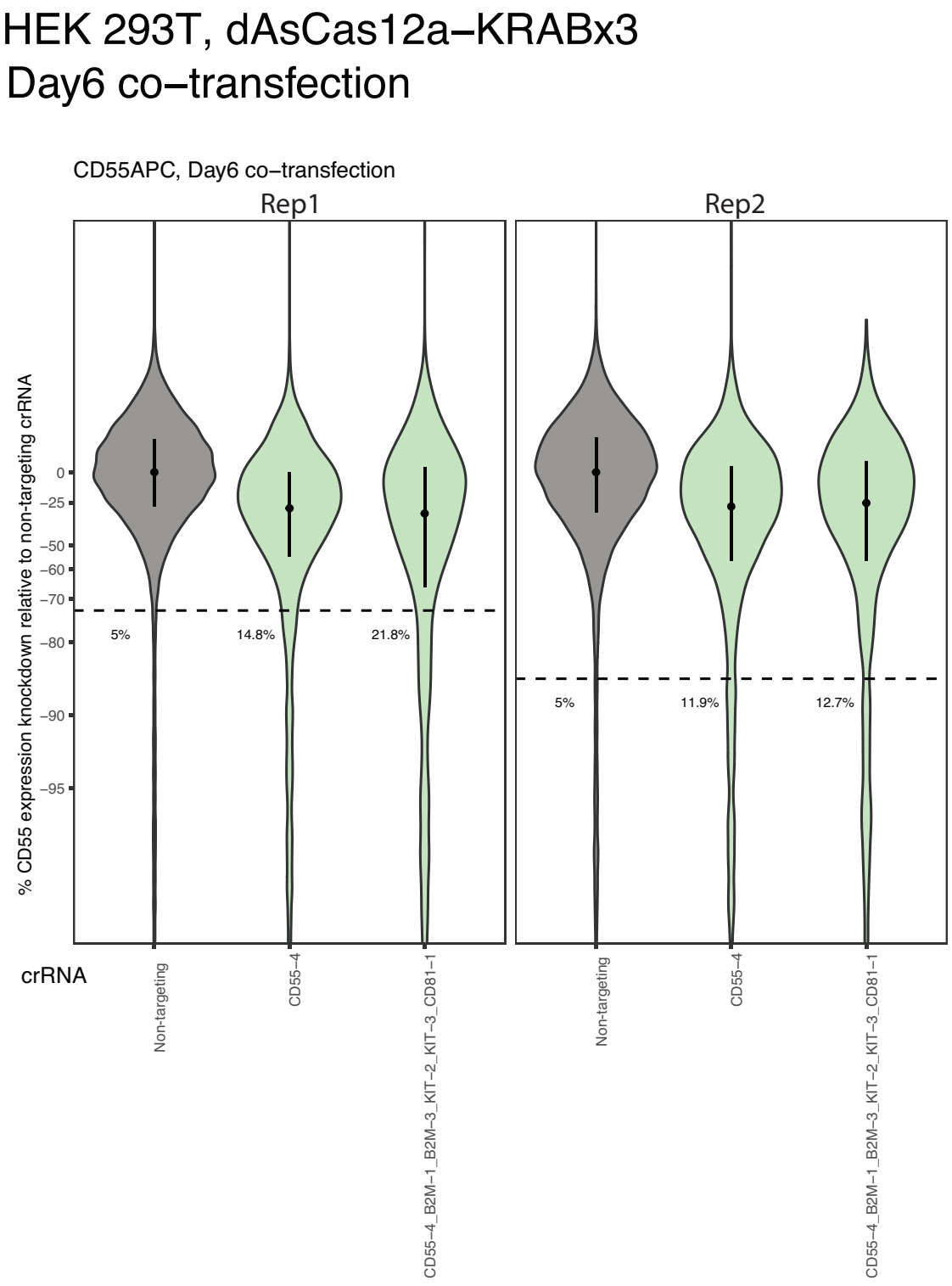


Figure S5 – dAsCas12a-KRABx3 CRISPRi by transient transfection in HEK 293T cells. HEK 293T cells were co-transfected with a plasmid encoding for dAsCas12a-KRABx3 and plasmids encoding for the indicated crRNA constructs targeting CD55. Cells were sorted 2 days after transfection for successful co-transfection based on BFP and GFP markers on the plasmids and CD55 expression was measured by antibody staining on flow cytometry 6 days after transfection. Violin plots of single-cell distributions of CD55 expression knockdown as a percentage of the median of non-targeting control are shown. Median and interquartile range are shown in the plot. The percentage of cells below the 5th percentile of the non-targeting control are also shown.

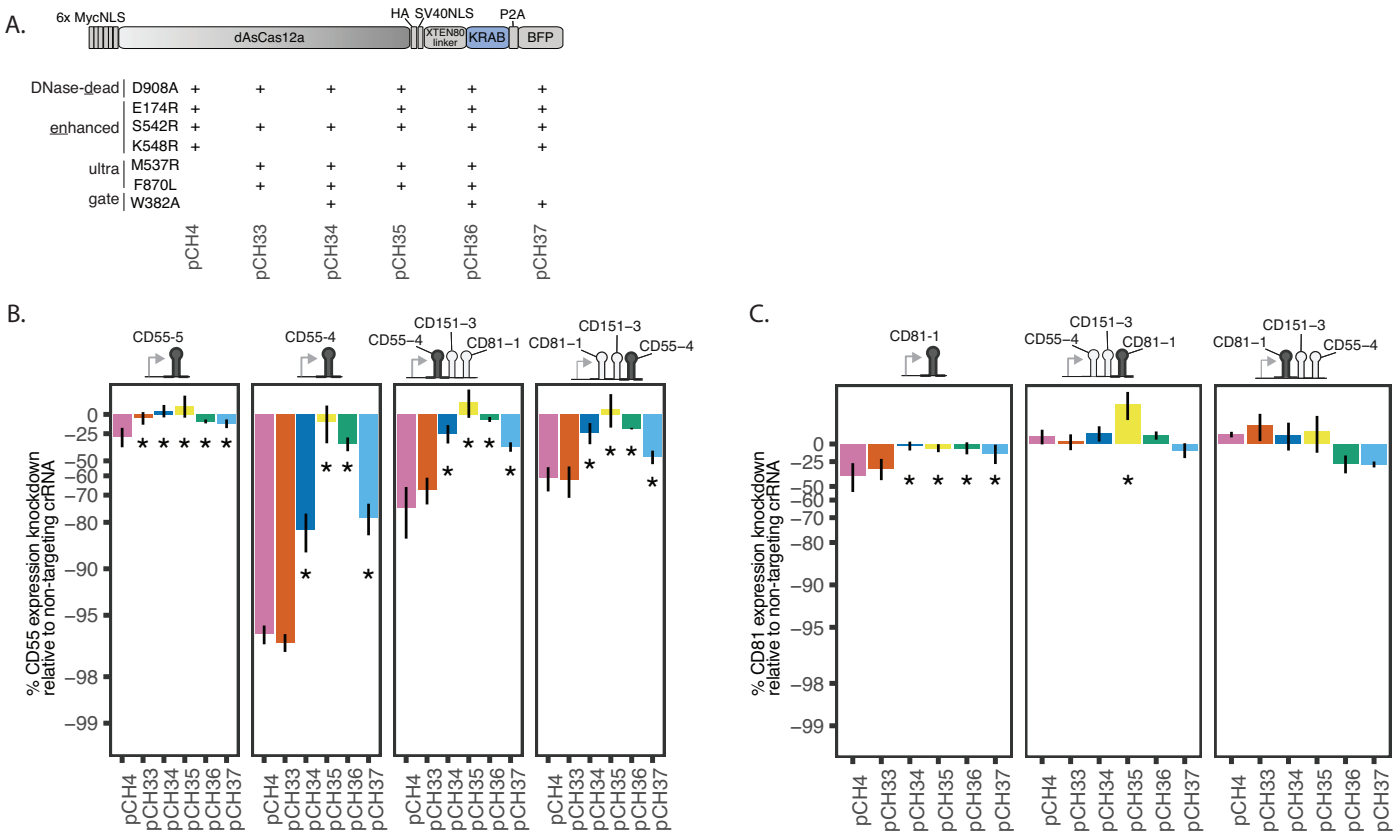


Figure S6 – Comparisons of dAsCas12a variant fusion CRISPRi constructs using up to 3-plex crRNA constructs.
A) The same fusion protein schematic as shown in Fig. 1C, labeled with construct IDs for ease of reference.
B) CD55 expression knockdown measured by flow cytometry using the indicated crRNA constructs and the panel of fusion protein constructs in A. Shown are averages of the median single-cell expression knockdown relative to non-targeting crRNA for 3 biological replicates (including the replicate for crCD55-4 shown in Fig. 1C) for all comparisons, except the comparison for crCD81-1_crCD151-3_crCD55-4 contains 2 replicates. One-sided Wilcoxon rank-sum test comparing denAsCas12a-KRAB (pCH4) to each of the other fusion constructs in the panel was performed. Asterisk indicates $p<0.01$ for all replicate-level comparisons for a given construct comparison.
C) Analogous to B, but for CD81 knockdown. Summaries shown for 3 biological replicates (including the replicate for crCD81-1 shown in Fig. 1C) for all comparisons, except the comparison for crCD81-1_crCD151-3_crCD55-4 contains 2 replicates.

