Tailoring a CRISPR/Cas-based Epigenome Editor for Programmable Chromatin Acylation and Decreased Cytotoxicity

Jacob Goell[1], Jing Li[1], Barun Mahata[1], Alex J. Ma[1], Sunghwan Kim[1], Spencer Shah[1], Shriya Shah[1], Maria Contreras[1], Suchir Misra[2], Daniel Reed[1], Guy C. Bedford[1], Mario Escobar[1], and Isaac B. Hilton[1,2]

Affiliations

1. Department of Bioengineering, Rice University, Houston, TX 77030, USA
2. Department of Biosciences, Rice University, Houston, TX 77030, USA

Corresponding author:
Isaac Hilton
Rice University
6500 Main Street
Department of Bioengineering
Houston, TX 77030, USA
E-mail: isaac.hilton@rice.edu

ABSTRACT
Engineering histone acylation states can inform mechanistic epigenetics and catalyze therapeutic epigenome editing opportunities. Here, we developed engineered lysine acyltransferases that enable the programmable deposition of acetylation and longer-chain acylations. We show that targeting an engineered lysine crotonyltransferase results in weak levels of endogenous enhancer activation yet retains potency when targeted to promoters. We further identify a single mutation within the catalytic core of human p300 that preserves enzymatic activity while substantially reducing cytotoxicity, enabling improved viral delivery. We leveraged these capabilities to perform single-cell CRISPR activation screening and map enhancers to the genes they regulate *in situ*. We also discover acylation-specific interactions and find that recruitment of p300, regardless of catalytic activity, to prime editing sites can improve editing efficiency. These new programmable epigenome editing tools and insights expand our ability to understand the mechanistic role of lysine acylation in epigenetic and cellular processes and perform functional genomic screens.

KEYWORDS
Epigenome editing; histone acetylation; histone crotonylation; EP300; CRISPR; Perturb-seq; CRISPRa; enhancer; protein engineering; off-target editing

INTRODUCTION

Over the past decade, catalytically deactivated CRISPR-Cas (dCas) systems have emerged as powerful technologies to study epigenetic regulatory mechanisms in eukaryotic cells[1–3]. Epigenome editing platforms leverage the dCas protein as a scaffold to recruit epigenetic modifiers, or effectors, to targeted loci of interest. This capability has been instrumental for understanding the fundamental principles of heritable gene silencing[4,5], enhancer biology[6–9], and transcriptional activation/repression[10–14]. By site-specifically depositing specific epigenetic modification(s) at the locus of interest, epigenome editing enables interrogation of whether the addition or erasure of a specific modification is necessary or instructive for the transcription of endogenous loci. The programmable nature of this class of technologies makes them flexible and highly useful for elucidating causal relationships between epigenetic modifications and genomic activity. Despite the potential of such tools, a better understanding of the cytotoxicity and off-target profiles these effector domains harbor is necessary for greater adoption and future clinical applications. Epigenome editing effector domains with active enzymatic activity are especially susceptible to pervasive guide RNA (gRNA) independent off-target effects and may require substantial engineering to overcome these concerns[4,15–17].

Lysine acylation is a post-translational modification (PTM) that plays a key role in regulating chromatin structure and gene expression[18]. In particular, acylation of lysine residues on histone proteins has been shown to play a crucial role in determining chromatin accessibility and recruitment of epigenetic reader domains in eukaryotes[19,20]. Lysine acylation is dynamic and is regulated by a number of enzymes, allowing for the regulation of gene expression in response to various stimuli[18–22]. It has also been shown that many different lysine acetyltransferases harbor the ability to deposit longer-chain acylations onto their substrates in addition to acetylation[26]. One such enzyme is EP300 (referred to as p300 hereafter), and its paralog CBP, which have been shown to deposit propionyl, butyryl, crotonyl, lactyl, and other acylations onto histones[27]. These acylations are largely derived through metabolic processes such as fatty acid oxidation and the tricarboxylic acid (TCA) cycle, where they are generated as an adduct to coenzyme A and used as a substrate by lysine acyltransferases and deposited onto histones and other proteins[20]. As a result, these acylations are highly context-specific and are present in cell types wherein specific metabolic pathways are active and corresponding feedstocks are available. Longer-chain acylations are chemically distinct from acetylation. However, they are thought to serve similar functions and have been linked to active transcription despite exhibiting different epigenetic reader affinities[28–30]. Notwithstanding this fundamental importance, there are currently no tools permitting interrogation of the roles that different acylations play at endogenous loci in living cells.

Here, we develop and benchmark a set of engineered lysine acyltransferases using both rational and high throughput mutagenesis of the human p300 HAT core domain. We build a programmable lysine crotonyltransferase and demonstrate that, in contrast to histone acetylation, histone crotonylation only weakly activates genes from endogenous human enhancers. However, both acylation reactions can drive potent transcription when targeted to promoters. Using a deep mutational scanning approach, we also discover a single point mutation (I1417N) that drastically reduces cytotoxicity associated with the overexpression of p300 WT, without sacrificing the ability to deposit H3K27ac and activate transcription. Using quantitative mass spectrometry, we demonstrate that these individual point mutations reshape p300's interactome and we identify longer-chain acylation-specific interactions as well as cytotoxicity-associated interactions. Differential interactions in DNA repair proteins led us to discover that p300 co-targeting with various DNA editing enzymes can improve prime editing efficiency. At the transcriptomic and epigenomic levels, these engineered acyltransferases display decreased off-targeting, a consideration of paramount importance for the adoption of epigenome editing methodologies in translational applications. Finally, we leveraged our dCas9-p300 I1417N fusion protein to perform single-cell CRISPR activation and benchmarked its performance against dCas9-p300 WT, demonstrating that targeting this co-activator to non-coding regions can map enhancers to the

genes they regulate. Overall, we envision the engineering strategies and tools that we have developed will be highly useful for understanding the role of lysine acylation in transcriptional control and for improving epigenome editing tools beyond those applied for programmable histone acetylation.

RESULTS

Chromatin crotonylation strongly activates transcription from human promoters

We first aggregated publicly available data to determine the histone modifications and transcription factors that colocalize with H3 lysine 18 crotonylation (H3K18cr) in HCT116 cells[21,31]. We found that H3K18cr closely correlates with histone modifications associated with transcriptionally active loci (**Fig. S1A**). This finding is consistent with previous studies at both reporter genes and endogenous loci which demonstrated that exogenous addition of crotonate increases transcription[23,32]. To determine how longer-chain acylations affect transcription at endogenous regulatory elements, we constructed CRISPR-based fusion proteins with different variants of the p300 core (referred to as p300 hereafter) domain. We fused p300 harboring a previously characterized mutation (p300 I1395G) to the C-terminus of catalytically deactivated *Streptococcus pyogenes* Cas9 (dCas9; **Fig. 1A**). The I1395G mutation has been shown to skew p300/CBP toward depositing crotonylation onto chromatin instead of acetylation[33]. When these constructs were expressed in HEK293T cells, a marked increase of crotonylated lysine residues was observed across the proteome in comparison to the dCas9-p300 WT fusion or a catalytically inactivated p300 control (dCas9-p300 D1399Y[34]; **Fig. 1B**). In addition, we found that *in vitro* purified dCas9-p300 I1395G crotonylated purified histone H3 peptides to a greater degree than purified dCas9-p300 WT, with slightly decreased amounts of H3 peptide acetylation observed (**Fig. 1C**). These results demonstrate that the I1395G mutation can crotonylate lysine residues on histone and non-histone proteins to a greater degree than p300 WT. We also transfected a range of plasmid amounts and, interestingly, found that dCas9 fused to p300 WT was expressed at much lower levels than dCas9-p300 I1395G or dCas9-p300 D1399Y (**Fig. S1B**). Further, we observed that, when expressed at high enough levels, dCas9-p300 WT appeared to have detrimental effects on the relative expression of the tubulin loading control, which may indicate inhibited cell growth, toxicity, or cellular senescence upon p300 overexpression as has been suggested previously[35,36].

Upon targeting human promoters using a pool of guide RNAs (gRNAs), we observed that dCas9-p300 I1395G activated transcription at a level comparable to dCas9-p300 WT from promoters (**Fig. 1D**, **S1C**), which was also observed in another cell type (**Fig. S1D**). However, in contrast to dCas9-p300 WT, we found that dCas9-p300 I1395G only weakly activated the transcription of downstream genes when targeted to enhancers (**Fig. 1D, S1E**). Interestingly, most bromodomain-containing proteins are unable to bind crotonylated histone residues[28,37,38], which may explain the weakened enhancer-mediated gene activation displayed by dCas9-p300 I1395G. To determine if dCas9-p300 I1395G retained residual acetylation capabilities at targeted loci, we performed CUT&RUN-qPCR at the promoter of a targeted gene (*OCT4*). We observed a drastically reduced enrichment with dCas9-p300 I1395G to a level indistinguishable from dCas9 alone or the enzymatically deactivated dCas9-p300 D1399Y. However, as expected, we found enrichment of H3K18ac and H3K27ac when the dCas9-p300 WT fusion proteins were targeted to either the *OCT4* promoter or the *OCT4* distal enhancer (**Fig. 1E, 1F**). These data indicate that crotonylation and acetylation can be functionally divergent at endogenous human enhancers but can be redundant at human promoters with respect to transcriptional activation. Furthermore, our results demonstrate that dCas9-p300 I1395G is sufficient to activate genes from human promoters in the absence of H3K18ac and/or H3K27ac.

