



RESEARCH ARTICLE

CRISPR-mediated Synergistic Epigenetic and Transcriptional Control

Antonia A. Dominguez,^{1,†,‡} Michael G. Chavez,^{1,†} Amanda Urke,¹ Yuchen Gao,^{1,2,§} Lizhong Wang,^{3,*} and Lei S. Qi^{1,4,5,*}

Abstract

Targeted activation of endogenous genes is an important approach for cell engineering. Here, we report that the nuclease-deactivated dCas9 fused to a transcriptional activator (VPR) and an epigenetic effector (the catalytic domain of histone acetyltransferase p300^{core}) simultaneously, sequentially, or as a single quadripartite effector can lead to enhanced activation of target genes. The composite activator, VPRP, behaves more efficiently than individual activators across a set of genes in different cell types. We characterize off-target effects for host chromatin acetylation and transcriptome using the effectors. Our work demonstrates that transcriptional and epigenetic effectors can be used together to enhance gene activation and suggests the need for further optimization of epigenetic effectors to reduce off-targets.

Introduction

Nuclease-deactivated dCas (e.g., dCas9) approaches provide the ability to up- and/or downregulate target genes, offering a toolbox for research and potentially gene therapies.^{1–6} Recent work utilizing CRISPR activation (CRISPRa) in mouse models demonstrated the use of *Streptococcus pyogenes* dCas9 (SpdCas9) fused to VP64 to upregulate a haploinsufficient gene to rescue an obesity phenotype and the use of *Staphylococcus aureus* dCas9 (SadCas9) fused to two copies of VP64 to upregulate a compensatory gene for the treatment of congenital muscular dystrophy in a mutation-independent manner.^{7,8} While these studies highlight the potential of CRISPRa to treat diseases caused by transcriptional misregulation, applications for cell reprogramming or metabolic engineering remain limited due to the inability to activate targeted genes efficiently.

Effectively upregulating an endogenous gene with CRISPRa is dependent on the genomic and epigenetic landscape of the gene, necessitating multiple mechanisms of activation. In one commonly used approach, dCas9 is fused to the tripartite transcriptional activator VPR (VP64-p65-Rta), which recruits subunits of the RNA

polymerase to initiate transcription at proximal promoters.¹ Alternatively, dCas9 fused to the catalytic core of p300 histone acetyltransferase (p300^{core}) induces epigenetic enzymatic acetylation at local histone H3 lysine 27 (H3K27), which can activate distal promoters and enhancers.⁵

Hybrid approaches employing both epigenetic and transcriptional activation of a single gene have only begun to be tested.⁹ Further analysis of how expression of dCas9–activator fusions affects targeted and global gene expression and chromatin modifications is needed prior to the therapeutic applications of CRISPRa.

Here, we examine and compare two broadly used effector domains, VPR and p300^{core}, for their efficiency in upregulating target genes and the propensity to cause off-target effects on nontargeted genes. We tested their individual or combinatorial fusions by simultaneously, sequentially, or spatially recruiting VPR and p300^{core} (Fig. 1A). Our data implies a synergistic effect between transcription and epigenetic activators for target gene activation using a composite quadripartite effector VPRP.

We performed RNA sequencing (RNA-seq) and H3K27ac chromatin immunoprecipitation sequencing

¹Department of Bioengineering, ²Cancer Biology Program, ³Department of Chemical and Systems Biology, and ⁵ChEM-H, Stanford University, Stanford, California, USA; and ⁴Department of Genetics, University of Alabama at Birmingham, Birmingham, Alabama, USA.

[†]These authors contributed equally to the work.

[‡]Current affiliation: Sana Biotechnology, South San Francisco, California, USA.

[§]Current affiliation: Mammoth Biosciences, South San Francisco, California, USA.

*Address correspondence to: Lizhong Wang, MD, PhD, Department of Genetics, University of Alabama at Birmingham, Birmingham, AL 35294, Email: lwang12@uab.edu; or Lei S. Qi, PhD, 290 Jane Stanford Way, Stanford University, Stanford, CA 94305, Email: stanley.qi@stanford.edu

(ChIP-seq) of individual VPR, p300^{core}, and VPRP. Our data showed no observable off-targets using VPR on transcriptome level and chromatin modifications, which is consistent with previous studies, but we saw some off-target effects when using p300^{core} individually or in combination. Our profiling implies that while CRISPR-mediated transcription control is specific, the specificity of epigenetic effector p300^{core} can be further optimized.

Methods

Single-guide RNA cloning

Single-guide RNAs (sgRNAs) were expressed from a lentiviral mouse U6-based expression vector that co-expresses Puro-T2A-BFP (pSLQ1373) from an EF1a promoter. New sgRNA sequences were generated by polymerase chain reaction (PCR) and introduced by InFusion cloning into the sgRNA expression vector that was digested with *Bst*XI and *Xho*I (see Supplementary Table S1 for sgRNA sequences).

Plasmid cloning

Individual constructs for dCas9 effectors used in this study are described in Supplementary Table S2, ordered by figure. To assemble the doxycycline (dox)-inducible dCas9-effector constructs, human codon-optimized SpdCas9 (a gift from Martin Jinek) was fused at the C-terminus with an HA tag and two SV40 nuclear localization signals, followed by the effector. The VPR effector¹ was assembled by fusing the activation domain of VP64 with the activation domain of p65 (p65AD) and RTA with two GS linkers with a SV40 NLS between VP64 and p65. The p300^{core} and p300^{core} (D1399Y) effec-

tors were a gift from Charles Gersbach (Addgene plasmid nos. 61357 and 61358).⁵ All transformations containing plasmids with p300^{core} were grown at 30°C.

For visualization, sfGFP was fused at the C-terminus of the effector following a P2A peptide. This cassette is driven by the TRE3G dox-inducible promoter. Zeocin resistance and TetOn 3G transactivator expression is driven by the EF1a promoter. These cassettes were cloned into a PiggyBac plasmid containing PiggyBac homology arms. The SadCas9 (a gift from Feng Zhang; Addgene plasmid no. 61594) constructs were driven by the SSFV promoter and contained a C-terminal P2A mCherry fusion.

Lentiviral production

293LX cells were seeded 24 h prior to transfection. On the day of transfection, the cells reached between 60% and 80% confluence. Lentivirus was produced by transfecting with the pHR-sgRNA expression vector in addition to pMD2.G and pCMV-dR8.91 using TransIT-LT1 transfection reagent (Mirus). A media exchange was conducted 24 h after transfection, and lentivirus was collected 48 h after transfection and filtered through a 0.45 μ m filter. Crude lentivirus was collected and stored immediately at -80°C .

To concentrate the lentivirus, the filtered supernatant was mixed with Lentivirus Precipitation Solution (A1Stem) according to the manufacturer's instructions. Following concentration, lentiviral pellets were resuspended 1/100 of the original volume using cold phosphate-buffered saline (PBS) and stored at -80°C . Crude extract was used to transduce HEK293T and A549 cells, and concentrated lentivirus was used to transduce human induced pluripotent stem cells (hiPSCs).