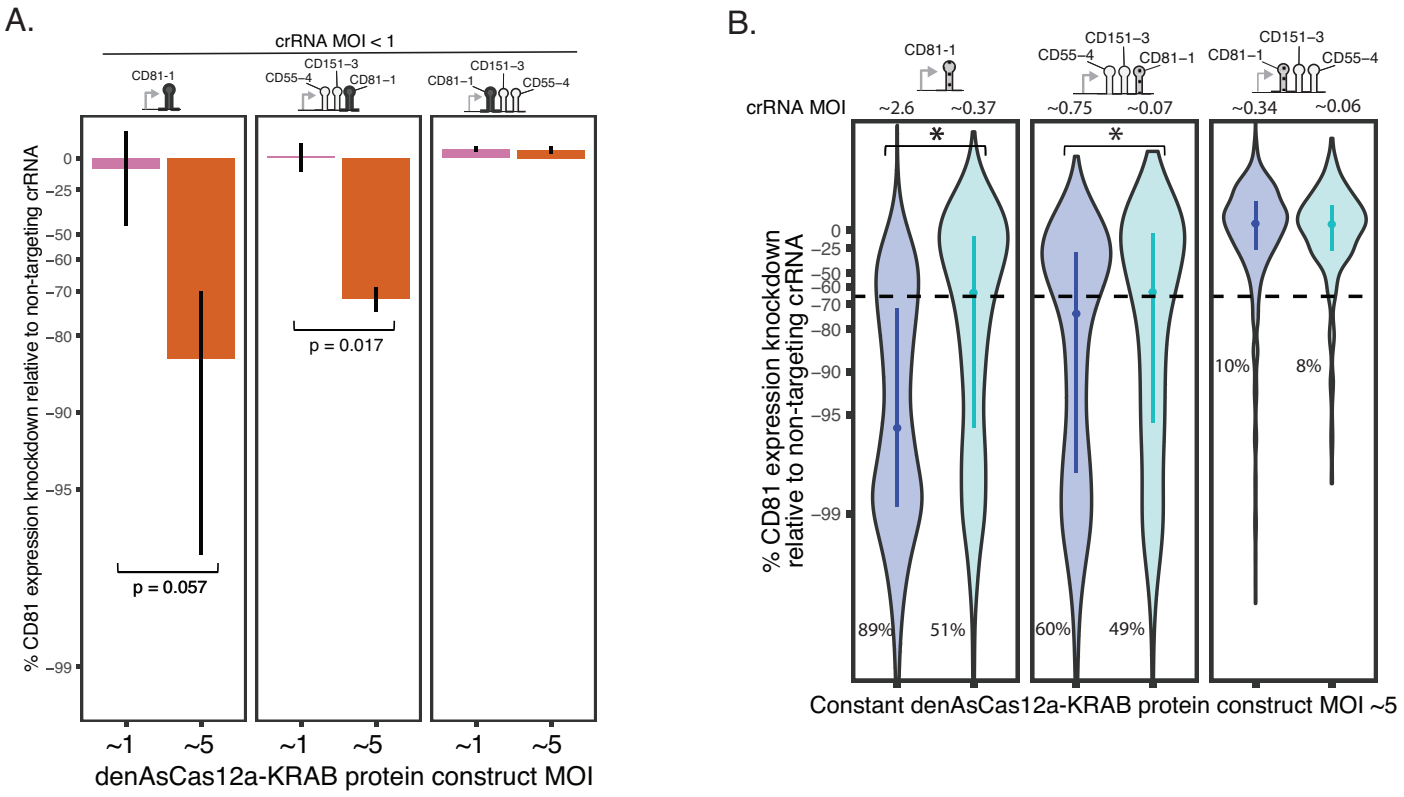


Figure S7 – Additional replicates testing effect of dose on denAsCas12a-KRAB CRISPRi activity

A) Summary of all replicates for experiment shown in Fig. 1D: shown are averages of median expression knockdown for each crRNA construct (N = 3-6 biological replicates for each crRNA construct, including the replicate shown in Fig. 1D). Error bars denote SEM. One-sided Wilcoxon rank-sum test was performed on the medians of single-cell expression knockdown of each replicate and p-values indicated where relevant.

B) Additional biological replicate for Fig. 1E testing denAsCas12a-KRAB CRISPRi activity with varying crRNA MOI while holding protein MOI constant; see Fig. 1E for details.