High throughput screening clarifies the p300 HAT domain activity-expression relationship

As an effector domain for epigenome editing, p300 (and its close paralog the CBP core)[9] is the only known direct writer of H3K27ac, and is also highly active when the core domain is isolated[39]. However, in many cell types, dCas9-p300 WT can be poorly expressed[9,40,41]. Given that the single I1395G point mutation altered both the acyltransferase activity and expression level of p300 (**Fig. 1B, S1B**), we hypothesized that other mutations within the catalytic domain of the enzyme could result in enhanced expression without sacrificing epigenome editing efficacy. To test this hypothesis, we developed an unbiased screening approach enabling us to characterize other mutations within the p300 core domain in high throughput. We used a reverse tetracycline repressor (rTetR) fusion-based approach[42] to recruit a library of single amino acid p300 mutants to a reporter cell line consisting of 9x TetO sites upstream of an SV40 core promoter driving a Citrine fluorescent reporter stably integrated into the AAVS1 locus (**Fig. 2A**).

We found that linking the rTetR to mCherry using a porcine teschovirus self-cleaving peptide (P2A) showed differing levels of expression across our previously characterized variants of p300, validating the rTetR system as a reporter on protein expression levels when either transduced and transiently transfected (**Fig. 2B**, **S2A**). Further, by directly fusing mCherry to the N-terminus of dCas9-p300-based fusions, we confirmed that this observation is consistent at the protein level, thus linking protein expression with stability in this context (**Fig. S2B**). This also corroborates our Western blot data demonstrating similar relative expression levels among dCas9 and dCas9-p300 variants (**Fig. S1B**). Collectively, these results suggest that down-regulation of exogenous p300 protein is correlated with higher levels of acetylation activity.

Having established that the rTetR fusion system enabled multiparametric screening of p300 variants on both transcriptional output and relative expression, we next designed a deep mutational scanning library consisting of 940 members encoding 800 point mutations, 40 single amino acid deletions, and 100 random negative controls. This library targeted the amino acids flanking a contiguous stretch of p300 (1381-1420aa) that contained the I1395G mutation and a known inactivating mutation (D1399Y)[43]. We also selected this region because it resides in the acyl-CoA binding pocket (~1330 - 1630aa)[34], which we hypothesized might enable reshaping of the enzymatic activity of p300 based on structural analysis (**Fig. 2C**). We synthesized and inserted this library into the p300 core domain and then delivered resulting variants at a multiplicity of infection (MOI) of ~0.3 via lentivirus into the HEK293T cells expressing our reporter system. After selection, we treated the resulting pool of transduced cells with doxycycline for 48 hours to induce recruitment of the library as described previously[42]. At day 9 post-transduction, we used a Sort-seq approach to bin our library into one of four quadrants demarcated by mCherry ($mCherry^{hi/lo}$) and Citrine ($Cit^{hi/lo}$) levels, sequenced the resulting domains, and used HT-recruit to compute the log_2(ON/OFF) ratios for each of the library members using the read counts of the sorted library populations within each of the four quadrants[42]. We consequently were able to assign ratios for 819 of the variants of p300.

As expected, mutations at the 1399 residue resulted in loss of transcriptional activation (**Fig. 2D, Supplementary Table 1**) which is consistent with the D1399Y catalytically deactivated p300 and the most prevalent mutations found in the Catalogue of Somatic Mutations in Cancer (CoSMIC) database (**Fig. S2D**)[44]. Mutations at this residue are also linked to the neurodevelopmental disorder, Rubinstein-Taybi Syndrome[45]. The residues known to encompass the acyl-CoA binding site (1398-1400 and 1410-1411) appeared to have a mixed effect when mutagenized in this system. For instance, only Asp1399 and Ser1400 showed disruptive effects on gene activation, which is reflected in these sites being the predominant residues where disease-causing variants are observed (**Fig. S2D**)[46]. Other residues known to contact acetyl-CoA appeared to be much more permissive of mutagenesis despite largely being hydrophobic. For example, Leu1398 and Ile1395 have been previously described to accommodate the methyl moiety of the acetyl group and were permissive of mutagenesis (**Fig. 2D**)[46]. Further, alteration at the Ile1395 imparted a greater propensity for crotonyl-CoA usage as described previously[27,33].

Therefore, site-specific mutagenesis at these two residues may enable further engineering of improved longer-chain acyltransferases.

Additionally, we found that deletions tended to disrupt transcriptional activation but had a mixed effect on protein expression. This result is concordant with the notion that loss of gene activation, and hence histone acetylation, can lead to greater expression at the protein level as was observed with p300 D1399Y (**Fig. 2D, 2E, Supplementary Table 1**). Despite the loss of transcriptional activation observed when mutagenizing the acyl-CoA binding site, we found that these mutations had minimal effect on protein expression, suggesting that these mutations in p300 drive catalytic inactivity as opposed to changes in structure and/or misfolding (**Fig. 2E**). Interestingly, some frequently mutated residues found in the CoSMIC database (i.e., Cys1385 and Tyr1414; **Fig. S2D**) did not appear to impact the transcriptional activity of p300, but instead, increased the relative expression of p300, suggesting that p300 may drive oncogenesis through loss of function and/or increased expression.

Due to the demonstrable difference the catalytically deactivated p300 displayed with respect to expression levels compared to the active enzyme, we tested if this observation held for other types of epigenome editing enzymes. Specifically, we integrated the catalytically active or inactive domains of p300, PRDM9, RING1B, and DNMT3A3L into a stability reporter system (**Fig. 2F**)[47,48]. We found that both activators (p300 and PRDM9) had reduced expression levels when catalytically intact, whereas the expression of repressors that we tested (RING1B and DNMT3A3L) was unaffected by catalytic activity. Loss of expression occurred primarily at the mRNA level (**Fig. 2F, S2E**), consistent with our data comparing both direct fusion and cistronic expression of p300 (**Fig. S2A-C**). This suggests that this mechanism of downregulation at the level of mRNA may extend to epigenome-modifying activators as a class and may be a result of transcriptional squelching or another mode of cellular toxicity driven by redistribution of stoichiometrically-controlled co-activators[49–51]. Taken together with our screening regime, this data further demonstrates the utility of multiparametric tuning of expression and transcriptional activation as a powerful approach to modulate the expression and activity of these epigenome editing tools.

Characterization of a stable and active variant of the p300 core domain

From our deep mutational scanning of the 1381-1420aa position within the p300 core domain, we identified two regions on the activity vs. stability landscape on which we performed validation studies (**Fig. 3A**). Specifically, we analyzed the top six scoring variants across a combination of both protein expression and activity (i.e., mChe^{hi}Cit^{hi}). Given that p300 I1395G worked as a stronger activator at the *OCT4* and 9x TetO promoters than p300 WT (**Fig. 1D, S2A**), we also performed follow-up studies on five variants with high protein expression but an intermediate level of gene activation (mChe^{hi}Cit^{mid}; **Fig. 3A**). We found that 5 out of 6 of the mChe^{hi}Cit^{hi} variants and 2 out of 5 of the mChe^{hi}Cit^{mid} variants had expression levels that exceeded p300 WT (**Fig. 3B**). We next fused each of these variants to dCas9 to assess relative performance at endogenous loci and found that four of the engineered variants (Q1390R, K1407Y, L1409K, and I1417N) performed comparably to p300 WT across native human promoters and enhancers (**Fig. 3C**). Three of these variants (Q1390R, K1407Y, I1417N) were derived from the mChe^{hi}Cit^{mid} region, supporting our rationale for enrichment of variants that activate genes similarly to p300 WT but had improved expression. Integrating these results together, we identified the I1417N mutation as a promising lead to characterize further given that its expression level was similar to the catalytically deactivated D1399Y enzyme (**Fig. 3B**) and its ability to activate enhancers and promoters was similar to p300 WT.

We validated that p300 I1417N activated transcription comparably to slightly less than the WT enzyme in a locus-dependent manner by targeting the regulatory elements controlling *OCT4* expression as a testbed (**Fig. 3D**, **S3A**). Further, we confirmed that p300 I1417N displayed similar efficacy to p300 WT in other cell types (**Fig. S3B-D**) and across different Cas species (ddCas12a

and dCasMINI; **Fig. S3E, S3F**)[52,53]. We also confirmed that protein expression correlated to the mCherry output of the polycistronic P2A system via Western blot, validating that dCas9-p300 I1417N is expressed much more highly than dCas9-p300 WT in HEK293T cells (**Fig. 3E, 3F**). To determine whether this improved expression level also resulted in improved cell viability upon transient transfection, we performed a cell viability assay across the different variants of p300. Strikingly, cells transfected with p300 WT displayed reduced viability compared to the p300 I1417N enzyme whereas the viability of cells expressing the engineered p300 I1417N variant was comparable to that of dCas9 alone (**Fig. 3G**).