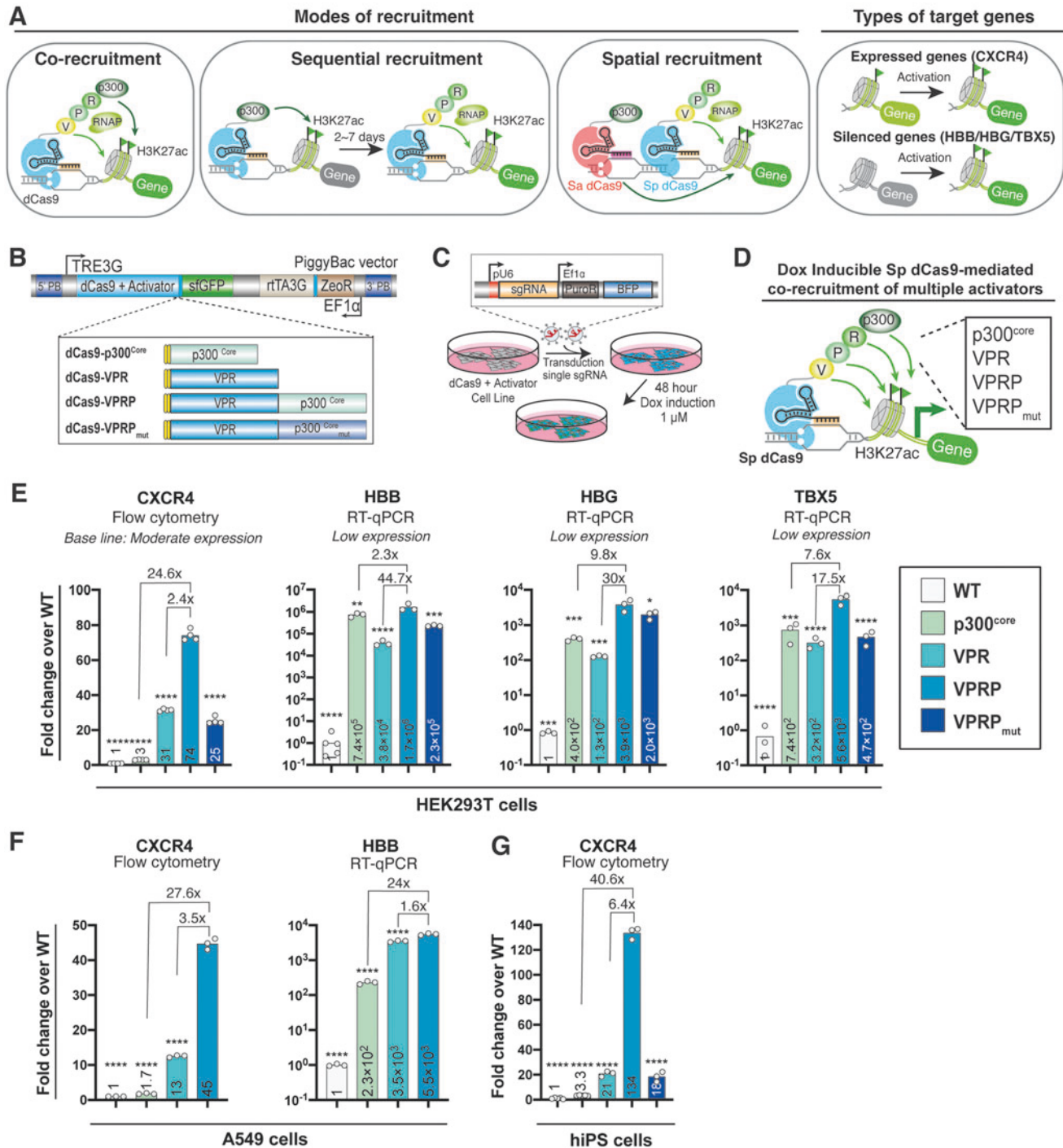
FIG. 1. Direct fusion of VPR and p300^{core} to dCas9. **(A)** Schematic of three modes of activator co-recruitment, from left to right, as a single fusion, sequential recruitment, and spatial recruitment. V, VP64; P, P65; R, Rta. Targeted activation of two types of endogenous genes, expressed or silenced, are shown on the right. **(B)** Constructs encoding the *Streptococcus pyogenes* dCas9 (SpdCas9) activator, including p300^{core}, VPR, VPRP, and VPRP_{mut}, used for stable integration into the host genome utilizing the PiggyBac transposase system. **(C)** Experimental design demonstrating stable integration of single-guide RNA (sgRNA) by lentiviral transduction and activation of inducible dCas9 activator cell line by addition of doxycycline (dox). **(D)** Schematic of target gene activation using co-recruitment of dCas9-p300^{core}, VPR, VPRP, and VPRP_{mut} fusions. **(E)** Comparison of activation efficiency across four activators and a sgRNA targeting C-X-C motif chemokine receptor 4 (*CXCR4*), hemoglobin subunit β (*HBB*), hemoglobin subunit γ (*HBG*), and T-box transcription factor 5 (*TBX5*) in HEK293T cells. Fold change is calculated relative to wild-type (WT) cells (no dCas9 activator) containing the same targeting sgRNA. All data represent three to four biological replicates. **(F)** Activation efficiency across three activators in A549 cells and a sgRNA targeting *CXCR4* and *HBB*. Data represent three biological replicates. **(G)** Activation efficiency across four activators in human induced pluripotent stem cells (hiPSCs) and a sgRNA targeting *CXCR4*. Data represent three to five biological replicates. In **(E)** and **(F)**, * $p < 0.05$; ** $p < 0.01$; **** $p < 0.001$.

Cell culture and stable line generation

HEK293T (Clontech) and A549 (ATCC) cells were cultured at 37°C and 5% CO₂ in Dulbecco's modified Eagle's medium (Gibco) supplemented with 10% fetal bovine serum (FBS). hiPSCs (WTC, Gladstone Stem Cell Core) were cultured at 37°C and 5% CO₂ in MTeSR1 (STEMCELL Technologies) and grown on

Matrigel (Corning). hiPSCs were passaged as single cells using Accutase (STEMCELL Technologies) with the addition of Thiazovivin ROCK inhibitor (STEMCELL Technologies).

To generate HEK293 and A549 cell lines stably expressing the Piggybac-based inducible dCas9-effector constructs, cells were seeded at 2×10^5 cells per well of a



12-well plate. The following day, cells were 60–80% confluent and transfected with 250 ng PiggyBac plasmid containing the dCas9–effector and 100 ng Super PiggyBac transposase (Systems Biosciences) using Mirus TransIT-LT1 reagent. Cells were incubated with the transfection complexes for 48 h. Selection for stable integration was done by the addition of Zeocin for 10 days.

To generate hiPSC cell lines stably expressing the Piggybac-based inducible dCas9–effector constructs, the Mirus TransIT-LT1 Reverse transfection protocol was utilized. Briefly, the transfection complex was mixed and incubated at room temperature for 20 min and then added to a well of a 12-well plate containing 1 mL mTeSR and ROCK inhibitor. hiPSCs (0.65×10^6) were then added to the well containing the transfection complexes and incubated at 37°C and 5% CO₂ for 24 h. Selection for stable integration into HEK293, A549, or hiPSCs was done by the addition of Zeocin for 10 days (see Supplementary Table S3 for all stable lines used in this study).

Flow cytometry analysis

For C-X-C motif chemokine receptor 4 (CXCR4) expression analysis, cells were dissociated with ReLeSR (STEMCELL Technologies) and stained with APC-labeled CXCR4 antibody (BioLegend #30651) in PBS +10% FBS. Cells were washed twice with PBS +10% FBS and then analyzed for fluorescence using a CytoFLEX S flow cytometer (Beckman Coulter). Cells were gated for viability and single cells, and 10,000 cells were collected for the population of interest (GFP+ for the SpdCas9 activation system, BFP+ or mCherry+ for sgRNA, and mCherry+ for the SadCas9 activation system) and analyzed for CXCR4 expression using FlowJo.

Expression fold changes are calculated relative to wild-type (WT) cells + dox and containing a matched sgRNA. All data are displayed as single points or as the mean $\pm$ standard error. A one-way analysis of variance (ANOVA) was used to compare the means between groups and to determine whether any of the means are statistically different.