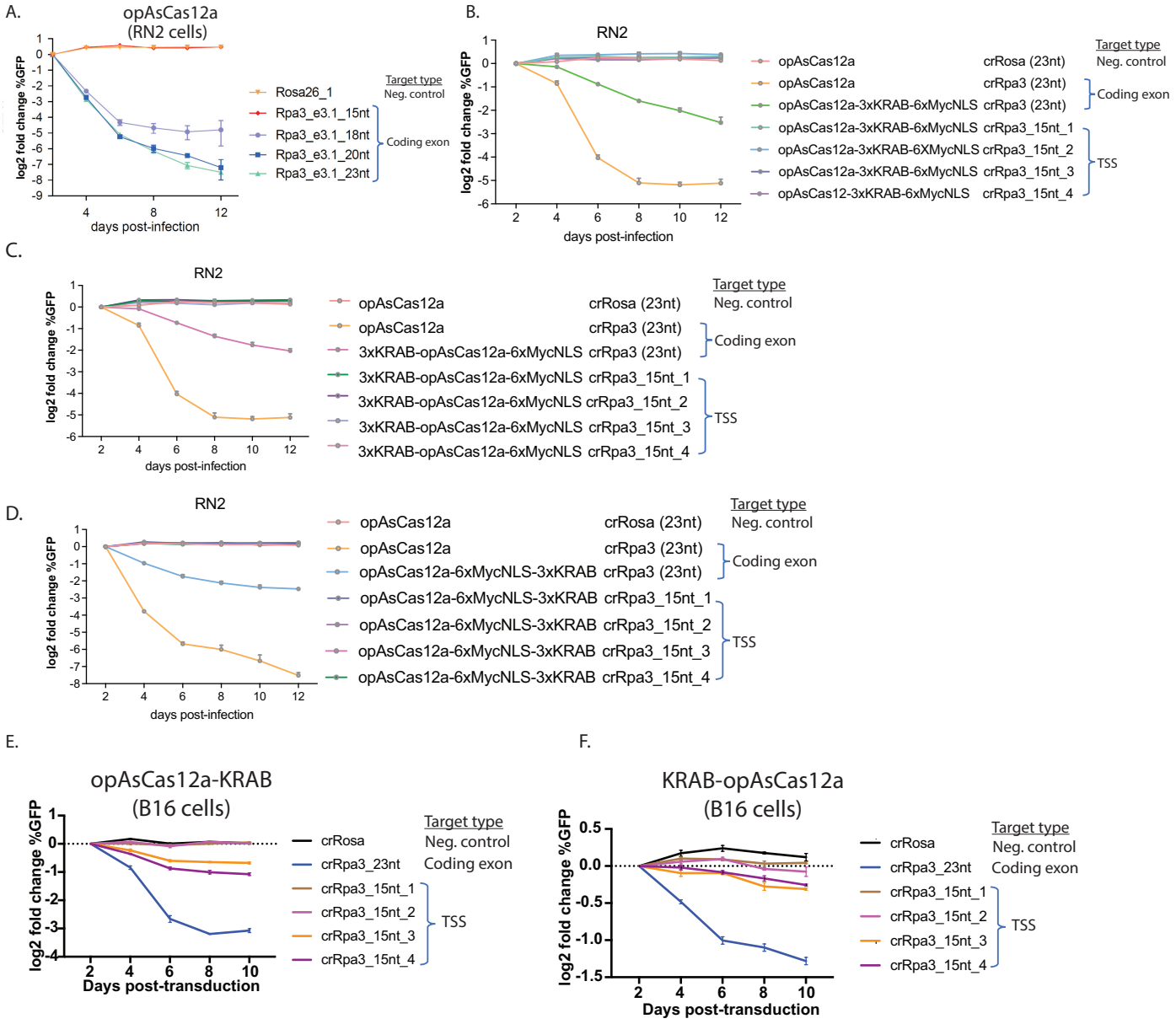


Figure S8 – Testing CRISPRi activity of lentivirally delivered truncated crRNAs. In all panels, the indicated cell line (RN2 or B16) was engineered for constitutive expression of the indicated fusion protein constructs by lentiviral transduction, followed by lentiviral transduction (at MOI between 0.3-0.4) of the indicated single-plex crRNA constructs containing spacers of the indicated lengths targeting Rpa3, an essential gene. The spacers target either the gene's coding exon, the TSS region, or the Rosa locus (negative control) as indicated in the legends. Cell fitness phenotype over time is measured in a competition assay by quantifying log2 fold change in percent of cells expressing the GFP marker on the crRNA expression constructs (relative to day 2). Error bars indicate SEM for N = 3 biological replicates for all panels.

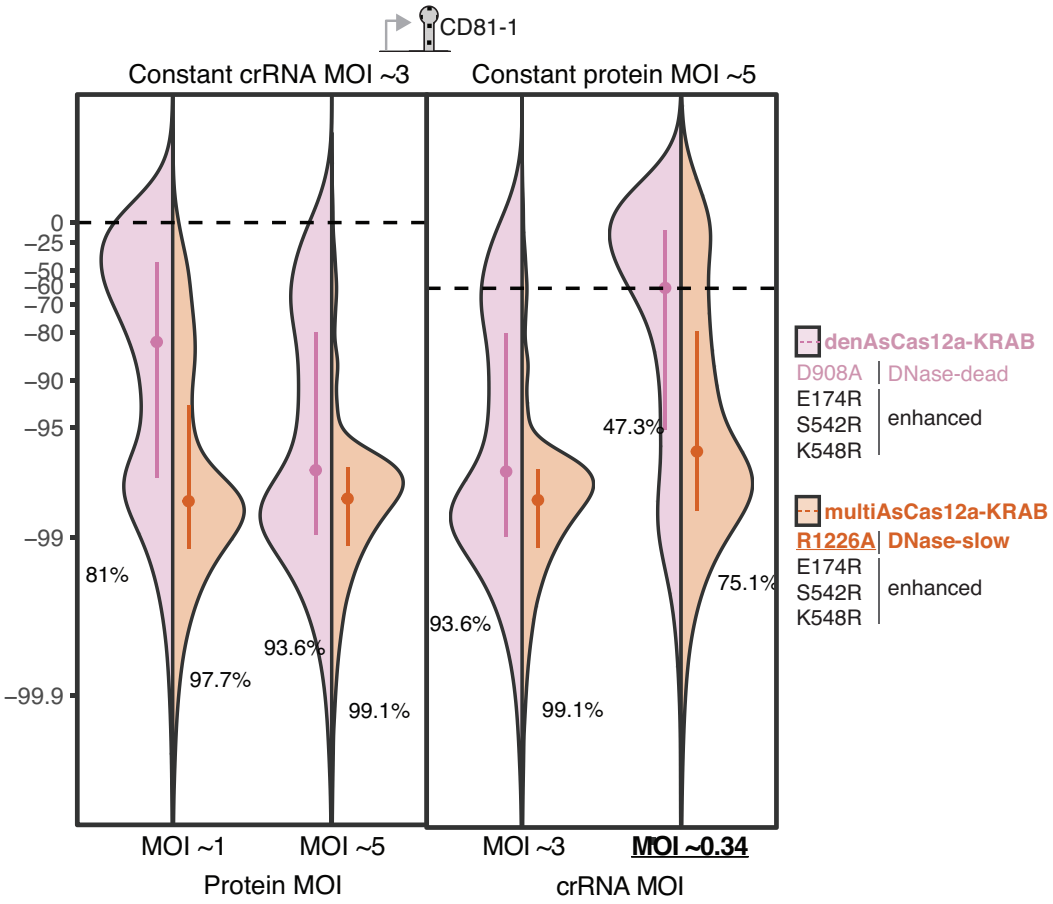


Figure S9 – CD81 knockdown denAsCas12a-KRAB vs. multiAsCas12a-KRAB at different protein and crRNA MOIs. Second biological replicate for testing denAsCas12a-KRAB vs. multiAsCas12a-KRAB CRISPRi activity in the setting of varying protein MOI or crRNA MOI, same as Fig. 2B. See Fig. 2B for further details.

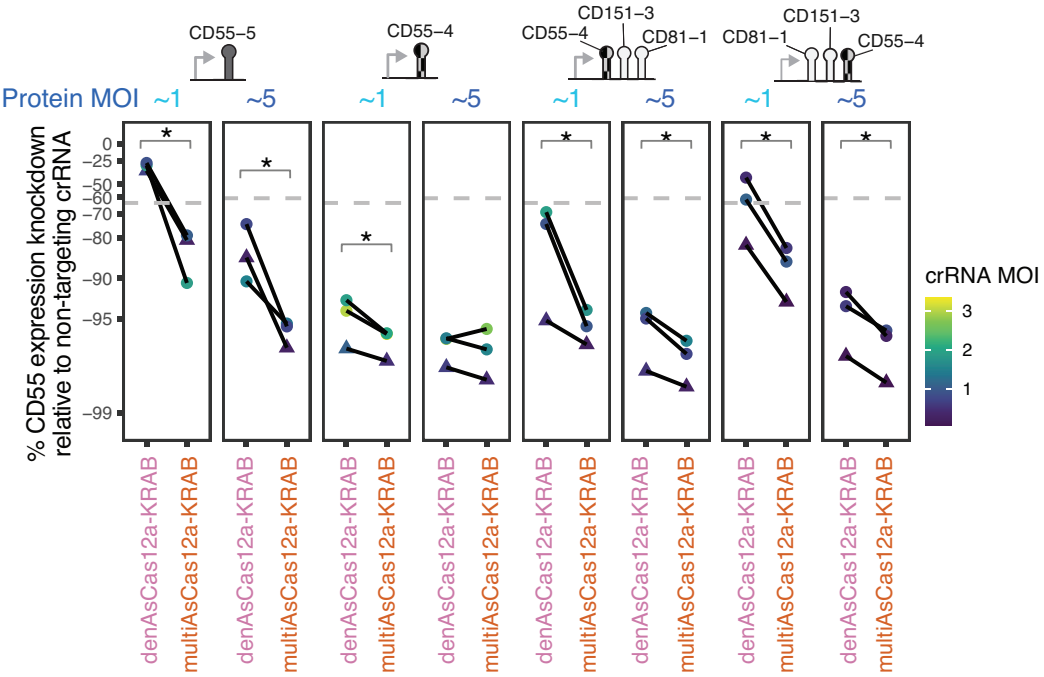
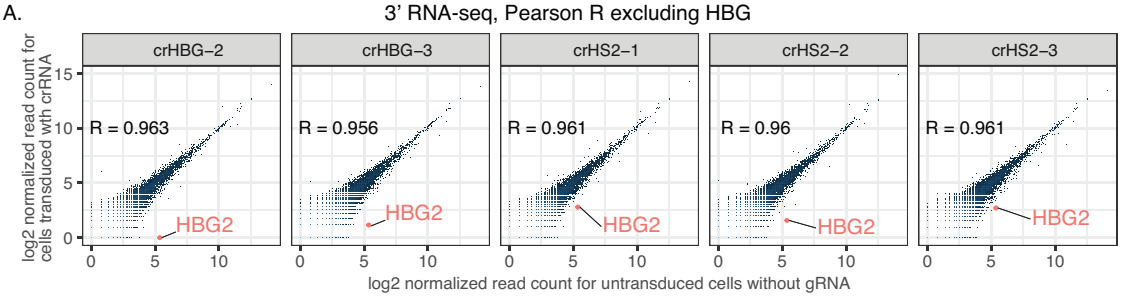
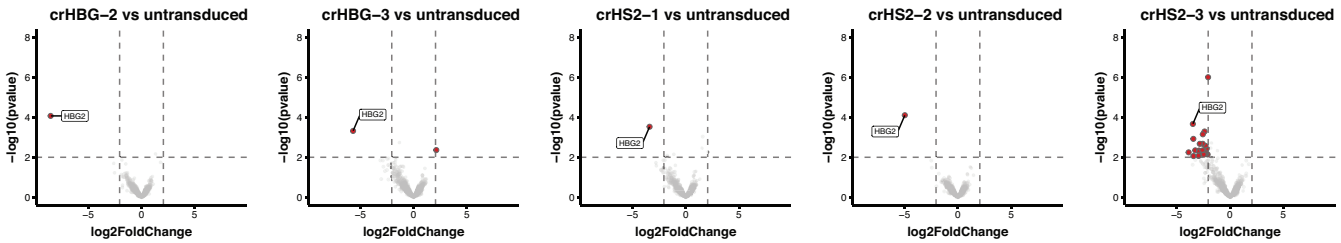