Constraints on stability and cell viability are potential bottlenecks in packaging virus for downstream cell engineering[54] and we hypothesized that the improvement in both protein stability and cell viability would enable improved packaging of dCas9-p300 I1417N relative to dCas9-p300 WT in viral-based systems. To test this hypothesis, we packaged dCas9, dCas9-p300 WT, and dCas9-p300 I1417N in lentiviral vectors and quantified their physical titers. We observed that dCas9-p300 I1417N had a 2-fold greater titer in these packaging experiments, thus enabling more efficient p300-based epigenome editor delivery via lentivirus (**Fig. S3G**). We also quantified lentiviral transduction efficiency in K562s transduced at a defined quantity (2%v/v) and observed a greater than 5-fold improvement in transduction efficiency as measured by a co-translated mCherry fluorophore (**Fig. 3H**).

We next created a similar mutation in CBP (I1453N) as the HAT core domains are highly conserved between the two proteins[40,43]. Upon targeting CBP WT and I1453N variants to the *HBG1* and *HS2* loci, we discovered that CBP I1453N lacks the ability to activate transcription at both promoters and enhancers (**Fig. 3I**). While this was surprising as p300 and CBP are structurally similar, and are thought to be functionally redundant in many cases, previous studies have found identical mutations in CBP/p300 that lead to only minor reductions in acetylation activity for one and abolished activity in the other, pointing to subtle structural differences between the two proteins not obvious at the level of conserved amino acid sequences (**Fig. S3H**) .

Discovery of protein-protein interactions driven by different variants of the p300 core domain

Due to the pronounced differences that p300 variants displayed relative to p300 WT in both transcriptional activation and expression levels, we suspected that these mutations were altering protein-protein interactions. To investigate this possibility, we performed immunoprecipitation followed by mass spectrometry (IP-MS) on HEK293T cells transiently transfected with selected mutants of dCas9-p300 and dCas9-p300 WT (**Fig. S4A**). We filtered out hits that were highly prevalent in the CRAPome (>50% occurrence) to remove false positive interactors that are common in IP-MS experiments[56]. We found that dCas9-p300 WT exhibited the most significant differential interactors (p_{val}<0.05; compared to a dCas9-only “bait”) with 236 hits, whereas the catalytically inactivated dCas9-p300 D1399Y mutant exhibited the fewest with 115 significant hits (p_{val}<0.05; compared to a dCas9-only bait, **Fig. 4A**). The crotonylation depositing dCas9-p300 I1395G variant and the improved stability dCas9-p300 I1417N variant had comparable interaction hits with 169 and 163, respectively. When compared against dCas9 as a negative control bait sample, and hierarchically clustered, patterns emerged that indicated which interactions arose as a function of the different activities of these p300 variants (**Fig. 4B**). For example, we observed that scaffolding-dependent interaction partners such as RNASEH2A, PTPN11, CHMP5, PREB, and CDK4 were conserved across all four variants of p300. Another hierarchical cluster contained interactors that were specific to certain acyl modifications such as MINK1 and CDK13. This cluster also contained interactions that required the acylation activity of p300 such as YTHDF1, BAG2, and SIRT2, the last of which is a histone deacetylase and known negative regulator of p300, demonstrating the close interplay between proteins that counterbalance epigenetic mechanisms, particularly histone acetylation/deacetylation cycling[57]. Interestingly, many hits in the dCas9-p300 WT condition were unique (159/458, 34.7%) and, of

these unique hits, were significantly depleted compared to the dCas9 only control (141/159, 88.6%; **Supplementary Table 2**). This suggests that the acetylation activity of p300 drives a greater subset of p300 interactors as opposed to its scaffolding function and may even repel potential interacting partners, as shown through the depletion of significant interactors.

Within our proteomics dataset, we found several known p300 interactions that are noted in the STRING database such as TAF6, WDR5, SUPT16H, and SIRT2 (**Fig. S4C**). Interestingly, TAF6, WDR5, and SUPT16H are depleted hits in our dataset, highlighting the function of the flanking domains of the p300 core that serve scaffolding functions and stabilize interactions with transcription factors and co-activators. As expected, many of the known hits are positive regulators of transcription and are only dCas9-p300 WT interactors, whereas SIRT2 interacts with all variants with acyltransferase activity (**Fig. S4C**). We also compared our proteomics dataset against another dataset of known acetylation targets of p300[58]. Despite retaining the capacity to acetylate histones, dCas9-p300 I1417N shared only 45% and 60% of significant differential interacting partners with dCas9-p300 WT and dCas9-p300 D1399Y, respectively, which suggests that dCas9-p300 I1417N may differ in activity from WT due to its lack of non-histone protein acetylation (**Fig. S4D**).

To further elucidate the contrasts between dCas9-p300 WT and dCas9-p300 I1417N, we computed the fold change differences in interactions between these two variants. The stable I1417N variant interacted with a greater number of proteins at a significant level reinforcing the notion that WT-level acetylation leads to a repulsion of interactor proteins (**Fig. 4C**). Interestingly, some of these proteins such as DAXX, ASF1A, and UBE2L3 are involved in apoptosis and protein degradation pathways[59–61]. We performed gene ontology analysis and discovered that many of the dCas9-p300 I1417N hits are implicated in transcription-coupled nucleotide-excision repair and other DNA repair pathways (**Fig. 4D**), which is intriguing as DNA damage response proteins have been observed to be acetylated by p300, and moreover, knockdown of p300 was recently shown to decrease prime editing efficiency[62,63].

These findings led us to evaluate how p300 co-targeting altered prime editing efficiency. We first co-targeted dCas9-p300 with the PEmax prime editor to an integrated GFP reporter in HEK293T cells and observed improved conversion to BFP with p300 WT, D1399Y, and I1417N co-targeting to the editing site (**Fig. 4E**). In addition, we found that the recently described transcriptional activator, NMS, also led to small increases in editing efficiency. Importantly, these trends in improved prime editing efficiency were also observed at other endogenous genes. Specifically, we found consistent improvements in prime editing efficiency at B2M, ATP1A1, and HBB for all variants of p300 co-targeted. NMS also showed consistent prime editing improvements across all sites it was targeted to (**Fig. 4E**). This is concordant with previous data demonstrating that transactivation domain co-targeting to editing sites improves prime editing outcomes[63,64]. However, our data indicates that catalytically inactive p300 can also improve editing efficiency, suggesting that transient gene activation is not necessary for these improved outcomes and points to a non-catalytic/scaffolding function of p300 that may enhancer prime editing at target sites. Further, we observed similar improvements in base editing experiments, wherein p300 D1399Y and I1417N increased base editing efficiencies (**Fig. S4E**). Co-targeting experiments using CRISPR-mediated integration of an HDR donor template led to reduced integration of a Citrine-containing donor template across all p300 variants, with p300 WT having the lowest editing efficiency (**Fig. S4F**). Altogether, these data suggest that the specific function of p300 can influence DNA editing efficiencies in contrasting ways dependent on the DNA repair pathway and/or deployed genome editing modality.

Off-target characterization of p300-mediated epigenome editing

We next performed RNA-seq to quantify how the exogenous expression of different p300 variants globally affected the human transcriptome. Overall, we found that transcriptional profiles were fairly consistent among tested p300 variants, with minimal off-targeting compared to dCas9.

As expected, we observed a pattern wherein dCas-p300 WT displayed the most off-targeting (R^2=0.9935) followed by dCas9-p300 I1417N (R^2=0.9947) and then dCas9-p300 D1399Y (R^2=0.9953) when compared against cells expressing dCas9 alone (**Fig. 5A, Supplementary Table 3**). Most of the observed differentially expressed genes occurred at lowly expressed loci (**Figs. S5A**). Interestingly, these data are consistent with a model in which the enzymatic activity of high levels of p300 can lead to aberrant H3K27ac enrichment at enhancers that is consequently responsible for relatively small (or undetectable) changes in gene expression[65,66].

To better understand both the on- and off-target profiles of dCas9-p300 WT, I1417N, and D1399Y variants at the epigenomic level, we used CUT&RUN to probe for H3K27ac in HEK293T cells in which these dCas9-based fusion proteins were targeted to the *HBG1* promoter. Expectedly, we observed high levels of on-target H3K27ac deposition via dCas9-p300 WT and I1417N but not the catalytically-inactive D1399Y (**Fig. 5B, S5B**). We next compared enrichment of H3K27ac across all transcription start sites (TSSs) in our genome-wide CUT&RUN datasets to characterize the off-target profile of these respective epigenome editors. Counterintuitively, we found that both p300 WT and p300 I1417N lead to substantially reduced global H3K27ac levels at TSSs compared to p300 D1399Y (**Fig. 5C**), which we also observed upstream of TSSs and into gene bodies (**Fig. S5C**). Genome-wide, we observed concordance with our transcriptomic profiles in which p300 I1417N exhibited an epigenomic profile most similar to both a dCas9 and p300 D1399Y-only negative control in comparison to p300 WT, which showed greatly reduced enrichment of H3K27ac across the genome (**Fig. S5D**). This relative reduction in global H3K27ac levels may be due to acetylation of other histone residues and/or transcription factors resulting in shunting of the available acetyl-CoA pool away from H3K27ac genome-wide[25,67,68].