Quantitative PCR analysis

Hemoglobin subunit β (*HBB*), hemoglobin subunit γ (*HBG*), and T-box transcription factor 5 (*TBX5*) levels were measured using quantitative PCR (qPCR). For sequential and orthogonal activation experiments, all cells were harvested using 0.25 Trypsin-EDTA (Thermo Fisher Scientific) or Accutase (STEMCELL Technologies), and total RNA was isolated using an RNeasy Plus Mini Kit (Qiagen), according to manufacturer's instructions. Reverse transcription was performed using iScript cDNA

Synthesis kit (Bio-Rad). Quantitative PCR reactions were prepared using iTaq Universal SYBR Green Supermix (Bio-Rad) and run on a Bio-Rad CFX384 real-time PCR system (Bio-Rad). Primers are summarized in Supplementary Table S4.

Expression fold changes are calculated relative to WT cells + dox and containing a matched sgRNA. Data represent the average of either three or four technical replicates for each biological replicate. All data are displayed as single points or as the mean $\pm$ standard error. A one-way ANOVA was used to compare the means between groups and to determine whether any of the means are statistically different.

Preparation of samples for RNA-seq and ChIP-seq

RNA-seq was performed from duplicates for each condition. To prepare samples, cells were induced with dox for 48 h, and GFP+/BFP+ double-positive cells were sorted on a BD Influx flow cytometer (BD Biosciences). WT cells containing only a sgRNA were sorted for BFP+ cells. One million cells were sorted for each sample, and total RNA was purified using an RNeasy Plus Mini Kit (Qiagen). Total RNA was sent to Novogene Corporation for library preparation and Illumina PE-150 sequencing.

Briefly, total RNA quantification and qualification was done by Nanodrop for RNA purity, agarose gel electrophoresis, and Agilent 2100 for RNA integrity. mRNA was purified using poly-T oligo-attached magnetic beads and fragmented. cDNA was synthesized using M-MuLV Reverse Transcriptase and DNA Polymerase I and purified using AMPure XP beads.

The 3' ends of the cDNA were adenylated, and a NEB-Next Adaptor with hairpin loop structure was ligated to prepare for hybridization. AMPure XP system (Beckman Coulter) was utilized to enrich for fragments of 150–200 bp length. Finally, the final library was obtained by PCR amplification and purification of PCR products by AMPure XP beads. Libraries were diluted to 1.5 ng/ μ L (Qubit2.0) and insert size detected by Agilent 2100. qPCR is used to quantify the library effective concentration accurately (>2 nM). Samples were then pooled and run on Illumina sequencers.

H3K27ac ChIP-seq was performed from duplicates for each condition. To prepare samples, cells were induced with dox for 48 h. To fix the cells, 1/10 volume of freshly prepared formaldehyde solution (11% formaldehyde, 0.1 M NaCl, 1 mM EDTA, 50 mM HEPES in water) was added to existing media for 15 min at room temperature. The fixation was stopped by adding 1/20 volume of 2.5 M glycine solution for 5 min at room temperature. After the glycine incubation cells were scraped from the culture surface, cells were collected and kept on ice

for the remainder of the procedure. The cells were centrifuged at 800 *g* for 10 min and then re-suspended in 10 mL chilled PBS-IgEPAL. Cells were centrifuged again and re-suspended in 10 mL PBS-IgEPAL and 100 μ L of PMSF (100 mM). Cells were centrifuged a third time, and cell pellets were snap-frozen on dry ice and stored at -80°C .

ChIP-seq was performed by Active Motif. Illumina sequencing libraries were prepared from the ChIP and Input DNAs by the standard consecutive enzymatic steps of end-polishing, dA-addition, and adaptor ligation. After a final PCR amplification step, the resulting DNA libraries were quantified and sequenced on Illumina's NextSeq 500 (75 nt reads, single end).

Bioinformatic analysis of RNA-seq and ChIP-seq

For RNA-seq, mapping and quantification was performed by Novogene. Briefly, mapping was performed with STAR v2.5 against hg19 with a mismatch parameter equal to 2.¹⁰ Initial quantification of transcript abundance was performed by HTSeq v0.6.1.¹¹ We performed differential expression analysis using DESeq2 v1.24.0.¹² MA plots were generated in DESeq2 by contrasting the two conditions using the log fold change shrinkage method.

For ChIP-seq, initial analysis was performed by Active Motif. Reads were aligned to hg19 using the BWA algorithm on default settings and stored in BAM format.¹³ Reads were extended *in silico* at their 3' ends to a length of 200 bp. The genome was divided into 32 nt bins, and fragment density was calculated in each bin. Peak locations were determined using the MACS algorithm, with a cut-off *p*-value of 10^{-7} .¹⁴ Regions were merged across samples for comparison across samples and stored in bigWig files. We imported these bigWig files directly to deepTools to create correlation results and visualization of H3K27Ac about a genomic locus.¹⁵

Results

Direct fusion of VPR and p300^{core} to dCas9 shows enhanced activation across genes and cell types

To allow for inducible expression of the dCas9 activators, we developed a dox-inducible system on a single Piggy-Bac transposon plasmid that can be stably integrated into the host genome (Fig. 1B). This construct was stably integrated via transposition into the genome of HEK293T cells, and sgRNA driven by a U6 promoter was introduced by lentiviral transduction (Fig. 1C).

To test our first strategy for the coordinated recruitment of VPR and p300^{core}, we developed four dCas9 activators (Fig. 1D). Single activators included VPR or p300^{core} fused to the C-terminus of SpdCas9. Our coordinated activators included the fusion of p300^{core} or p300^{core}_{mut}, the catalytically inactive (D1399Y) form of

p300^{core}, to the C-terminus of VPR, which resulted in VPRP or VPRP_{mut}, respectively (Fig. 1B and D). The dCas9 activators were induced by the addition of dox, and cells were assayed 48 h later by flow cytometry or quantitative reverse transcription (qRT)-PCR (Fig. 1C). The level of activation is calculated as fold-change over WT cells that do not contain an activation system but do contain the targeting sgRNA.

To test activation across a panel of genes, we chose *CXCR4*, *HBB*, *HBG*, and *TBX5*. All these genes are relevant to disease modeling or gene therapy. For *CXCR4*, since it is a surface protein with an available antibody for immunostaining, we performed flow cytometry to quantify its expression on the protein level in single cells. For other genes, we performed qPCR to quantify expression on the transcript level in a bulk cell population. dCas9-VPRP activated *CXCR4*, a gene that is moderately expressed in HEK293T cells, by 74-fold (Fig. 1E). In comparison, dCas9-VPR and dCas9-VPRP_{mut} activated *CXCR4* by only 31- and 25-fold, respectively. Notably, dCas9-p300^{core} activated *CXCR4* by only threefold.

For *HBB*, *HBG*, and *TBX5*, which are all expressed at low levels in HEK293T cells, we observed that the levels of activation were significantly higher than *CXCR4*. dCas9-VPRP activated *HBB* by 1.7×10^6 -fold, *HBG* by 3,900-fold, and *TBX5* by 5,600-fold (Fig. 1E). Interestingly, p300^{core} outperformed VPR in terms of activation efficiency for all three genes, which was different from activation of *CXCR4*. This likely suggests that p300^{core} might be more potent at originally silenced genes, while VPR works better at already expressed genes.

We also observed comparable activation efficiency of using VPRP_{mut} and VPR, thus confirming the additional activation of VPRP was indeed induced by the acetyltransferase catalytic activity of the p300^{core}. Overall, VPRP showed a positive combinatorial effect (more than additive) of gene activation compared to VPR or p300^{core} alone, suggesting cooperativity between transcription and epigenetic regulation.