Figure S10 – CD55 knockdown by denAsCas12a-KRAB vs. multiAsCas12a-KRAB at different protein MOIs. Comparison of CD55 knockdown by lentivirally delivered denAsCas12a-KRAB vs. multiAsCas12a-KRAB at protein MOI ∼1 vs. ∼5 across a panel of single and 3-plex crRNA constructs, while holding constant crRNA MOI for each paired fusion protein comparison for each crRNA construct. Dashed gray line indicates 5th percentile of non-targeting crRNA control. crRNA MOI indicated by color scale. Lines connect paired replicates. One-sided Wilcoxon rank-sum tests were performed on single-cell distributions for each replicate, and asterisk denotes $p<0.01$ for all paired replicates within each condition. Dots indicate flow cytometry measurement 10 days after crRNA transduction; triangles indicate flow cytometry measurement 16 days after crRNA transduction.

A.



B.



C.

crHBG-3 differentially expressed genes

	log2FoldChange	pvalue
COL6A3	-2.111480489	0.032227

crHS2-3 differentially expressed genes

	log2FoldChange	pvalue
COL6A3	-2.111480489	0.032227
ITGA3	-1.754297196	0.039295
PHLDA1	-1.627213967	0.143479
MALAT1	-1.321521059	0.001853
UCHL1	-2.374917208	0.011372
HBG2	-5.687020138	0.00048
TM4SF1	-1.471229856	0.105152
CD44	-1.689959013	0.025554
IGFBP7	-1.39428827	0.077142
CTSA	-1.778803563	0.033146
ALDH1A1	-1.336719338	0.071234
KRT81	-1.53327734	0.072989
THBS1	-1.179123518	0.252654
APP	-1.867798594	0.019031
CCND1	-1.879166671	0.066226
ANXA1	-1.720793771	0.046504
ALDH3A1	-1.879322885	0.03484
TGFBI	-1.58211392	0.030056
NEAT1	-1.670821249	0.013372
HMOX1	-1.762705604	0.067726
PLS3	-2.890643152	0.014401
SERPINE1	-1.635079753	0.067127

All crisprVerse off-target predictions

chr	TSS coordinate	Gene name
chr2	120211053	MTATP6P26
	120211726	MTCO3P43
	120212934	MTND3P10
	120213343	MTND4LP14
	120213630	MTND4P26
chr11	5249857	HBG1
	5254783	AC104389.6
chr12	7838913	AC006517.6
	7954215	AC006511.7
chr17	57531768	AC007431.3

Figure S11 – RNA-seq analysis of crRNA specificity.

A) K562 cells lentivirally engineered (MOI ∼5) to constitutively express multiAsCas12a-KRAB were either transduced with the indicated crRNA's at MOI <0.3, followed by sorting for crRNA-transduced cells based on GFP marker, or received no crRNAs. RNA was isolated from the sorted cells 32 days of culture after crRNA transduction and subjected to 3' RNA-seq. Scatter plot of normalized mRNA expression levels for crRNA transduced (1 biological replicate each) vs. cells without crRNA (2 biological replicates), and Pearson correlation coefficient calculated for the transcriptome, excluding HBG. RT-qPCR quantifications are shown in Fig. 5A.

B) Volcano plots of p-values vs. log2FoldChange from differential expression analysis using DE-seq2 are shown. Genes that fall beyond p-value and log2FoldChange cutoffs (dashed lines) are highlighted.

C) Lists of differentially expressed genes (other than HBG) from the analysis in B for the crHBG-3 and crHS2-3 transduced cells are shown. For comparison, a list of all off-target predictions generated by crisprVerse are shown for all crRNAs in the panel in A and B.

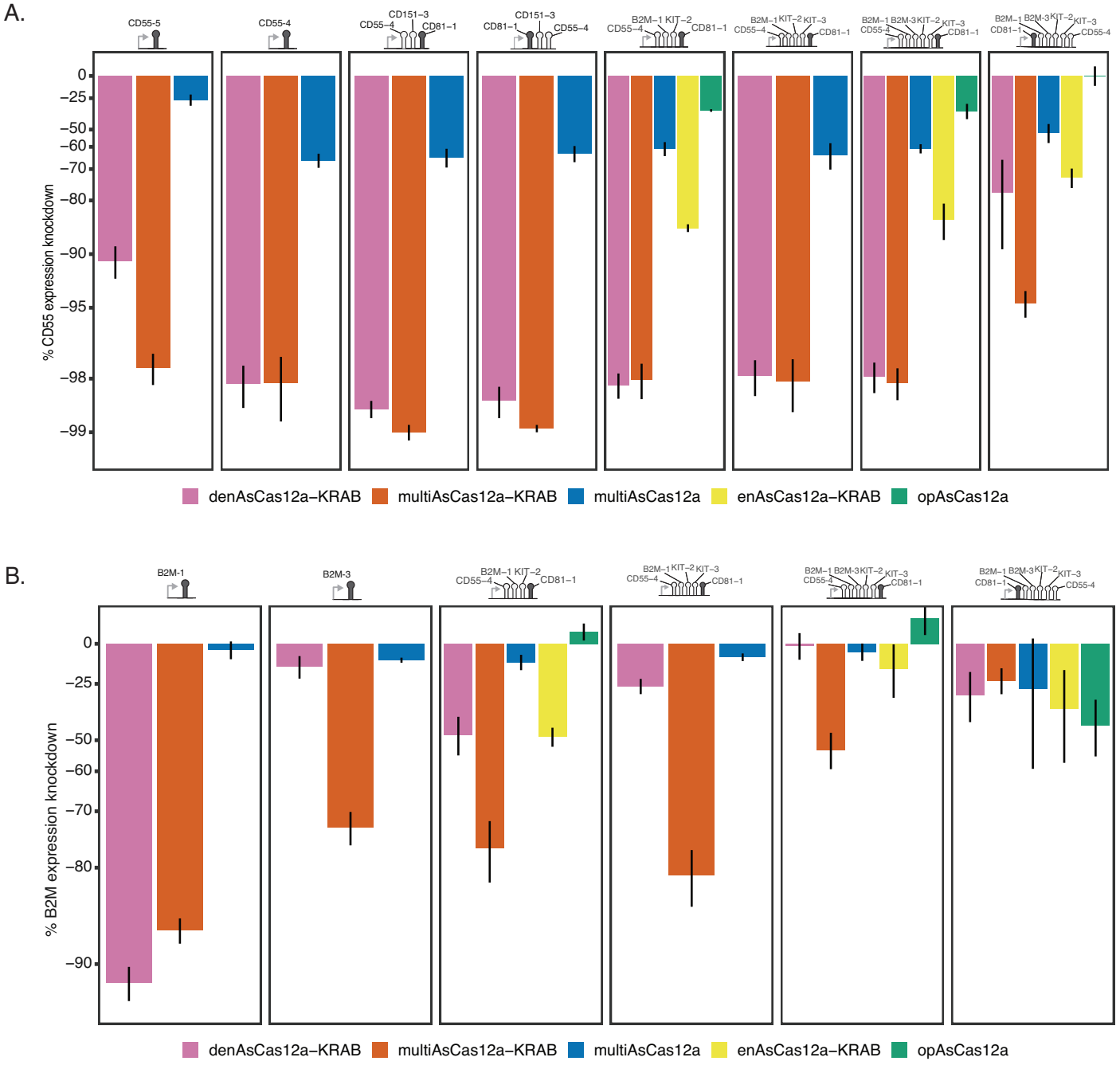


Figure S12 – CRISPRi knockdown of CD55 and B2M using up to 6-plex crRNA arrays. Analogous to Fig. 3B-C, but shown for CD55 and B2M knockdown on day 6 after crRNA transduction, measured by antibody staining of those targets using flow cytometry. Shown are averages of median single-cell expression knockdown from 2-5 biological replicates for each crRNA construct, with error bars indicating SEM.