Additionally, we performed CUT&RUN-qPCR across several other histone modifications to assess the epigenomic status at the *HBG1* target locus. We found H2BK20ac, a p300/CBP-specific histone modification that is thought to designate active enhancers[69], was deposited at similar levels by both active enzymes (**Fig. 5D**). We also probed for H3K4me3, as this modification has been previously found to have substantial crosstalk with H3K27ac[13,70,71] and indeed observed high levels of co-enrichment among these modifications (**Fig. 5E**). Lastly, we quantified H3K27me3 as a negative control and found no significant deposition of this heterochromatin-associated histone modification at targeted sites (**Fig. 5F**). These data support the notion that p300 WT and I1417N are highly similar in terms of their capacity to acetylate histones and activate transcription while relatively divergent with respect to protein expression and effects on cell viability. Our collective data indicate that dCas9-p300 I1417N has a reduced transcriptomic off-targeting profile yet retains comparable enzymatic activity in human cells in comparison to dCas9-p300 WT. Further, our data indicates that widespread H3K27ac perturbation does not necessarily lead to dramatic alterations to the transcriptome, which indicates that H3K27ac can be dispensable for enhancer activity and may not be strictly necessary for transcriptional activity at promoters in all cases[66,72,73].

Benchmarking of p300-based Perturb-seq

We next sought to assess both p300 WT and p300 I1417N as a tool for functional genomics profiling[8,74,75]. We generated monoclonal K562 cell lines expressing either dCas9-p300 WT or I1417N and validated transgene expression levels and transactivation efficacy (**Fig. S6A and S6B**). We generated a gRNA library to target candidate promoters and enhancers sourced from a previous study demonstrating a single-cell RNA-seq-based CRISPRa proof of concept mapping enhancers to promoters[75]. We used a similar experimental framework to integrate gRNA libraries at a high MOI, performing a vector copy number titration to ensure delivery of MOI > 10 library members per cell (**Fig. S6C**). By coupling scRNA-seq with gRNA capture, cells containing subsets of the library can be partitioned into a set containing a specific gRNA and the complement set, thus linking the transcriptional profile of that group as previously described[65,75,76] (**Fig. 6A)**.

K562 cells were transduced at 1% v/v lentivirus and selected for puromycin resistance for 10 days before harvesting for scRNA-seq and gRNA sequence capture. After quality control, we captured 9,660 and 11,073 single-cell transcriptomes with a median of 19 and 20 guides captured per cell and 392 and 484 cells assigned per gRNA for the p300 WT and p300 I1417N expressing lines, respectively (**Fig. S6C and S6D**). To perform differential expression testing, we used a conditional resampling approach (SCEPTRE) that was previously developed for improved calibration on single-cell CRISPR datasets to link perturbations with gene expression changes[77]. Using the SCEPTRE pipeline, we grouped together gRNAs targeting each individual genomic element and performed pairwise tests against each gene located within 1Mb of each gRNA location for gene expression perturbation. We observed strong calibration for our controls with non-targeting gRNAs having no effect (**Fig. S6E**).

Of the 434 human genome-targeting gRNAs used, 88 unique elements were targeted between promoters (38 unique elements) and enhancers (50 unique elements) previously identified and validated using CRISPRi screens[65]. Across all hits, the mean activation was similar between p300 WT and p300 I1417N (**Fig. 6B**). We identified 48 and 24 gRNA group-gene response hits for p300 WT and p300 I1417N, respectively, with 17/24 hits found in the p300 I1417N dataset shared by that of p300 WT (**Fig. 6C and 6D, S6F**). While p300 WT had twice the number of statistically significant hits using the SCEPTRE framework, the average activation from called hits was comparable. Indeed, at individual promoters such as *BIK*, *GMPR*, and *GNB2* we observed comparable gene activation between p300 WT and p300 I1417N (**Fig. 6E**). Similar trends are observed at enhancers regulating *LINC01033* (chr5.1314), *THUMPD2* (chr2.1584), and *RNPEP* (chr1.10703).

Because p300 has direct catalytic activity and therefore can function through different mechanisms than other transactivation domains, we also compared our results to a previously performed scCRISPRa study that used VP64 or VPR[75]. We found several promoters that were responsive to p300 WT-mediated upregulation but that were not activated by the transactivation domains VP64 or VPR (e.g. *ANO5*, *IRS1*, *SCN2A*; **Fig. S6G, Table S4**). Conversely, only one promoter-targeting hit, *FOXP1*, was found in the transactivation domain dataset but not identified as a hit in the p300 WT dataset. We also found enhancers that activate downstream genes by p300 WT but not by VP64 or VPR (e.g. chr2.1584-*THUMPD2*, chr5.1314-*LINC01033*, chr6.3216-*SLC35B2;* **Fig. S6H, Table S4**). Interestingly, only one enhancer was found that was activated by VP64/VPR but not p300 (chr1.7358-*TMEM56*). Altogether, we found 19 enhancer-promoter pairs by targeting p300 WT whereas previous studies using transactivation domains only uncovered 8 pairs. Of these 8 previously identified pairs, 7 were shared by p300 WT. Hence p300 appears to have a unique strength in terms activation of genes from enhancers relative to VP64 or VPR (**Fig. S6H**).

DISCUSSION

Here we comprehensively studied the ability of the human p300 core domain to acylate chromatin *in situ*, developing approaches to manipulate its function and reduce its toxicity for epigenome editing. Through single point mutations in the enzyme's binding pocket, we benchmarked a variant of p300 tailored towards crotonyltransferase activity and established that this variant can activate promoters comparably to an acetylation-competent version of the p300 core. Conversely, our data suggest that histone crotonylation may not be sufficient to activate genes from enhancers. This is in line with recent reports suggesting that YEATS-domain containing GAS41 can bind to H3K27cr and repress transcription[78]. While we did observe low levels of transcriptional activation at enhancers marked with engineered crotonylation as opposed to no activation, this may be due to low residual acetylation activity of p300 I1395G or competition by transcription-activating YEATS domain-containing proteins such as ENL or AF9[79]. Given the significant role that BRD4 and other bromodomain-containing proteins have in enhancer activity coupled with the inability of the bromodomain family to bind crotonylated residues, we hypothesize

that histone crotonylation may drive weaker enhancer-mediated transcriptional regulation in tissues where crotonylation is more prevalent[19,21].

As an effector domain for epigenome editing, the p300 core and its close paralog CBP are the only known writers of H3K27ac, a modification that demarcates active promoters and enhancers[40]. The ability to write this PTM is thus critical for understanding its role in enhancer activation and for mapping enhancers to the genes they regulate[80]. Given its promiscuity to acylate other targets beyond H3K27ac in an "acetyl spray" mechanism, p300 has been associated with widespread off-targeting in previous studies[58,81]. To address this issue, we developed an unbiased screening approach to characterize other mutations within the p300 core domain in a high throughput manner and, in turn, identify high expressing p300 variants that retain the ability to activate genes from both promoters and enhancers. We were able to successfully nominate and characterize the I1417N mutation, which we validated as a novel mutation that both improves the stability of p300 while also retaining the molecule's ability to potently stimulate gene expression.

In addition, we were able to link improvements in p300 I1417N protein expression to a reduction in genome-wide off-targeting and downstream transcriptomic changes. Interestingly, the I1417N point mutation also enabled higher efficiency packaging into lentivirus and transduction. This suggests that systematic epigenetic dysregulation may be driving minor transcriptomic alterations as well as functional changes in cells in which epigenome editors are expressed. Thus, the approach delineated in this study may prove useful for optimization of other genome or epigenome editor effector domains. Indeed, we observed the same activity versus expression tradeoff across other epigenome editing effector domains associated with gene upregulation.

The protein-protein interaction (PPI) profiling we performed also provides an important resource in understanding the interactome of the p300 core domain and how its specific enzymatic activity regulates its engagement with other proteins. By performing IP-MS across p300 core variants with differing catalytic activities, we were able to dissect which interactions were driven by the scaffolding function of p300 compared to acetylation and crotonylation functions. It is important to note that here we used the p300 core domain to interrogate direct PPIs associated with the catalytic activity of p300 variants. This is opposed to PPIs associated with the extensive intrinsically disordered domain-rich scaffolding that forms the flanking regions of the core domain, which are known to interact with transcription factors and other chromatin regulators[40]. By isolating the core domain, this dataset is a more accurate representation of histone acetylation interactors as opposed to the complex self-regulatory and hub-like function that the full-length protein plays. Our finding that p300 co-targeted to prime editing sites improves editing efficiency is concordant with previous findings demonstrating similar phenomena co-targeting transactivation domains to these sites[63,64]. This points to a transcriptional activity-dependent or chromatin accessibility-mediated mechanism. However, our data suggests that this may also be due to non-catalytic/scaffolding functions of p300, as a catalytically inactive p300 variant also improves prime editing efficiency to a degree similar to p300 WT and the transcriptional activator, NMS. Interestingly, many of these transactivation domains interact with endogenous p300, raising the intriguing possiblity that recruitment of these domains may serve an intermediary role in bringing p300 and its interactors to these sites.

Recently, several studies have utilized CRISPRi and CRISPRa to perturb the function of the non-coding genome and thereby link enhancers to the genes that they regulate. Largely, these studies have been conducted using dCas9-KRAB and derivatives thereof[65,76], which provide a lens towards active enhancers and their cognate genes. However, the toolbox to perturb latent or silent enhancers remains underdeveloped, with only one other study performed to date developing CRISPRa at non-coding regions for Perturb-seq[75]. We extend this toolbox to include both p300 WT and p300 I1417N. While our engineered p300 I1417N only encompasses ~50% of the hits of p300 WT using an existing Perturb-seq statistical framework, its improved transduction

efficiency, decreased off-target effects, and reduced toxicity provide benefits when performing functional genomics studies on non-cancerous and primary cells.