Besides HEK293T cells, we observed similarly improved activation using dCas9-VPRP in A549 lung epithelial carcinoma cells and hiPSCs (Fig. 1F and G). These gene activation effects were sgRNA dependent, as demonstrated using a nontargeting sgRNA (Supplementary Fig. S1A and B).

Sequential recruitment of VPR and p300^{core} enhances transcriptional activation at silenced genes at earlier time points than expressed genes

We next tested sequential recruitment of p300^{core} and VPR to the same genomic site to allow for histone acetylation prior to transcriptional activation, which may

increase chromatin accessibility and improve activation of the target gene compared to either modality alone (Fig. 2A).¹⁶

To express SpdCas9 fused to p300^{core} or p300^{core}_{mut} in the dox-inducible dCas9-VPR HEK293T cells, we

transiently transfected cells with constructs encoding SpdCas9-p300^{core}-P2A-mCherry or SpdCas9-p300^{core}_{mut}-P2A-mCherry driven by the constitutive SFFV promoter (Fig. 2A). Targeted histone modifications by dCas9-p300^{core} were permitted to occur before dox induction of

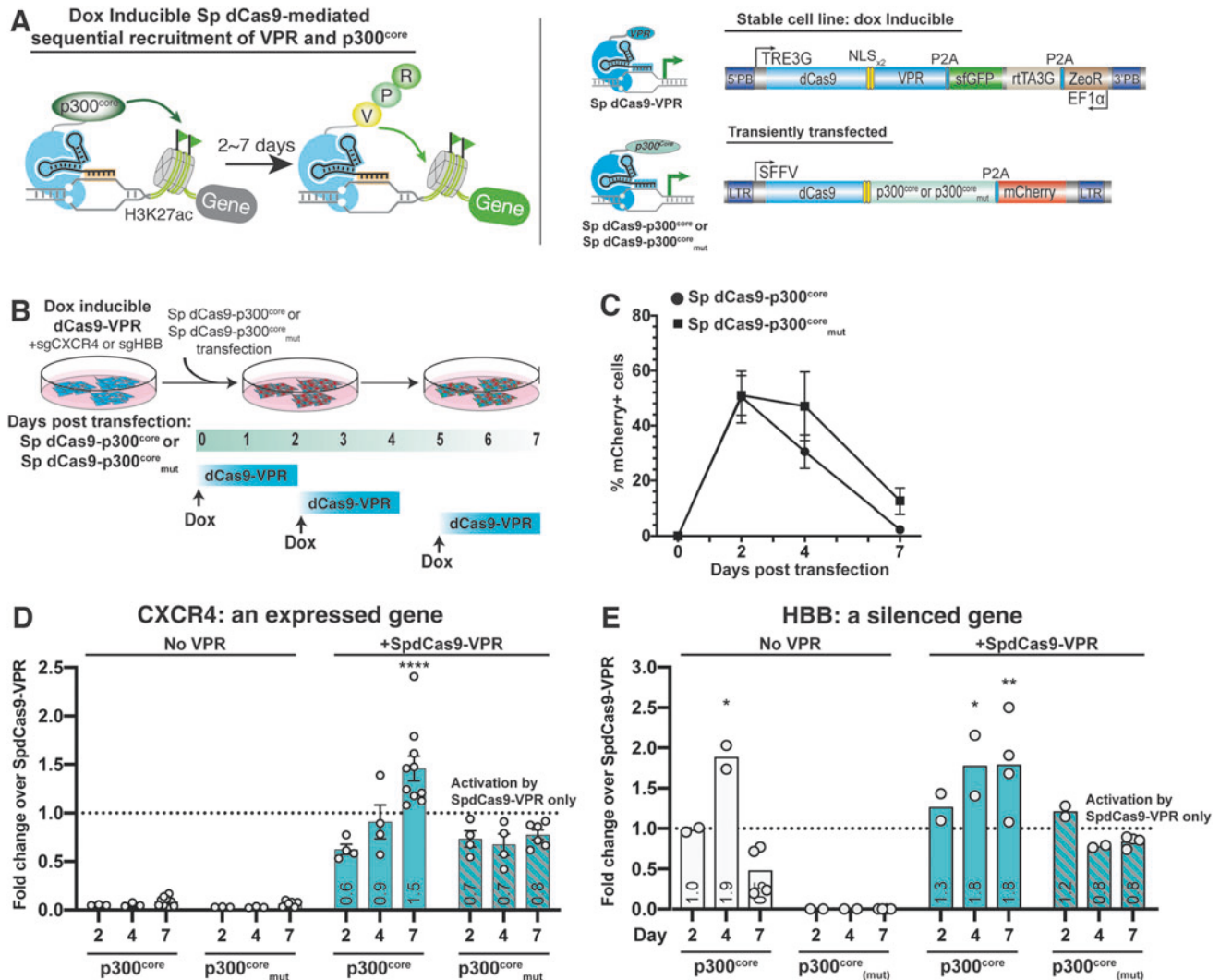


FIG. 2. Testing sequential recruitment of dCas9-VPR and dCas9-p300^{core} for endogenous gene activation. **(A)** Schematic of sequential recruitment of VPR and p300^{core}. The experiment was performed using transient expression of dCas9-p300^{core} or dCas9-p300^{core}_{mut} in stable cell lines encoding dCas9-VPR for dox-inducible expression. Constructs encoding inducible dCas9-VPR and constitutively expressed dCas9-p300^{core} or dCas9-p300^{core}_{mut} are shown. **(B)** Experimental design testing gene activation for sequential recruitment of dCas9-VPR and dCas9-p300^{core} or dCas9-p300^{core}_{mut} using the same targeting sgRNA. **(C)** Stability of transient expression of SpdCas9-p300^{core} and SpdCas9-p300^{core}_{mut} over time. **(D)** Results of sequential activation on CXCR4. Time indicates 2 days after dox induction of dCas9-VPR. Fold change is calculated relative to activation by dCas9-VPR alone and is indicated by the dashed line. Data represent between 4 and 10 biological replicates. *****p* < 0.001. **(E)** Results of sequential activation on HBB. Time indicates 2 days after dox induction of dCas9-VPR. Fold change is calculated relative to activation by dCas9-VPR alone and is indicated by the dashed line. Data represent between two to four biological replicates. **p* < 0.05; ***p* < 0.01; *****p* < 0.001.

dCas9-VPR for 0, 2, or 5 days (Fig. 2B). Transient expression of dCas9-p300^{core} in cells, measured by mCherry, peaked on day 2 and dropped to a low level by day 7 (Fig. 2C). Cells were assayed for gene activation 48 h after induction of dCas9-VPR on days 2, 4, or 7, respectively.

For *CXCR4* (an expressed gene), while time points day 2 or 4 did not show increased activation relative to dCas9-VPR alone, we saw a modest increase (1.5-fold) in *CXCR4* activation only later on day 7 (Fig. 2D). For *HBB* (a silenced gene) targeting, the sequential recruitment showed better activation at earlier time points starting on day 2 (Fig. 2E). This observation likely suggests deposition of H3K27ac to local chromatin facilitates transcriptional activation, which is more pronounced for originally silenced chromatin locus (lacking H3K27ac) compared to expressed genes (already possessing H3K27ac). The p300^{core}_{mut} fusion did not increase the target gene activation compared to dCas9-VPR at all time points, confirming that these effects are attributable to the catalytic activity of the p300^{core}.