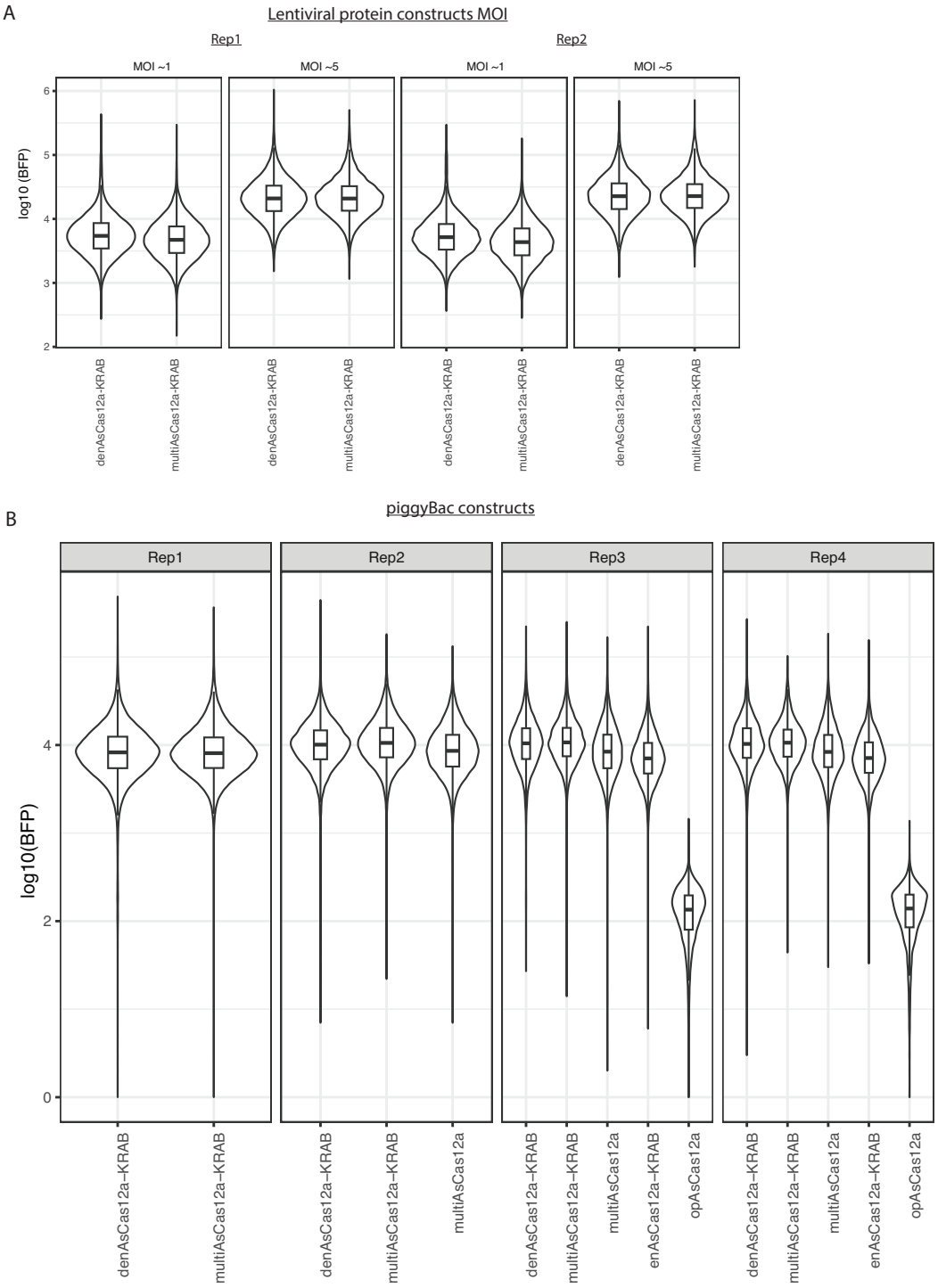


Figure S13 – Monitoring P2A-BFP reporter as proxy of fusion protein expression level.
A) K562 cells lentivirally engineered (at MOI ∼1 or MOI ∼5) to constitutively express the indicated fusion protein constructs were monitored for P2A-BFP expression levels by flow cytometry. Shown are representative biological replicates from routine monitoring.
B) Same as A for K562 cells piggyBac-engineered to constitutively express the indicated fusion protein constructs. opAsCas12a does not contain BFP reporter and is shown as fluorescence negative control. Shown are representative biological replicates from routine monitoring.