CRISPR-based technologies that are reliant upon the recruitment of other effector domains such as base editing, prime editing, gene writing, and epigenome editing, must contend with CRISPR-independent off-target effects. Our data suggests that certain domains, particularly enzymatic modifiers, may impose toxicity which must be addressed prior to translation into clinically relevant cells and *in vivo* usage. Therefore, comprehensive studies such as those performed here using the p300 core domain are necessary to investigate toxicities that occur beyond the established yet evolving practices of studying DNA disruption. Our study validates a systematic engineering approach that can be applied to other epigenome editing effector domains to overcome toxicities while retaining efficacy and suggests that other gene-activation linked domains may require, and be amenable to, similar optimization.

Limitations of the study

While we were able to directly detect enriched crotonylation using *in vitro* experiments, we were unable to reliably detect lysine crotonylation likely due to the lack of high-specificity antibodies. Therefore, other approaches may be necessary to quantify endogenous histone crotonylation site-specifically. Additionally, all gRNAs used in this study, besides those used for our Perturb-seq screening, were previously selected for p300 WT, and mutation of the catalytic pocket of p300 or biasing towards histone crotonylation may have altered optimal gRNA activity. In our mutagenesis screen, we selected only a fragment of the p300 core domain, and it is likely that other regions harbor beneficial mutations for improved expression (i.e. within the bromodomain). Furthermore, while our protein-protein interaction study nominated differential hits for varying p300 activities, we have not validated whether, and if so to what extent, these divergent interactomes drive different functional effects.

The scCRISPRa experiment gRNA library was selected based on a validated single-cell gRNA library[75]. gRNAs chosen in this library were largely derived from previous studies using CRISPRi and may not reflect optimal targeting characteristics of p300 or transcriptional activators generally. Moreover, the use of different assay conditions and processing pipelines used between the two studies may contribute to some of the differences observed between the hits found in the two datasets. Finally, an additional bias away from the discovery of inducible enhancers also likely occurred through the selection of gRNAs targeting regions with biochemical markers predictive of enhancers (i.e. DNAse hypersensitivity, H3K27ac) which is indicative of existing p300 activity at these sites.

REFERENCES

1. Goell, J. H. & Hilton, I. B. CRISPR/Cas-Based Epigenome Editing: Advances, Applications, and Clinical Utility. *Trends in Biotechnology* **39**, 678–691 (2021).
2. Thakore, P. I., Black, J. B., Hilton, I. B. & Gersbach, C. A. Editing the epigenome: technologies for programmable transcription and epigenetic modulation. *Nat Methods* **13**, 127–137 (2016).
3. Nakamura, M., Gao, Y., Dominguez, A. A. & Qi, L. S. CRISPR technologies for precise epigenome editing. *Nat Cell Biol* **23**, 11–22 (2021).
4. Nuñez, J. K. *et al.* Genome-wide programmable transcriptional memory by CRISPR-based epigenome editing. *Cell* **184**, 2503-2519.e17 (2021).
5. Cappelluti, M. A. *et al.* Durable and efficient gene silencing in vivo by hit-and-run epigenome editing. *Nature* 1–8 (2024) doi:10.1038/s41586-024-07087-8.
6. Li, K. *et al.* Interrogation of enhancer function by enhancer-targeting CRISPR epigenetic editing. *Nat Commun* **11**, 485 (2020).
7. Klann, T. S. *et al.* CRISPR–Cas9 epigenome editing enables high-throughput screening for functional regulatory elements in the human genome. *Nat Biotechnol* **35**, 561–568 (2017).

8. Fulco, C. P. *et al.* Systematic mapping of functional enhancer-promoter connections with CRISPR interference. *Science* **354**, 769–773 (2016).
9. Wang, K. *et al.* Systematic comparison of CRISPR-based transcriptional activators uncovers gene-regulatory features of enhancer–promoter interactions. *Nucleic Acids Research* **50**, 7842–7855 (2022).
10. O'Geen, H. *et al.* dCas9-based epigenome editing suggests acquisition of histone methylation is not sufficient for target gene repression. *Nucleic Acids Research* **45**, 9901–9916 (2017).
11. Matharu, N. *et al.* CRISPR-mediated activation of a promoter or enhancer rescues obesity caused by haploinsufficiency. *Science* **363**, eaau0629 (2019).
12. Li, J. *et al.* Programmable human histone phosphorylation and gene activation using a CRISPR/Cas9-based chromatin kinase. *Nat Commun* **12**, 896 (2021).
13. Mahata, B. *et al.* Compact engineered human mechanosensitive transactivation modules enable potent and versatile synthetic transcriptional control. *Nat Methods* 1–13 (2023) doi:10.1038/s41592-023-02036-1.
14. Escobar, M. *et al.* Quantification of Genome Editing and Transcriptional Control Capabilities Reveals Hierarchies among Diverse CRISPR/Cas Systems in Human Cells. *ACS Synth. Biol.* **11**, 3239–3250 (2022).
15. Hofacker, D. *et al.* Engineering of Effector Domains for Targeted DNA Methylation with Reduced Off-Target Effects. *International Journal of Molecular Sciences* **21**, 502 (2020).
16. Pflueger, C. *et al.* A modular dCas9-SunTag DNMT3A epigenome editing system overcomes pervasive off-target activity of direct fusion dCas9-DNMT3A constructs. *Genome Res.* **28**, 1193–1206 (2018).
17. Gemberling, M. P. *et al.* Transgenic mice for in vivo epigenome editing with CRISPR-based systems. *Nat Methods* **18**, 965–974 (2021).
18. Ali, I., Conrad, R. J., Verdin, E. & Ott, M. Lysine Acetylation Goes Global: From Epigenetics to Metabolism and Therapeutics. *Chem. Rev.* **118**, 1216–1252 (2018).
19. Nitsch, S., Zorro Shahidian, L. & Schneider, R. Histone acylations and chromatin dynamics: concepts, challenges, and links to metabolism. *EMBO reports* **22**, e52774 (2021).
20. Dai, S.-K. *et al.* Histone crotonylation regulates neural stem cell fate decisions by activating bivalent promoters. *EMBO reports* **22**, e52023 (2021).
21. Fellows, R. *et al.* Microbiota derived short chain fatty acids promote histone crotonylation in the colon through histone deacetylases. *Nat Commun* **9**, 105 (2018).
22. Trefely, S., Lovell, C. D., Snyder, N. W. & Wellen, K. E. Compartmentalised acyl-CoA metabolism and roles in chromatin regulation. *Molecular Metabolism* **38**, 100941 (2020).
23. Sabari, B. R. *et al.* Intracellular crotonyl-CoA stimulates transcription through p300-catalyzed histone crotonylation. *Mol Cell* **58**, 203–215 (2015).
24. Sheikh, B. N. & Akhtar, A. The many lives of KATs — detectors, integrators and modulators of the cellular environment. *Nat Rev Genet* **20**, 7–23 (2019).
25. Shvedunova, M. & Akhtar, A. Modulation of cellular processes by histone and non-histone protein acetylation. *Nat Rev Mol Cell Biol* **23**, 329–349 (2022).
26. Tan, M. *et al.* Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell* **146**, 1016–1028 (2011).
27. Kaczmarska, Z. *et al.* Structure of p300 in complex with acyl-CoA variants. *Nat Chem Biol* **13**, 21–29 (2017).
28. Flynn, E. M. *et al.* A Subset of Human Bromodomains Recognizes Butyryllysine and Crotonyllysine Histone Peptide Modifications. *Structure* **23**, 1801–1814 (2015).
29. Filippakopoulos, P. *et al.* Histone Recognition and Large-Scale Structural Analysis of the Human Bromodomain Family. *Cell* **149**, 214–231 (2012).
30. Li, Y. *et al.* Molecular Coupling of Histone Crotonylation and Active Transcription by AF9 YEATS Domain. *Molecular Cell* **62**, 181–193 (2016).