Spatial co-recruitment of VPR and p300^{core} enhances transcriptional activation at silenced genes but not expressed genes

We next tested spatial co-recruitment of p300^{core} and VPR effectors. We used an orthogonal pair of dCas9, SadCas9 and SpdCas9, which do not interact with each other's sgRNA and protospacer adjacent motifs. We co-expressed the SadCas9 fused to p300^{core} or p300^{core}_{mut} with SpdCas9-VPR in the same cells (Fig. 3A).¹⁷ The sgRNAs are designed to be overlapping (Sa-C1 and Sp-C1 for *CXCR4*, Sa-H2 and Sp-H1 for *HBB*) or nonoverlapping (Sa-C2 and Sp-C1 for *CXCR4*, Sa-H1 and Sp-C1 for *HBB*; Fig. 3B). It should also be noted that in our previous study, we observed that SpdCas9 exhibited a higher activity for gene activation than SadCas9 using the same activator.¹⁸

We did not observe better gene activation when co-recruiting SadCas9-p300^{core} and SpdCas9-VPR for already expressed gene *CXCR4* compared to SpdCas9-VPR alone (dotted line in Fig. 3C). However, spatial co-recruitment of SadCas9-p300^{core} and SpdCas9-VPR resulted in improved activation than SpdCas9-VPR alone for silenced gene *HBB* (5.1-fold for the nonoverlapping sgRNAs and 2.1-fold for the overlapping sgRNAs; orange bars in Fig. 3D). The nonoverlapping guide (Sa-H2) showed better enhancement for *HBB* activation compared to the overlapping guide (Sa-H1), which is consistent with less steric hindrance between the two dCas9/sgRNA complexes at the local chromatin.

The p300^{core}_{mut} showed no enhancement over VPR alone for *CXCR4* or *HBB*, suggesting this enhancement of activation was dependent on the catalytic function of

p300^{core}. These results are consistent with temporal recruitment that deposition of H3K27ac to silenced genes but not actively expressing genes enhances gene activation.

Genome-wide acetylation characterization of CRISPRa fusions

We next carried out genome-wide measurement on transcriptomic and epigenetic levels to characterize the quadripartite activator dCas9-VPRP system and compared its performance to VPR, p300^{core}, and p300^{core}_{mut} fusions. We performed H3K27ac chromatin immunoprecipitation sequencing (ChIP-seq) to profile genome-wide acetylation and RNA sequencing to assay global transcriptome changes in biological replicates, including WT HEK293T cells expressing an *HBB*-targeting sgRNA and HEK293T cells expressing dCas9-VPR, dCas9-p300^{core}, dCas9-VPRP, and dCas9-VPRP_{mut}. We chose to activate *HBB* in these samples (rather than *CXCR4*) because *HBB* activation likely has less indirect regulatory effects on other genes.

Measurement of H3K27ac confirmed that dCas9 fused to p300^{core} or VPRP acetylated genomic regions around the *HBB* locus (Fig. 4A). We compared the genome-wide H3K27ac profiles by correlating the H3K27ac peak intensities between samples and observed that samples expressing p300^{core} fusions (including p300^{core} alone and VPRP) segregated from samples without the acetyltransferase activity (WT, VPR, and VPRP_{mut}; Fig. 4B).

We analyzed the peak intensities of each fusion protein compared to WT HEK293T cells (Fig. 4C). While dCas9-VPR and dCas9-VPRP_{mut} demonstrated somewhat similar H3K27ac profiles to the WT cells, surprisingly, samples with p300^{core} (p300^{core} and VPRP) showed deacetylation across the genome. This is likely due to a nonspecific function of overexpressing exogenous p300^{core} in human cells, which may induce proteomic interactions between p300^{core} and other co-factors. It should also be noted that we observed moderate increase of H3K27ac peak intensities around the sgRNA target site, suggesting the inactive D1399Y form of p300^{core} may still possess residual activity when used with VPR.

We analyzed putative off-target genomic locations with altered H3K27ac peak intensities in the sample of dCas9-VPRP. We submitted the sequence of sgHBB to Cas-OFF Finder, allowing for up to 3 bp mismatches and no DNA or RNA bulges, to obtain all predicted off-target sites.¹⁹ The average ChIP-seq score for the locations 1 kb up- and downstream (2 kb total window) of each putative off-target binding site was calculated using deepTools.¹⁵ Interestingly, a majority of predicted off-targets showed no apparent increase of acetylation around putative off-binding sites, as most of the data