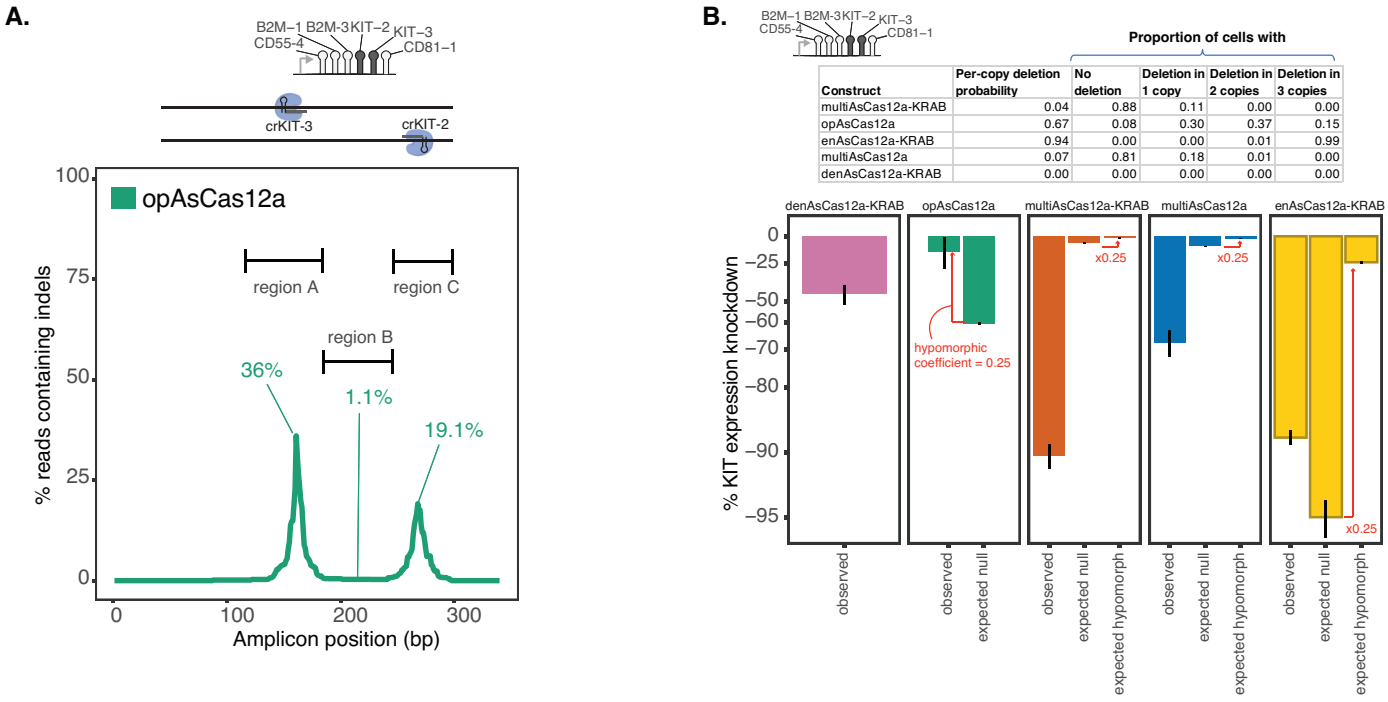


Figure S14 – Indel quantification and gene expression knockdown simulation for dual-targeting of the KIT TSS region.
A) Related to Fig. 3D. Indel quantification of PCR amplicon near the KIT TSS region in K562 cells lentivirally engineered to constitutively express opAsCas12a 15 days after transduction of the indicated 6-plex crRNA array (sorted for crRNA transduced cells on 2 days after transduction). Note that opAsCas12a is encoded in a different expression backbone using a puromycin selectable marker and thus is not directly matched to other fusion constructs in Fig. 3D in transgenic expression level.
B) Based on the observed indel allelic frequencies in A and Fig. 3D, we calculated the expected proportion of cells that harbor a specified number of DNA copies containing indels of any size within the PCR amplicon, assuming indels induced by crKIT-2 and crKIT-3 occur independently across DNA copies within each cell. We purposefully overestimate these indel frequencies as an adversarial test in favor of the impact of indels on gene expression (see Methods). Based on these proportions we simulated the expected distribution of single-cell gene expression levels under the assumption that knockdown were solely due to genetically null deletions of any size abolishing KIT expression in cis ("expected null"). Expected knockdown under this genetic null assumption exceeds that observed for opAsCas12a (fully active DNase), demonstrating the genetic null assumption is an overestimate of gene expression effects of indels in this region. To correct for this overestimate, we use the ratio of observed vs. expected null median expression knockdown by opAsCas12a as an estimate of the hypomorphic effect of deletions in this region ("hypomorphic coefficient"). We multiply the expected null median expression knockdown for all other fusion proteins by this hypomorphic coefficient to obtain an "expected hypormorph" median expression knockdown, which we propose as our final estimate of the effects arising from indels. The observed knockdown values are the same as shown in Fig. 3D.

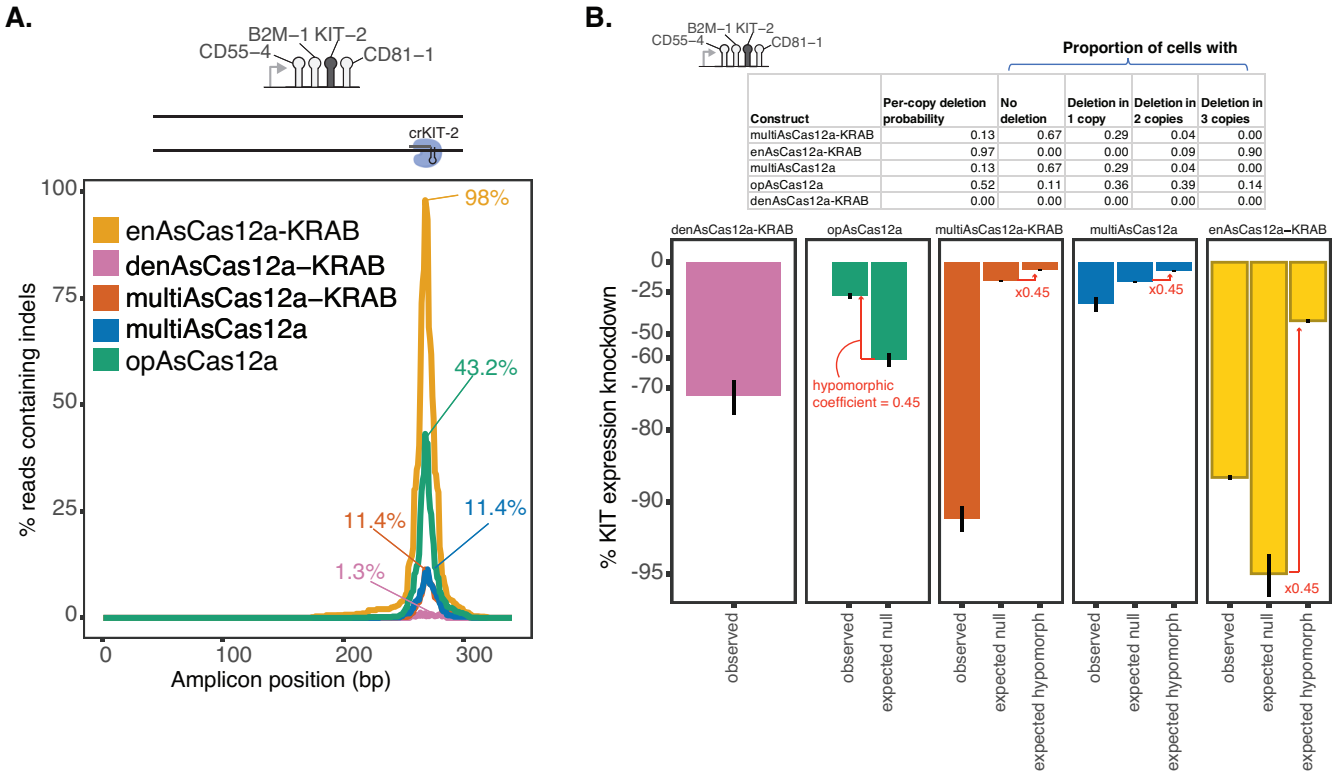


Figure S15 – Indel quantification and gene expression knockdown simulation for single-targeting of the KIT TSS region.
A) Indel quantification analogous to Fig. 3D, but for a single-site targeting of the KIT TSS region using crKIT-2 encoded within the indicated 4-plex crRNA array.
B) Analogous to Fig. S14B, we simulated expected gene expression knockdown (see Methods) under the assumption that knockdown were solely due to genetically null deletions of any size abolishing KIT expression in cis ("expected null"). Expected knockdown under this genetic null assumption exceeds that observed for opAsCas12a (fully active DNase), demonstrating the genetic null assumption is an overestimate of gene expression effects of indels in this region. To correct for this overestimate, we use the ratio of observed vs. expected null median expression knockdown by opAsCas12a as an estimate of the hypomorphic effect of deletions in this region ("hypomorphic coefficient"). We multiply the expected null median expression knockdown for all other fusion proteins by this hypomorphic coefficient to obtain an "expected hypormorph" median expression knockdown, which we propose as our final estimate of the effects arising from indels.