31. Dunham, I. *et al.* An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).
32. Goudarzi, A. *et al.* Dynamic Competing Histone H4 K5K8 Acetylation and Butyrylation Are Hallmarks of Highly Active Gene Promoters. *Molecular Cell* **62**, 169–180 (2016).
33. Liu, X. *et al.* MOF as an evolutionarily conserved histone crotonyltransferase and transcriptional activation by histone acetyltransferase-deficient and crotonyltransferase-competent CBP/p300. *Cell Discov* **3**, 1–17 (2017).
34. Ortega, E. *et al.* Transcription factor dimerization activates the p300 acetyltransferase. *Nature* **562**, 538–544 (2018).
35. Olzscha, H., Fedorov, O., Kessler, B. M., Knapp, S. & La Thangue, N. B. CBP/p300 Bromodomains Regulate Amyloid-like Protein Aggregation upon Aberrant Lysine Acetylation. *Cell Chemical Biology* **24**, 9–23 (2017).
36. Sen, P. *et al.* Histone Acetyltransferase p300 Induces De Novo Super-Enhancers to Drive Cellular Senescence. *Mol Cell* **73**, 684-698.e8 (2019).
37. Hnisz, D. *et al.* Super-Enhancers in the Control of Cell Identity and Disease. *Cell* **155**, 934–947 (2013).
38. Sabari, B. R. *et al.* Coactivator condensation at super-enhancers links phase separation and gene control. *Science* **361**, eaar3958 (2018).
39. Hilton, I. B. *et al.* Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat Biotechnol* **33**, 510–517 (2015).
40. Dancy, B. M. & Cole, P. A. Protein Lysine Acetylation by p300/CBP. *Chem. Rev.* **115**, 2419–2452 (2015).
41. Tycko, J. *et al.* Development of compact transcriptional effectors using high-throughput measurements in diverse contexts. 2023.05.12.540558 Preprint at https://doi.org/10.1101/2023.05.12.540558 (2023).
42. Tycko, J. *et al.* High-Throughput Discovery and Characterization of Human Transcriptional Effectors. *Cell* **183**, 2020-2035.e16 (2020).
43. Delvecchio, M., Gaucher, J., Aguilar-Gurrieri, C., Ortega, E. & Panne, D. Structure of the p300 catalytic core and implications for chromatin targeting and HAT regulation. *Nat Struct Mol Biol* **20**, 1040–1046 (2013).
44. Tate, J. G. *et al.* COSMIC: the Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Research* **47**, D941–D947 (2019).
45. Zimmermann, N., Acosta, A. M. B. F., Kohlhase, J. & Bartsch, O. Confirmation of EP300 gene mutations as a rare cause of Rubinstein–Taybi syndrome. *Eur J Hum Genet* **15**, 837–842 (2007).
46. Maksimoska, J., Segura-Peña, D., Cole, P. A. & Marmorstein, R. Structure of the p300 Histone Acetyltransferase Bound to Acetyl-Coenzyme A and Its Analogues. *Biochemistry* **53**, 3415–3422 (2014).
47. Słabicki, M. *et al.* The CDK inhibitor CR8 acts as a molecular glue degrader that depletes cyclin K. *Nature* **585**, 293–297 (2020).
48. Policarpi, C., Munafò, M., Tsagkris, S., Carlini, V. & Hackett, J. A. Systematic Epigenome Editing Captures the Context-dependent Instructive Function of Chromatin Modifications. 2022.09.04.506519 Preprint at https://doi.org/10.1101/2022.09.04.506519 (2022).
49. Lin, H., McGrath, J., Wang, P. & Lee, T. Cellular Toxicity Induced by SRF-Mediated Transcriptional Squelching. *Toxicological Sciences* **96**, 83–91 (2007).
50. Gillespie, M. A. *et al.* Absolute Quantification of Transcription Factors Reveals Principles of Gene Regulation in Erythropoiesis. *Molecular Cell* **78**, 960-974.e11 (2020).
51. Jones, R. D. *et al.* An endoribonuclease-based feedforward controller for decoupling resource-limited genetic modules in mammalian cells. *Nat Commun* **11**, 5690 (2020).

52. Campa, C. C., Weisbach, N. R., Santinha, A. J., Incarnato, D. & Platt, R. J. Multiplexed genome engineering by Cas12a and CRISPR arrays encoded on single transcripts. *Nat Methods* **16**, 887–893 (2019).
53. Xu, X. *et al.* Engineered miniature CRISPR-Cas system for mammalian genome regulation and editing. *Molecular Cell* **81**, 4333-4345.e4 (2021).
54. Maunder, H. E. *et al.* Enhancing titres of therapeutic viral vectors using the transgene repression in vector production (TRiP) system. *Nat Commun* **8**, 14834 (2017).
55. Bordoli, L. *et al.* Functional analysis of the p300 acetyltransferase domain: the PHD finger of p300 but not of CBP is dispensable for enzymatic activity. *Nucleic Acids Research* **29**, 4462–4471 (2001).
56. Mellacheruvu, D. *et al.* The CRAPome: a contaminant repository for affinity purification–mass spectrometry data. *Nat Methods* **10**, 730–736 (2013).
57. Black, J. C., Mosley, A., Kitada, T., Washburn, M. & Carey, M. The SIRT2 Deacetylase Regulates Autoacetylation of p300. *Molecular Cell* **32**, 449–455 (2008).
58. Weinert, B. T. *et al.* Time-Resolved Analysis Reveals Rapid Dynamics and Broad Scope of the CBP/p300 Acetylome. *Cell* **174**, 231-244.e12 (2018).
59. Huang, L. *et al.* DAXX represents a new type of protein-folding enabler. *Nature* **597**, 132–137 (2021).
60. Eldridge, M. J. G., Sanchez-Garrido, J., Hoben, G. F., Goddard, P. J. & Shenoy, A. R. The Atypical Ubiquitin E2 Conjugase UBE2L3 Is an Indirect Caspase-1 Target and Controls IL-1β Secretion by Inflammasomes. *Cell Rep* **18**, 1285–1297 (2017).
61. Wu, Y., Li, X., Yu, J., Björkholm, M. & Xu, D. ASF1a inhibition induces p53-dependent growth arrest and senescence of cancer cells. *Cell Death Dis* **10**, 1–15 (2019).
62. Dutto, I., Scalera, C. & Prosperi, E. CREBBP and p300 lysine acetyl transferases in the DNA damage response. *Cell Mol Life Sci* **75**, 1325–1338 (2018).
63. Li, X. *et al.* Chromatin context-dependent regulation and epigenetic manipulation of prime editing. 2023.04.12.536587 Preprint at https://doi.org/10.1101/2023.04.12.536587 (2023).
64. Chen, R. *et al.* Enhancement of a prime editing system via optimal recruitment of the pioneer transcription factor P65. *Nat Commun* **14**, 257 (2023).
65. Gasperini, M. *et al.* A Genome-wide Framework for Mapping Gene Regulation via Cellular Genetic Screens. *Cell* **176**, 377-390.e19 (2019).
66. Zhang, T., Zhang, Z., Dong, Q., Xiong, J. & Zhu, B. Histone H3K27 acetylation is dispensable for enhancer activity in mouse embryonic stem cells. *Genome Biology* **21**, 45 (2020).
67. Bulusu, V. *et al.* Acetate Recapturing by Nuclear Acetyl-CoA Synthetase 2 Prevents Loss of Histone Acetylation during Oxygen and Serum Limitation. *Cell Reports* **18**, 647–658 (2017).
68. Soaita, I. *et al.* Dynamic protein deacetylation is a limited carbon source for acetyl-CoA–dependent metabolism. *Journal of Biological Chemistry* **299**, (2023).
69. Narita, T. *et al.* Acetylation of histone H2B marks active enhancers and predicts CBP/p300 target genes. *Nat Genet* **55**, 679–692 (2023).
70. Jain, K. *et al.* An acetylation-mediated chromatin switch governs H3K4 methylation read-write capability. 2022.02.28.482307 Preprint at https://doi.org/10.1101/2022.02.28.482307 (2023).
71. Zhao, W. *et al.* Investigating crosstalk between H3K27 acetylation and H3K4 trimethylation in CRISPR/dCas-based epigenome editing and gene activation. *Sci Rep* **11**, 15912 (2021).
72. Martin, B. J. E. *et al.* Transcription shapes genome-wide histone acetylation patterns. *Nat Commun* **12**, 210 (2021).
73. Millán-Zambrano, G., Burton, A., Bannister, A. J. & Schneider, R. Histone post-translational modifications — cause and consequence of genome function. *Nat Rev Genet* **23**, 563–580 (2022).

74. Yao, D. *et al.* Multi-center integrated analysis of non-coding CRISPR screens. 2022.12.21.520137 Preprint at https://doi.org/10.1101/2022.12.21.520137 (2022).
75. Chardon, F. M. *et al.* Multiplex, single-cell CRISPRa screening for cell type specific regulatory elements. 2023.03.28.534017 Preprint at https://doi.org/10.1101/2023.03.28.534017 (2023).
76. Morris, J. A. *et al.* Discovery of target genes and pathways at GWAS loci by pooled single-cell CRISPR screens. *Science* **380**, eadh7699 (2023).
77. Barry, T., Wang, X., Morris, J. A., Roeder, K. & Katsevich, E. SCEPTRE improves calibration and sensitivity in single-cell CRISPR screen analysis. *Genome Biology* **22**, 344 (2021).
78. Liu, N. *et al.* Histone H3 lysine 27 crotonylation mediates gene transcriptional repression in chromatin. *Molecular Cell* **83**, 2206-2221.e11 (2023).
79. Schulze, J. M., Wang, A. Y. & Kobor, M. S. YEATS domain proteins: a diverse family with many links to chromatin modification and transcriptionThis paper is one of a selection of papers published in this Special Issue, entitled CSBMCB's 51st Annual Meeting – Epigenetics and Chromatin Dynamics, and has undergone the Journal's usual peer review process. *Biochem. Cell Biol.* **87**, 65–75 (2009).
80. Gasperini, M., Tome, J. M. & Shendure, J. Towards a comprehensive catalogue of validated and target-linked human enhancers. *Nat Rev Genet* **21**, 292–310 (2020).
81. Dominguez, A. A. *et al.* CRISPR-Mediated Synergistic Epigenetic and Transcriptional Control. *The CRISPR Journal* **5**, 264–275 (2022).
82. Zulkower, V. & Rosser, S. DNA Chisel, a versatile sequence optimizer. *Bioinformatics* **36**, 4508–4509 (2020).
83. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, e47 (2015).
84. Levesque, S., Verma, A. & Bauer, D. E. Nucleotide metabolism constrains prime editing in hematopoietic stem and progenitor cells. 2023.10.22.563434 Preprint at https://doi.org/10.1101/2023.10.22.563434 (2023).
85. Xu, L., Liu, Y. & Han, R. BEAT: A Python Program to Quantify Base Editing from Sanger Sequencing. *The CRISPR Journal* **2**, 223–229 (2019).
86. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
87. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
88. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**, 550 (2014).
89. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* **36**, 411–420 (2018).