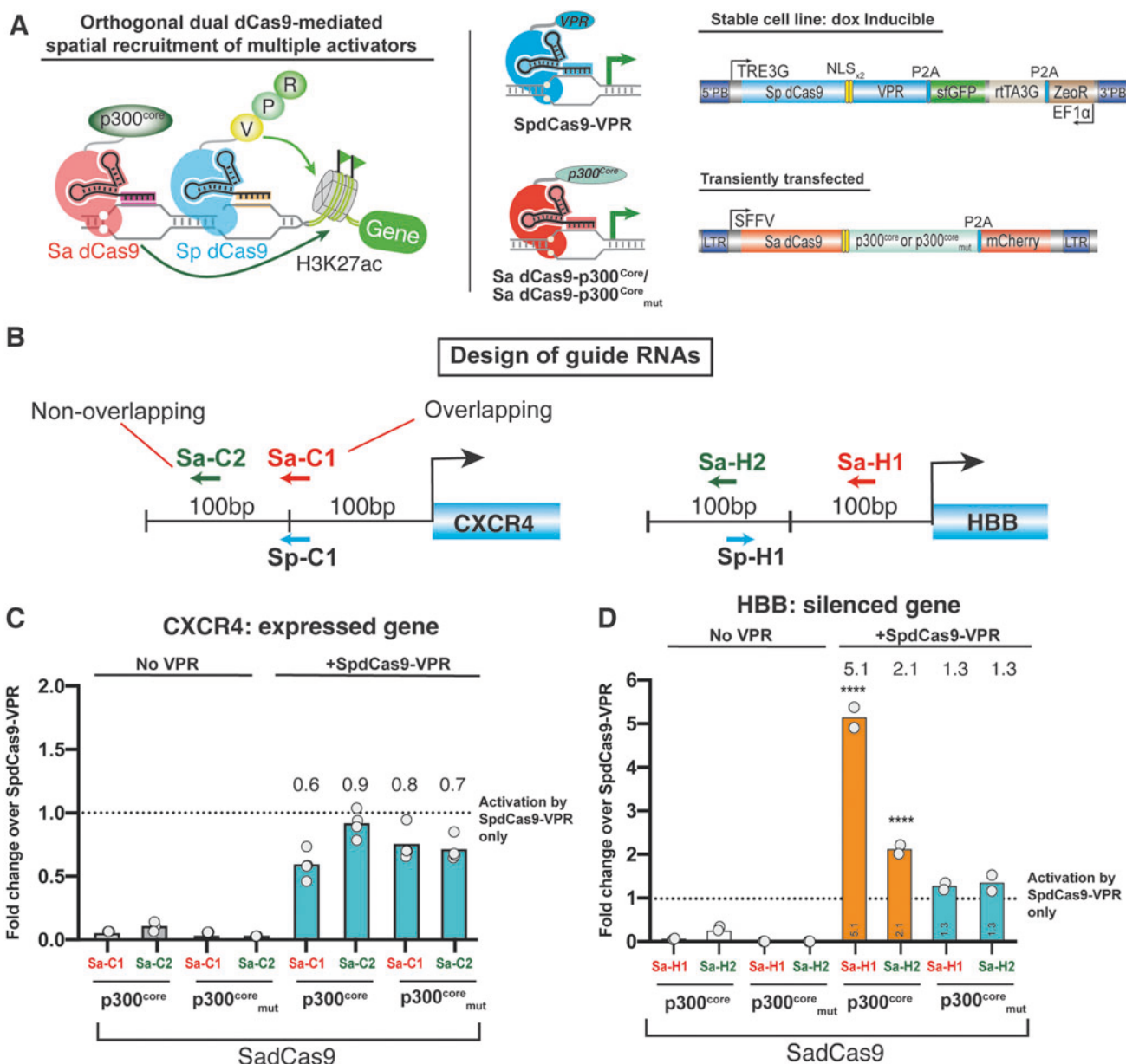


FIG. 3. Testing spatial recruitment of VPR and p300^{core} with dCas9 orthologs for endogenous gene activation. **(A)** Schematic of spatial recruitment of VPR and p300^{core}. The experiment was performed using transient expression of *Staphylococcus aureus* dCas9 (SadCas9)-p300^{core} or SadCas9-p300^{core}_{mut} in stable cell lines encoding SpdCas9-VPR for dox-inducible expression. Constructs encoding inducible SpdCas9-VPR and constitutively expressed SadCas9-p300^{core} or SadCas9-p300^{core}_{mut} are shown. **(B)** sgRNAs used in orthogonal dCas9-mediated activation of *CXCR4* and *HBB*. SadCas9 sgRNAs shown in purple and SpdCas9 targeting sgRNAs are shown in blue. **(C)** Results of orthogonal activation on *CXCR4*. Time indicates 2 days after dox induction of SpdCas9-VPR. Fold change is calculated relative to activation by SpdCas9-VPR alone and is indicated by the dashed line. Data represent three to four biological replicates. **p* < 0.05. **(D)** Results of orthogonal activation on *HBB*. Time indicates 2 days after dox induction of SpdCas9-VPR. Fold change is calculated relative to activation by SpdCas9-VPR alone and is indicated by the dashed line. Data represent two biological replicates. *****p* < 0.001.

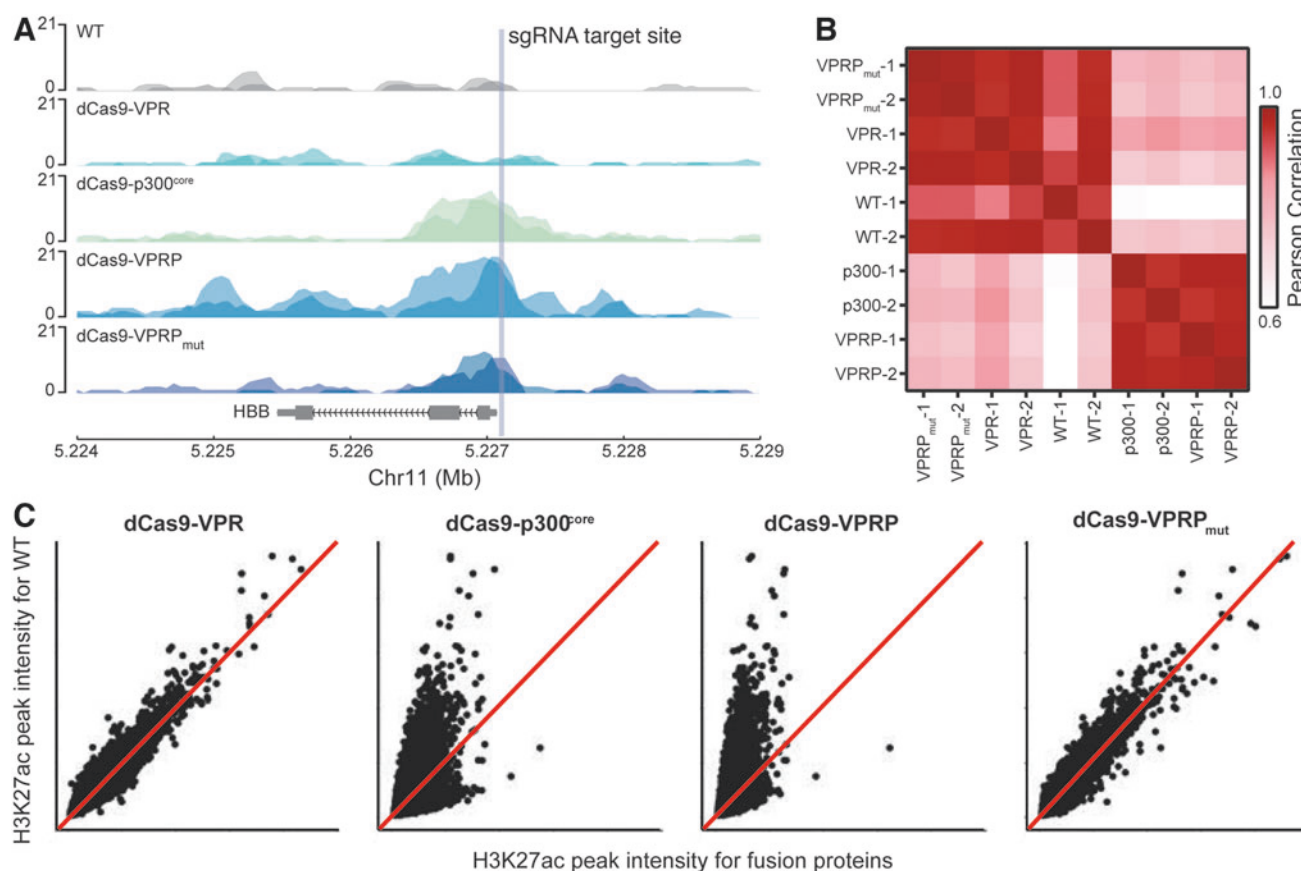


FIG. 4. Whole-genome chromatin acetylation analysis in cells with VPR, p300^{core}, VPRP, and VPRP_{mut}. **(A)** H3K27ac enrichment at the *HBB* locus as measured by ChIP-seq. Two colors in each group indicate two technical replicates. **(B)** Pairwise correlation of genome-wide H3K27ac peak intensities across all samples. Each biological group has two replicates. Peak intensities and correlations generated by deepTools. **(C)** H3K27ac peak intensities of each fusion compared to WT cells. Black dots represent the intensity at a specific locus.

exists closer to genome-wide average acetylation (blue line in Supplementary Fig. 2A) than the acetylation score around sgHBB locus (red line in Supplementary Fig. 2A).

We further zoomed out the *HBB* locus in Figure 4A to visualize a ~60 kb window around the sgRNA binding site (Supplementary Fig. 2B). Importantly, H3K27ac marks were clustered around the transcriptional start site of the cells expressing dCas9-VPRP compared to WT cells. There were no apparent H3K27ac peaks outside of the targeted regions, indicating that off-target increases in acetylation were not clustered around the bound locus.

Genome-wide transcriptomic characterization of CRISPRa fusions

We analyzed RNA-seq of cells expressing the dCas9 fusions targeted to *HBB*. Corroborating the qRT-PCR data

for gene activation (Fig. 1E–G), while WT cells and dCas9-VPRP with a nontargeting sgRNA displayed no *HBB* expression, the other dCas9 fusions upregulated *HBB* mRNA levels, with dCas9-VPRP being the most efficient (Fig. 5A).

Differential expression analysis of cells expressing dCas9-VPRP with an sgRNA targeting *HBB* versus a nontargeting sgRNA indicated that there were a few off-target effects, which was consistent with the previous results for similar comparisons between targeting versus nontargeting sgRNAs (Fig. 5B).⁵ Most differentially expressed genes belonged to low expressing genes, suggesting their difference can be largely due to stochastic gene expression. Similar as the genome-wide acetylation profiles, the transcriptomes of samples clustered depending on whether the active p300^{core} was present (Fig. 5C), indicating p300^{core} itself affected global transcript expression.