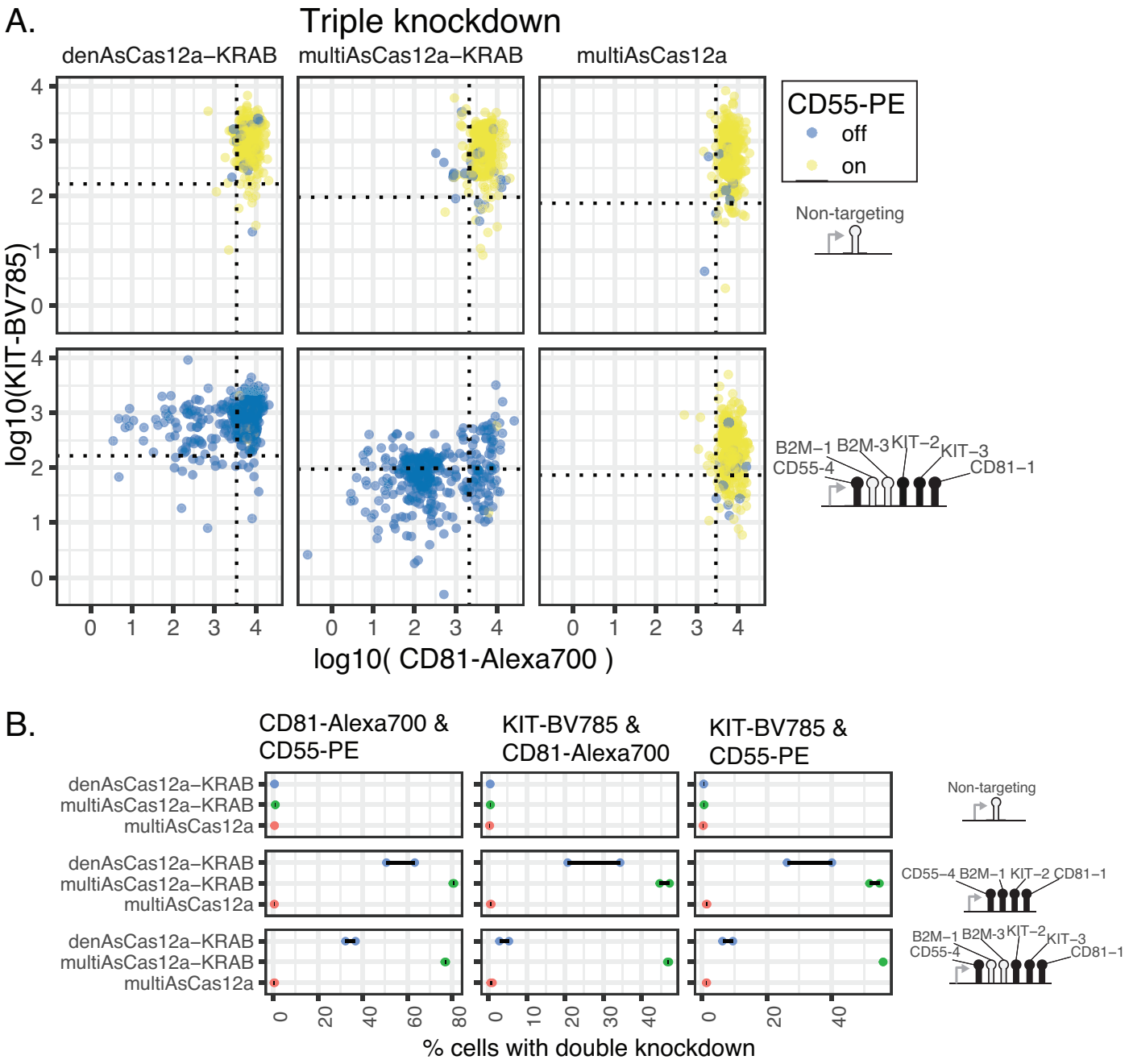


Figure S16 – Double and triple gene knockdown by CRISPRi using higher-order crRNA arrays.
A) Single-cell view of CD81, KIT, and CD55 3-way knockdown using a 6-plex crRNA construct in K562 cells piggyBac-engineered to constitutively express each of the indicated fusion protein constructs, measured by multiplexed flow cytometry. Summary of percentage of cells with triple knockdown is shown in Fig. 3G.
B) Quantification of the fraction of cells showing double-knockdown of pairs of target genes in the experiment described in A for 4-plex and 6-plex crRNA arrays. Double knockdown is defined as the fraction of cells with expression below the 5th percentile for the non-targeting crRNA for the expression of a given pair of target genes.

A.

Library 1 (single-plex)	TSS-targeting constructs	Negctrl-intergenic	Negctrl-unmapped
# constructs	77387	524	445
# genes targeted	559		
For constructs with ≥1 RPM			
# constructs	76293	524	445
# constructs targeting TTTV PAM	3326		
# target genes with >0 crRNA exceeding fitness score of 1%tile of neg controls	450		
# target genes shared with prior dCas9 CRISPRi screens	240		

B.

Library 2 (6-plex)	Sublibrary A (positional shuffling)	Neg-ctrl only arrays
# constructs	84275	2000
For constructs with ≥1 RPM		
# constructs	67534	324
# constructs with TSS-targeting spacer in test position	44802	NA
# constructs with neg control spacer in test position	12029	NA
# constructs with TSS-targeting spacer in test position exceeding fitness score of 1%tile of neg controls in Library 1 screen	2987	NA
# unique TSS-targeting spacers in test position exceeding fitness score of 1%tile of neg controls in Library 1 screen	123	NA
# unique neg control spacers in test position	506	NA

C.

Library 2 (6-plex)	Sublibrary B (MYC locus)		Neg-ctrl only arrays
# constructs	6370		2000
For constructs with ≥1 RPM			
# constructs	1823		324
Promoter-targeting	**Enhancer-targeting (≥1 for each listed)**	**# constructs**	
≥1 promoter targeting spacer	none	48	
none	none	197	
≥1 promoter targeting spacer	e1	187	
none	e1	65	
≥1 promoter targeting spacer	e2	108	
none	e2	50	
≥1 promoter targeting spacer	e3	90	
none	e3	33	
≥1 promoter targeting spacer	e1e2	216	
none	e1e2	164	
≥1 promoter targeting spacer	e1e3	161	
none	e1e3	121	
≥1 promoter targeting spacer	e2e3	84	
none	e2e3	90	
≥1 promoter targeting spacer	e1e2e3	64	
none	e1e2e3	145	

Figure S17 – Summaries of Library 1 and Library 2 screens.
A) Summary of crRNA constructs in the Library 1 screen.
B) Summary of crRNA constructs in the Library 2, Sublibrary A screen.
C) Summary of crRNA constructs in the Library 2, Sublibrary B screen.

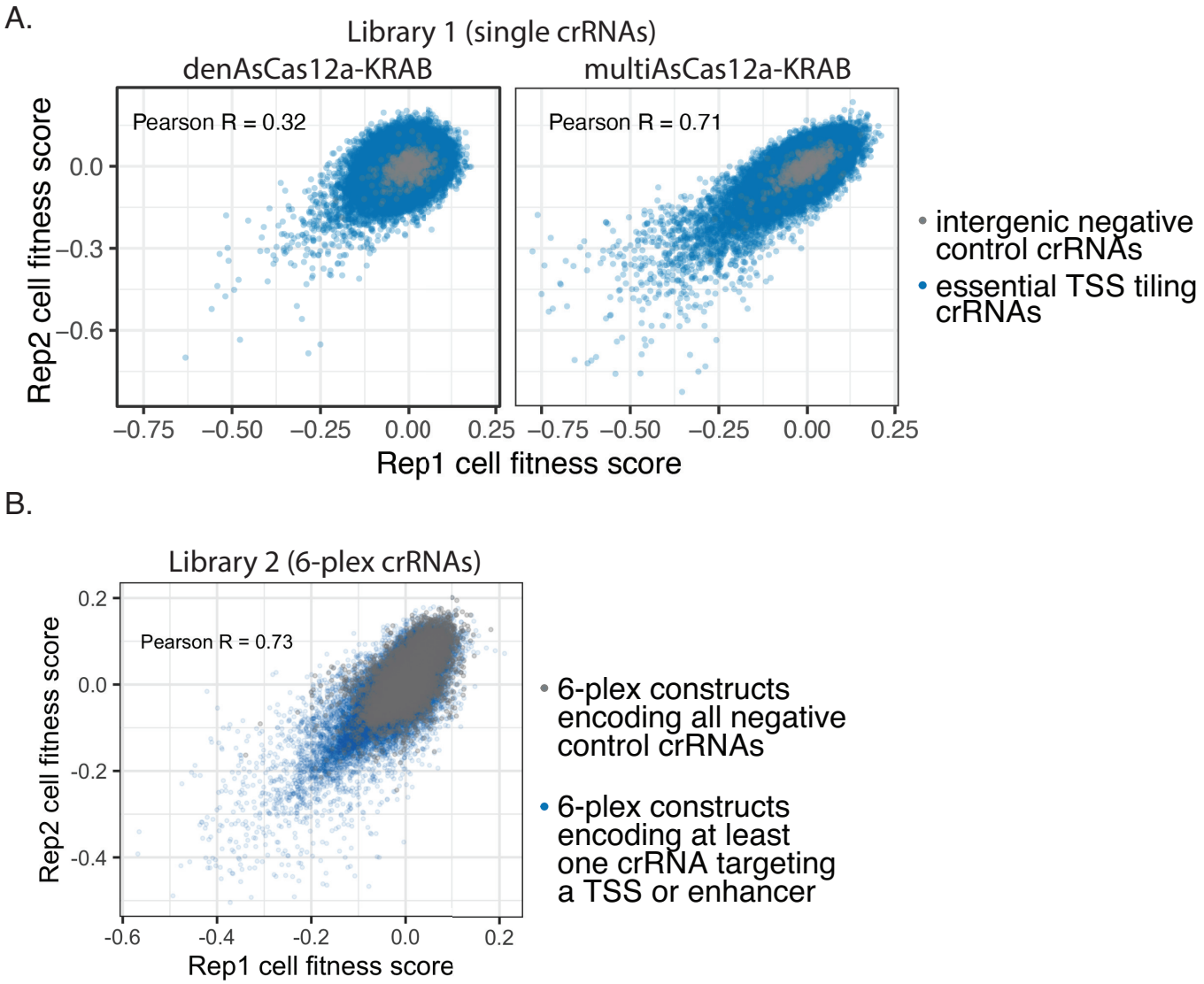


Figure S18 – Screen replicate concordance for Library 1 and Library 2. Shown are 2D density plots of cell fitness scores for individual crRNA constructs in **A)** Library 1 and **B)** Library 2 (Sublibrary A and Sublibrary B), with the Pearson correlation coefficients calculated for all constructs in each library shown.