Figure Legends

Figure 1: Targeting of p300 variants to OCT4 regulatory elements results in divergent transcriptional responses.

(A) Construct designs for testing dCas9-p300 variants.

(B) Western blot analysis of whole proteome lysine crotonylation probed using anti-KCr antibody in HEK293T cells overexpressing dCas9-p300 variant.

(C) *In vitro* acylation activity assay of dCas9-p300 variants measured using the acyl-CoA co-factor indicated on the X and Y axes. Activity was measured using the N-terminal H3 peptide (n = 5, mean ± sem).

(D) Top: Schematic of the OCT4 locus with red rectangles indicating target gRNA location relative to transcription start site. Bottom: RT-qPCR of OCT4 gene expression upon dCas9-p300 variant targeting in HEK293T cells (n = 3, mean ± sem).
(E) Relative enrichment of indicated histone modification measured by CUT&RUN-qPCR following delivery of dCas9-p300 variant and OCT4 distal enhancer targeting gRNAs in HEK293T cells (n = 4-6, mean ± sem).
(F) Same as (E) with OCT4 promoter targeting gRNAs (n = 4-6, mean ± sem).

Figure 2: Developing a high throughput screen to assess protein expression and transcriptional activity in mammalian cells.
(A) Schematic of the screening pipeline. A reporter gene is stably integrated into the AAVS1 locus with 5x upstream TetO sites in HEK293T cells. A library of p300 variants fused to rTetR is transduced into the reporter cell line, selected with blasticidin, and induced with doxycycline for 48 hours before sorting and sequencing.
(B) PDB (5LKU) crystal structure of the p300 core domain bound to Coenzyme A.
(C) Citrine and mCherry fluorescence levels of an individual denoted variant of rTetr-p300 were measured by flow cytometry 11 days after transduction.
(D) Heatmap illustrating effects of all single mutations on citrine activation reporter. Deleterious mutations are denoted in magenta and beneficial mutations in green. Squares in yellow did not appear in the screening results. Ratios are shifted to the median of negative degenerate sequence controls (n=2).
(E) Heatmap illustrating effects of all single mutations on mCherry expression reporter. Deleterious mutations are denoted in black and beneficial mutations in red. Squares in yellow did not appear in the screening results. Ratios are shifted to the median of negative degenerate sequence controls (n=2).
(F) Top: Construct design for testing epigenome effector domains protein stability against mRNA expression levels. Bottom: Representative flow cytometry plots of mRNA expression (mChe) across 4 different epigenome editor effector domains with and without catalytic activity.

Figure 3: I1417N mutation reduces p300 cytotoxicity and improves viral delivery while preserving activity
(A) Scatter plot depicting the relationship between reporter activation (x-axis) and protein expression (y-axis). In green/magenta are the top 100/bottom variants summing the log ratios of both fluorescent readouts. Hits from shaded regions were selected for individual validation.
(B) % of mCherry-positive cells across each single construct validation in transient transfection measured by flow cytometry.
(C) RT-qPCR of HBG1 gene expression upon dCas9-p300 variant targeting to HBG1 (x-axis) and HS2 enhancer (y-axis) (n = 3, mean ± sem).
(D) RT-qPCR of OCT4 gene expression upon dCas9-p300 variant targeting across OCT4 distal enhancer, proximal enhancer, and promoter, respectively (n = 3, mean ± sem).
(E) Western blot analysis of HEK293T cells overexpressing dCas9-p300 variant probing for anti-Cas9 with anti-Tubulin as loading control.

(F) Fluorescence microscopy image of HEK293T cells transiently transfected with dCas9-p300 variant bicistronically expressed with mCherry and scrambled gRNA. Top panel: mCherry channel fluorescence. Bottom panels: DAPI staining. Scale bar: 100um.
(G) Cell viability as measured using alamarBlue™ cell viability reagent 48 hours post-transient transfection with dCas9-p300 variant and scrambled gRNA in HEK293T cells.
(H) % of mCherry-positive cells in 2%v/v transduced K562 cells measured by flow cytometry.
(I) RT-qPCR of HBG1 gene expression upon dCas9-p300 or CBP variant targeting to HBG1 (x-axis) and HS2 enhancer (y-axis) (n = 3, mean ± sem).

Figure 4: Immunoprecipitation-mass spectrometry (IP-MS) reveals different classes of protein interactions within different p300 variants.
(A) Venn diagram of hits called in IP-MS experiment (n = 2, P < 0.05)
(B) Dot chart showing top 84 protein hits by greatest -log(P-value) relative to dCas9 only bait. Classes of interactions are hierarchically clustered and subgroups are named on characteristic similarity.
(C) Volcano plot of the protein interactome of the p300 I1417N variant relative to the wild type. Red highlights correspond to P < 0.05 (n = 2)
(D) Gene ontology analysis of 1417N-enriched hits.
(E) Percentage editing as measured by BFP+ cells in GFP-BFP conversion assay or % editing via prime editing co-transfected with epigenome editor targeting the same site in HEK293T cells (n = 2-3, mean ± sem).

Figure 5. dCas9-p300 I1417N has reduced off-targeting activity.
(A) Transcriptomes of HEK293T cells transiently expressing dCas9-p300 WT compared against dCas9 (n = 2 biological replicates). Each black dot indicates the expression level of an individual gene. Red dashed lines indicate the 2x difference between sample groups.
(B) Genomic coordinates spanning ~5,248,000 to ~ 5,255,256 bp of human chromosome 11 (GRCh38/hg38) are shown along with H2BK20ac, H3K27ac, H3K4me3, and H3K27me3 CUT&RUN enrichment data with the indicated dCas9-p300 variant and 4 HBG1-targeting sgRNAs in HEK293T cells (n = 2 biological replicates).
(C) Profile of H3K27ac enrichment at transcription start sites.
(D-F) Relative enrichment of indicated histone modification (H2BK20ac, H3K4me3, or H3K27me3) measured by CUT&RUN-qPCR following delivery of dCas9-p300 variant and HBG1 promoter targeting gRNAs in HEK293T cells (n = 4, mean ± sem).

Figure 6. A high MOI Perturb-seq screen to identify effector-specific cis regulatory elements.
(A) Schematic of the experimental framework. Monoclonal KS62 cells expressing dCas9-p300 fusions were transduced at high MOI (~20) to introduce sgRNAs targeting regulatory elements of interest followed by scRNA-seq.
(B) Violin plots of expression change over the null control of all statistically significant hits across p300 WT and p300 I1417N (n=48 and 24 hits, respectively).

(C-D) Volcano plot shows the statistically significant hits of each effector. Red highlights correspond to P < 0.01. An FDR of 0.10 was used after a Benjamini-Hochberg correction was used to call hits.
(E-F) Violin plots of normalized expression levels compared to the null control of select genes across p300 WT and p300 I1417N targeting promoters/enhancers.

STAR METHODS

KEY RESOURCE TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, peptides, and recombinant proteins		
Doxycycline hydrochloride	Sigma	D3447
DMEM	Gibco	11-995-073
RPMI	Gibco	11875119
FBS	Millipore Sigma	F2442-500ML
Penicillin-Streptomycin (10,000 U/mL)	Gibco	15140122
TRYPSIN 0.25% EDTA	Gibco	25200114
PBS, Phosphate Buffered Saline, 10X Solution,	Fisher	BP3994
Opti-MEM™ I Reduced Serum Medium	Gibco	31-985-070
Lipofectamine 3000	Thermo Fisher	L3000075
Triton X-100	Sigma	T9284-100ML
Tween20	Fisher	65-520-4100ML
Puromycin	Fisher	MIR940
Blasticidin	Invivogen	NC9016621
Polybrene	Fisher	TR1003G
LB Broth	Fisher	BP1426-2
LB Agar	Fisher	BP1425-2
10x Tris/Glycine/SDS Buffer	Bio-Rad	1610732
4–20% Mini-PROTEAN® TGX Stain-	Bio-Rad	4568096