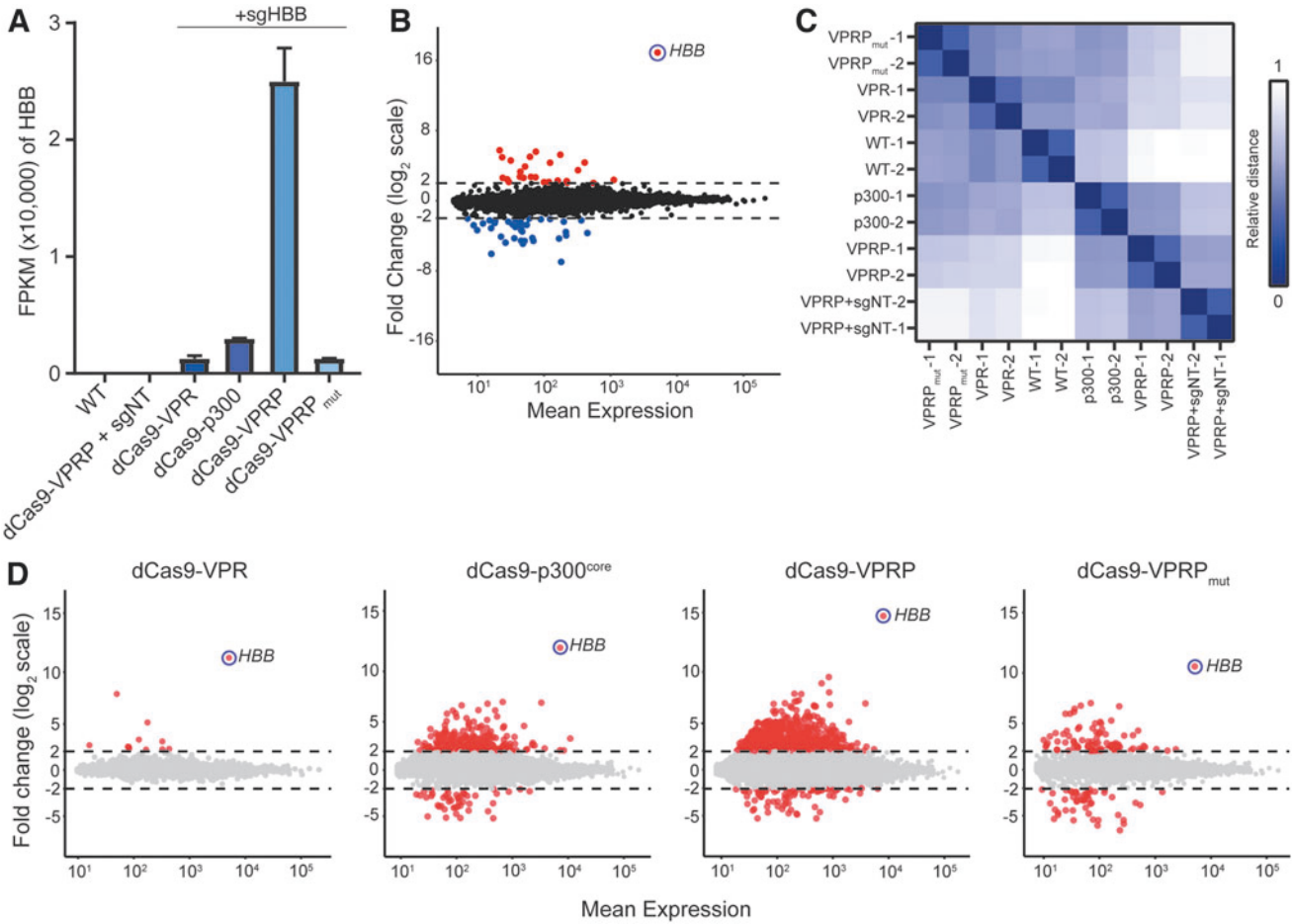


FIG. 5. Whole-cell transcriptome analysis in cells with VPR, p300^{core}, VPRP, and VPRP_{mut} using RNA-seq. **(A)** HTSeq defined FPKM of *HBB* in HEK293T cells using various CRISPR activation (CRISPRa) fusions targeting *HBB* as well as WT HEK293T cells and dCas9-VPRP with a nontargeting guide (sgNT). **(B)** Differential expression analysis between dCas9-VPRP using sgRNA targeting *HBB* versus a nontargeting sgRNA. **(C)** Pairwise correlation of whole transcriptome between samples. **(D)** Differential expression analysis between the dCas9 fusions targeting *HBB* and WT cells in HEK293T cells. Plots were generated by DESeq2. Dotted line represents a 2 log₂ fold change, red dots represent genes that were >2 or <-2 fold change and are statistically significant.

We further analyzed the 26 genes that were differentially upregulated by dCas9-VPRP using sgHBB compared to sgNT (red dots in Fig. 5B) and investigated whether these sites were due to off-target binding of sgRNA that targets *HBB*. We submitted the sequence for sgHBB to Cas-OFF Finder, allowing for up to 5 bp mismatches and no DNA or RNA bulges.¹⁹ We searched for each of the 26 genes a sequence range between 10 kb upstream of the transcriptional start site to 10 kb downstream of the transcriptional termination site. Other than *HBB*, *COL16A1* was the only gene that was upregulated in dCas9-VPRP + sgHBB samples that had an off-target binding site within this window. However, the off-target binding site had four mismatches on the dis-

tal end of the sgRNA (cAcaAcCCAAGGACAGGTA; lower case nucleotides are mismatches) and was 13 kb downstream of the transcriptional start site within the gene coding sequence. The general lack of off-target binding sites around the differentially expressed genes likely suggest the differentially expressed genes are not due to the off-target binding. Instead, they are likely due to long-range interactions from VPRP's chromatin alteration. We further analyzed the dCas9 fusions compared to the WT cells. We observed that dCas9-VPR showed very few off-target effects compared to the WT cells, with most changed genes being low differentially expressed genes (log₂ fold change of >2 or <-2). In

comparison, cells expressing dCas9-p300^{core} and dCas9-VPRP showed changes of many transcripts (Fig. 5D). The VPRP_{mut} fusion showed greatly decreased number of off-target effects compared to its active counterpart, further suggesting that it was the catalytic function of p300^{core} that induced global transcriptome change.

Together, the data imply that almost all off-target effects of p300^{core} or VPRP are induced by the effector p300^{core} rather than the dCas9 or VPR system. The detailed characterization of the effectors on both chromatin modification and RNA sequencing levels provides a resource for their future optimization with reduced off-target effects.

Discussion

In this study, we report that a transcriptional activator, VPR, and an epigenetic activator, the catalytic unit of the histone acetyltransferase p300^{core}, can synergistically enhance target gene activation. We designed a set of co-recruitment experiments to recruit both effectors temporally or spatially, which suggested that the epigenetic modification H3K27ac can enhance VPR for transcriptional activation. We show that while VPR is more effective for activating already expressed genes (e.g., *CXCR4*), p300^{core} is better at activating those silenced genes (e.g., *HBB*). This makes our engineered VPRP particularly suitable for activating epigenetically silenced endogenous chromatin and genes.

The combinatorial enhancing effects are most pronounced for genes that were originally silenced, such as *HBB*, compared to those already expressed genes, such as *CXCR4*. The quadripartite effector VPRP showed stronger activation of endogenous genes than VPR or p300^{core}, which can be potentially used to activate genes for applications that require overexpression of the target genes such as in cell reprogramming or metabolic engineering.

Our genome-wide ChIP-seq and RNA-seq confirmed the introduction H3K27ac marks to local histones and enhanced transcription. We observed almost no off-targets when using the transcriptional effector VPR but some off-targets when using the epigenetic effector p300^{core}, which is consistent with a recent study on the off-target effects of p300^{core} domain fusion.²⁰ Given the off-target effects on H3K27ac of the VPRP activator, we recommend using it with caution for research. For therapeutics, further protein engineering is required to reduce the off-target effects.