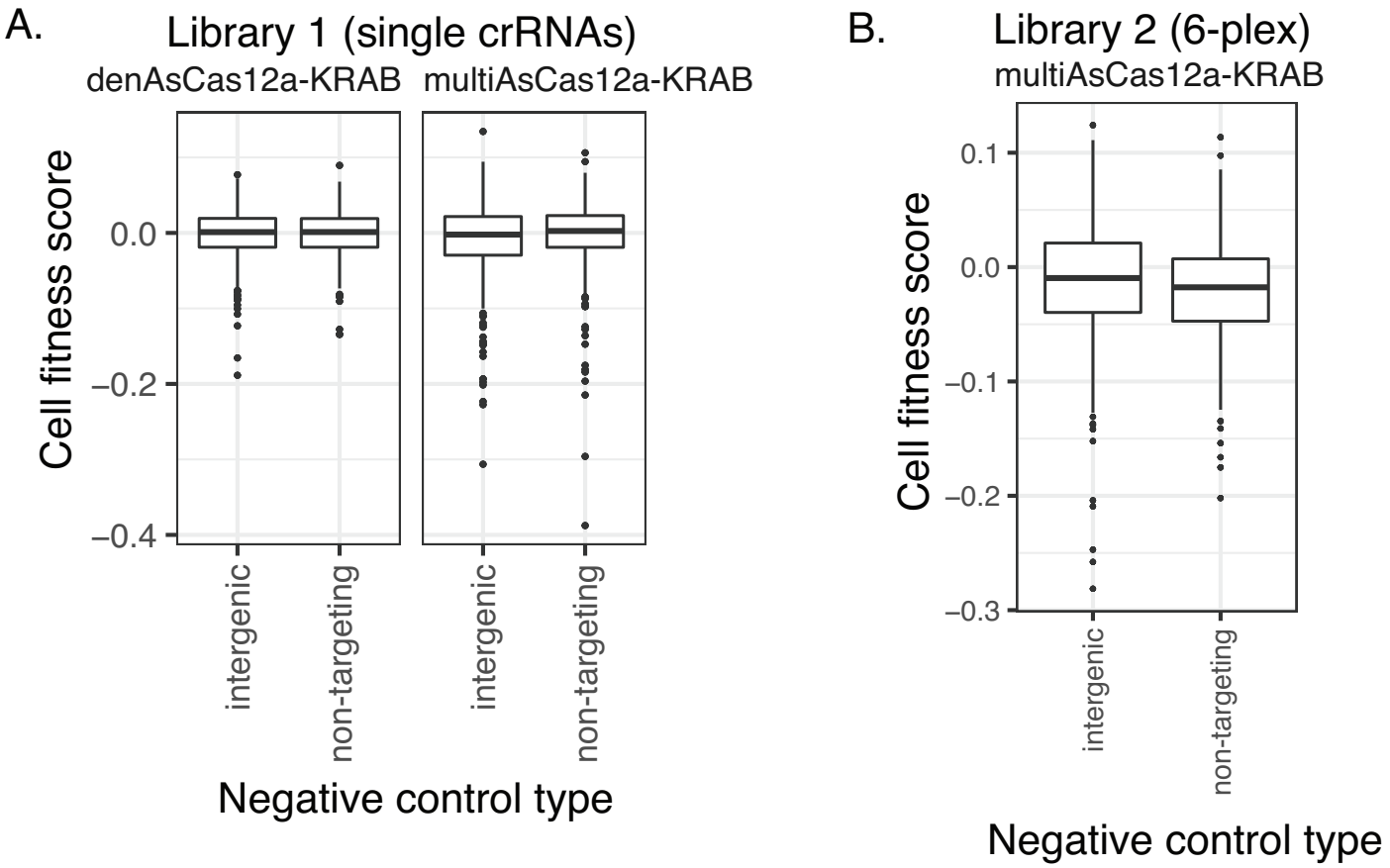


Figure S19 – Cell fitness score distributions of intergenic vs. non-targeting negative control crRNAs. Boxplots of cell fitness scores for **A)** Library 1 single-crRNA constructs, and **B)** Library 2 6-plex constructs, categorized by whether the construct encodes exclusively intergenic vs. non-targeting negative control crRNAs. Boxplots display median, interquartile range, whiskers indicating 1.5x interquartile range, and outliers.

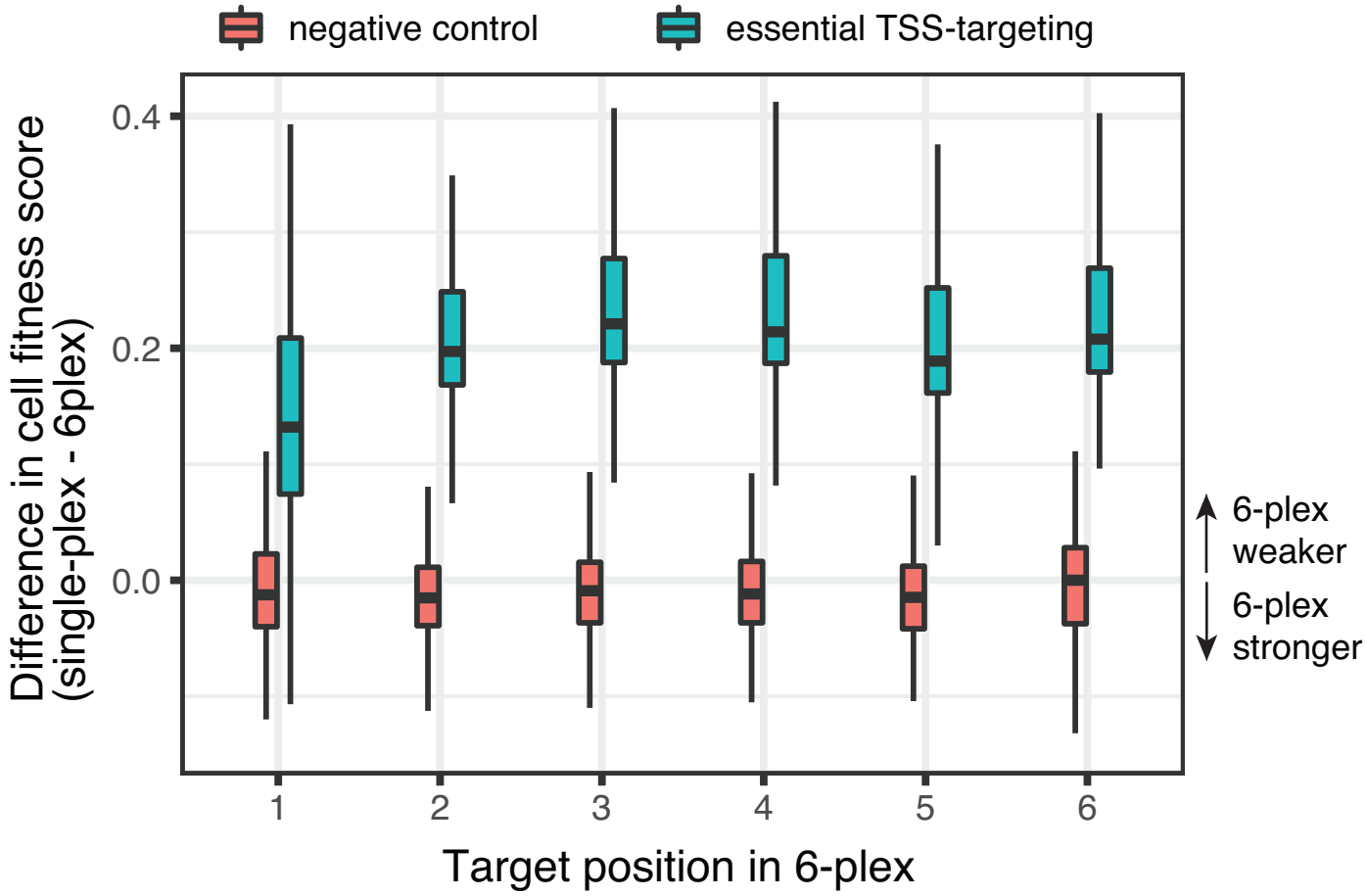


Figure S20 – Difference in cell fitness scores for 6-plex vs. 1-plex crRNA constructs. Library 2 Sublibrary A (6-plex crRNA construct with one test position targeting essential TSS) screen cell fitness scores for a given test position spacer are compared to the cell fitness scores of the same spacer in the Library 1 (single-plex crRNA) screen. Difference in cell fitness scores (6-plex minus single-plex) are shown as boxplots, which display the median, interquartile range, and whiskers indicating 1.5x interquartile range.