Furthermore, the increased payload size of VPRP compared to VPR or VP64 should be considered for use on delivery formulations such as lentivirus or mRNA delivery. The trade-off between increased efficacy for gene ac-

tivation and decreased delivery efficiency is common to effector engineering, which should be carefully evaluated for each application.

While recent work has extensively characterized the off-target effects of CRISPR-Cas nucleases and base editors, more characterization of epigenetic and transcriptional systems is necessary in the future.^{21–25} There is a need to engineer better epigenetic effectors further such that they exhibit high efficiency for desired function (transcription control or different types of epigenetic modifications) and low off-target effects. Recent work showed that introducing mutations into the DNA binding regions of the DNMT3 protein can reduce its off-target effects,^{26,27} suggesting further protein engineering is important to adapt broad epigenetic effectors with minimized off-target effects.

Acknowledgments

The authors thank the Qi lab members for reading and commenting on the manuscript. Cell sorting for this project was done on instruments in the Stanford Shared FACS Facility. Cloning and sequencing oligos were synthesized at the Protein and Nucleic Acid (PAN) facility at Stanford. The data sets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request. Plasmids related to the work are available on Addgene (https://www.addgene.org/Stanley_Qi/).

Author Disclosure Statement

L.S.Q. is a co-inventor on patent applications filed by Stanford University relating to the work in this study. L.S.Q. is founder and advisor of Epicrispr Biotechnologies.

Funding Information

A.A.D. received support from fellowship support from the ALS Foundation and a Postdoc Enrichment Program Fellowship from Burroughs Wellcome Fund. M.C. and Y.G. received fellowship support from NSF Graduate Research Fellowships Program (GRFP). L.S.Q. acknowledge support from Pew Charitable Trusts, Alfred P. Sloan Foundation, and Li Ka Shing Foundation. The work is supported by Department of Defense Breast Cancer Research Program Breakthrough Award (W81XWH-17-1-0018), National Institutes of Health Common Fund 4D Nucleome Program (Grant No. U01 DK127405), the National Science Foundation CAREER Award under Grant No. 2046650.

This project was partly supported by the Stanford Maternal and Child Health Research Institute through the Uytengsu-Hamilton 22q11 Neuropsychiatry Research Award Program.

Supplementary Material

Supplementary Figure S1
 Supplementary Figure S2
 Supplementary Table S1
 Supplementary Table S2
 Supplementary Table S3
 Supplementary Table S4

References

- Chavez A, Scheiman J, Vora S, et al. Highly efficient Cas9-mediated transcriptional programming. *Nat Methods* 2015;12:326–328. DOI: 10.1038/nmeth.3312.
- Qi LS, Larson MH, Gilbert LA, et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* 2013;152:1173–1183. DOI: 10.1016/j.cell.2013.02.022.
- Gilbert LA, Larson MH, Morsut L, et al. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* 2013;154:442–451. DOI: 10.1016/j.cell.2013.06.044.
- Yeo NC, Chavez A, Lance-Byrne A, et al. An enhanced CRISPR repressor for targeted mammalian gene regulation. *Nat Methods* 2018;15:611–616. DOI: 10.1038/s41592-018-0048-5.
- Hilton IB, D'Ippolito AM, Vockley CM, et al. Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat Biotechnol* 2015;33:510–517. DOI: 10.1038/nbt.3199.
- Dominguez AA, Lim WA, Qi LS. Beyond editing: repurposing CRISPR-Cas9 for precision genome regulation and interrogation. *Nat Rev Mol Cell Biol* 2016;17:5–15. DOI: 10.1038/nrm.2015.2.
- Matharu N, Rattanasopha S, Tamura S, et al. CRISPR-mediated activation of a promoter or enhancer rescues obesity caused by haploinsufficiency. *Science* 2019;363:eaau0629. DOI: 10.1126/science.aau0629.
- Kemaladewi DU, Bassi PS, Erwood S, et al. A mutation-independent approach for muscular dystrophy via upregulation of a modifier gene. *Nature* 2019;572:125–130. DOI: 10.1038/s41586-019-1430-x.
- Li K, Liu Y, Cao H, et al. Interrogation of enhancer function by enhancer-targeting CRISPR epigenetic editing. *Nat Commun* 2020;11:485. DOI: 10.1038/s41467-020-14362-5.
- Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 2013;29:15–21. DOI: 10.1093/bioinformatics/bts635.
- Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* 2015;31:166–169. DOI: 10.1093/bioinformatics/btu638.
- Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014;15:550. DOI: 10.1186/s13059-014-0550-8.
- Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 2009;25:1754–1760. DOI: 10.1093/bioinformatics/btp324.
- Feng J, Liu T, Zhang Y. Using MACS to identify peaks from ChIP-Seq data. *Curr Protoc Bioinformatics* 2011;34:2.14. DOI: 10.1002/0471250953.bi0214s34.
- Ramírez F, Ryan DP, Grüning B, et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* 2016;44:W160–W165. DOI: 10.1093/nar/gkw257.
- Andersson R, Sandelin A, Danko CG. A unified architecture of transcriptional regulatory elements. *Trends Genet* 2015;31:426–433. DOI: 10.1016/j.tig.2015.05.007.
- Ran FA, Cong L, Yan WX, et al. *In vivo* genome editing using *Staphylococcus aureus* Cas9. *Nature* 2015;520:186–191. DOI: 10.1038/nature14299.
- Gao Y, Xiong X, Wong S, et al. Complex transcriptional modulation with orthogonal and inducible dCas9 regulators. *Nat Methods* 2016;13:1043–1049. DOI: 10.1038/nmeth.4042.
- Bae S, Park J, Kim JS. Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics* 2014;30:1473–1475. DOI: 10.1093/bioinformatics/btu048.
- Gemberling MP, et al. Transgenic mice for *in vivo* epigenome editing with CRISPR-based systems. *Nat Methods* 2021;18:965–974. DOI: 10.1038/s41592-021-01207-2.
- Wienert B, Wyman SK, Richardson CD, et al. Unbiased detection of CRISPR off-targets *in vivo* using DISCOVER-Seq. *Science* 2019;364:286–289. DOI: 10.1126/science.aav9023.
- Kim D, Kim D, Lee G, et al. Genome-wide target specificity of CRISPR RNA-guided adenine base editors. *Nat Biotechnol* 2019;37:430–435. DOI: 10.1038/s41587-019-0050-1.
- Gaudelli NM, Komor AC, Rees HA, et al. Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature* 2017;551:464–471. DOI: 10.1038/nature24644.
- Gaudelli NM, Lam DK, Rees HA, et al. Directed evolution of adenine base editors with increased activity and therapeutic application. *Nat Biotechnol* 2020;38:892–900. DOI: 10.1038/s41587-020-0491-6.
- Xu X, Chemparathy A, Zeng L, et al. Engineered miniature CRISPR-Cas system for mammalian genome regulation and editing. *Mol Cell* 2021;81:4333–4345.e4. DOI: 10.1016/j.molcel.2021.08.008.
- Hofacker D, Broche J, Laistner L, et al. Engineering of effector domains for targeted DNA methylation with reduced off-target effects. *Int J Mol Sci* 2020;21:502. DOI: 10.3390/ijms21020502.
- Pflueger C, Tan D, Swain T, et al. A modular dCas9-SunTag DNMT3A epigenome editing system overcomes pervasive off-target activity of direct fusion dCas9-DNMT3A constructs. *Genome Res* 2018;28:1193–1206. DOI: 10.1101/gr.233049.117.