Free™ Protein Gels		
Critical commercial assays		
CUT&RUN assay kitCUTANA™ ChIC/CUT&RUN Kit	Epicypher	14-1048
RNeasy Plus Mini Kit (250)	Qiagen	74136
iScript™ Advanced cDNA Synthesis Kit	Bio-Rad	1725038BUN
Trans-Blot® Turbo™ RTA Mini LF PVDF Transfer Kit	Bio-Rad	1704274
SsoAdvanced Univ SYBR Grn Suprmx 10x5 ml	Bio-Rad	1725275
Invitrogen™ Qubit™ dsDNA HS and BR Assay Kits (500 reactions)	Fisher	Q32854
QIAprep Spin Miniprep Kit (250)	Qiagen	27106
QIAGEN Plasmid Maxi Kit (25)	Qiagen	12163
QIAquick Gel Extraction Kit (250)	Qiagen	28706
Zymo DNA Clean & Concentrator	Zymo	D4004
QIAquick PCR Purification Kit (250)	Qiagen	28106
Deposited data		
IP-MS	This paper	PXD050050
DMS library screen	This paper	GSE255610
RNA-seq	This paper	GSE255610
CUT&RUN	This paper	GSE255610
scCRISPRa	This paper	GSE255610
Experimental models: Mammalian cell lines		

HEK293T	ATCC	CRL-3216
K562	ATCC	CRL-243
U2OS	ATCC	HTB-96
HeLa	ATCC	CRM-CCL-2
Umbilical Cord-Derived Mesenchymal Stem Cells; Normal, Human	ATCC	PCS-500-010
Bacterial and Virus Strains		
Endura electrocompete nt cells	Lucigen	60242-2
One Shot™ Stbl3™ Chemically Competent E. coli	Invitrogen	C737303
BL21(DE3) Competent E. coli	NEB	C2527H
Oligonucleotides		
See Supplementary Table 5	IDT	N/A
Oligonucleotide Libraries	Twist	N/A
Software and algorithms		
HT-recruit	HT-recruit Analyze	https://github.com/bintulab/HT-recruit-Analyze
DiffBind	DiffBind 3.10.1	https://bioconductor.org/packages/release/bioc/html/DiffBind.html
CellRanger	Cell Ranger Count v7.1.0	https://www.10xgenomics.com/support/software/cell-ranger/latest
SCEPTRE	SCEPTRE	https://github.com/Katsevich-Lab/sceptre
CFX Maestro	BioRad	https://www.bio-rad.com/
DESeq2	DESeq2 1.40.2	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
STAR	STAR Version 2.7.9a	https://github.com/alexdobin/STAR
feature_counts	Liao et al., 2014	https://subread.sourceforge.net/subread.html
FlowJo	FlowJo V10	https://www.flowjo.com
Prism	Graphpad	https://www.graphpad.com
Antibodies		
anti-KCr	PTM BioLabs	PTM-501
anti-Cas9	Diagenode	C15200203-100

anti-FLAG	Millipore Sigma	F1804-200UG
Anti-Tubulin hFAB Rhodamine	BioRad	12004165
Anti-Histone H3 (acetyl K18) antibody	Abcam	ab1191
anti-H3K27ac	Abcam	ab4729
anti-H3K27ac	EpiCypher	13-0045
anti-H3K4me3	Epicypher	13-0041
anti-H2BK20ac	Abcam	ab177430
anti-H3K27me3	Abcam	13-0055
Recombinant DNA		
See Supplementary Table 5	This paper	N/A

LEAD CONTACT AND MATERIALS AVAILABILITY

Lead contact

Information and requests for resources and reagents should be directed to and will be fulfilled upon reasonable request by the corresponding author, Isaac Hilton (isaac.hilton@rice.edu)

Materials availability

Materials used in this study will be provided upon reasonable request and available upon publication. Key plasmids will be deposited on Addgene.

METHODS

Cell culture

All experiments were performed within 20 passages of cell stock thaws. HEK293T (ATCC, CRL-11268), HeLa (ATCC, CCL-2), U2OS (ATCC, HTB-96), and K562 (ATCC, CRL-243) cells were purchased from American Type Cell Culture (ATCC, USA) and cultured in ATCC-recommended media supplemented with 10% FBS (Sigma-Aldrich) and 1% pen/strep (100units/ mL penicillin, 100µg/ mL streptomycin; Gibco) at 37° C and 5% CO2.

Plasmid construction

For spdCas9 encoding vectors (Addgene #52961 or 180293), cloning backbones were modified to have C-terminal entry sites for insertion of different effector domains. The p300 core domain was amplified from pLV-dCas9-p300-P2A-PuroR (Addgene #83889). Mutations within p300 were installed by PCR amplifying fragments of p300 with the variants encoded within and assembled into backbones via NEBuilder HiFi DNA Assembly (NEB, E2621). Similar strategies were employed for other backbones if C-terminal entry sites did not already exist, including ddCas12a (Addgene #128136), dCasMINI (Addgene #176269), and lenti pEF-rTetR(SE-G72P)-3XFLAG-LibCloneSite-T2A-mCherry-BSD-WPRE (Addgene #161926). To generate the entry vector for mutagenesis library cloning, p300 was shuttled into an intermediate vector (Addgene #79770), linearized by PCR at the site of interest, and an oligo encoding Esp3I sites was inserted in NEBuilder HiFi DNA Assembly. To generate the reporter plasmid with an SV40 core promoter, subcloning was performed to insert a gBlock (IDT) in the parental reporter vector, AAVS1-PuroR-9xTetO-minCMV-IGKleader-hIgG1_FC-Myc-PDGFRb-T2A-Citrine-PolyA (Addgene #161928). Protein sequences of all dCas9 constructs are shown in **Supplementary Note 1**.

gRNA cloning

Oligonucleotides encoding protospacer sequences were designed using the Design CRISPR Guides tool on Benchling, ordered from IDT, and cloned as described previously[10,67]. Briefly, gRNA cloning backbone (SpdCas9 - Addgene #47108, ddCas12a - Addgene #128136, dCasMINI - Addgene #180280, SpdCas9 for scCRISPRa experiment - Addgene #192506) was digested with corresponding enzymes (Esp3I, SapI, or BbsI). Oligos were annealed, phosphorylated, and ligated into the cloning vector using T4 Ligase (NEB). All gRNA protospacer sequences used in this study for spdCas9, dCas12a, and dCasMINI are listed in **Supplementary Table 5**.

Transfection Strategies

RT-qPCR and Western blot experiments: Transient transfections were performed in 24-well plates using 375ng of respective dCas9 expression vector and 125 ng of single gRNA vectors or equimolar pooled gRNA expression vectors. Plasmids were mixed with Lipofectamine 3000 (Invitrogen, L3000015) as per the manufacturer's instruction. Cells were harvested for analysis 72 hours post-transfection.

Immunoprecipitation experiments: HEK293T cells were transfected in 15 cm dishes with Lipofectamine 3000 and 37.5 µg of respective dCas9 expression vector and 12.5 µg of scrambled gRNA expression vector as per manufacturer instruction. Cells were harvested for analysis 48 hours post-transfection.

Single construct flow cytometry validation: Transient transfections were performed in 24-well plates using 500ng rTetR-p300 variant expression vectors. Plasmids were mixed with Lipofectamine 3000 (Invitrogen, L3000015) as per the manufacturer's instruction. Cells were harvested for flow cytometry 48 hours post-transfection.

Prime Editing experiments: On Day 0, 1.5e5 HEK293T cells were plated in a 24-well plate. On Day 1, cells were transfected with 187.5ng of respective dCas9 expression vector, 187.5ng of PEmax expression vector (Addgene #174820), 62.5ng of pegRNA, and 62.5ng of nsgRNA targeting the region of interest using Lipofectamine 3000. Cells were harvested for analysis 48 hours post-transfection.

Cas9 DNA donor integration experiments: On Day 0, 1.5e5 HEK293T cells were plated in a 24-well plate. On Day 1, cells were transfected with 125ng of respective dCas9 expression vector, 125ng of Cas9 expression vector, 125ng of DNA donor template (Addgene #161927), and 125ng of gRNA targeting the AAVS1 locus using Lipofectamine 3000. On Day 3, cells were passaged into a new 24-well plate and maintained as needed. On Day 14, cells were harvested for analysis.

Base Editing experiments: On Day 0, 1.5e5 HEK293T cells were plated in a 24-well plate. On Day 1, cells were transfected with 125ng of respective dCas9 expression vector, 125ng of AncBE4max expression vector (Addgene #112094), 125ng of BFP reporter (Addgene #138272), and 125ng of gRNA (Addgene #138271) using Lipofectamine 3000. Cells were harvested for analysis 48 hours post-transfection.

Western blotting

Cells were lysed in RIPA buffer (Thermo Scientific, 89900) with 1X protease inhibitor cocktail (Thermo Scientific, 78442), lysates were cleared by centrifugation, and protein quantitation was performed using the BCA method (Pierce, 23225). 15-50µg of lysate were separated using precast 7.5%, 10%, or 4–20% SDS-PAGE (Bio-Rad) and then transferred onto PVDF membranes using the Transblot-turbo system (Bio-Rad). Membranes were blocked using 5% BSA in 1X TBST and incubated overnight with primary antibody (anti-Cas9; 1:1000 dilution, Diagenode #C15200216, Anti-FLAG; 1:2000 dilution, Sigma- Aldrich #F1804, anti-β-Tubulin; 1:1000 dilution, Bio-Rad #12004166). Then membranes were washed with 1X TBST 3 times (5 mins each wash) and incubated with respective HRP-tagged secondary antibodies (1:2000 dilution) for 1 hr. Next membranes were washed with 1X TBST 3 times (5 mins each wash). Membranes were then

incubated with ECL solution (BioRad # 1705061) and imaged using a Chemidoc-MP system (BioRad). The β-tubulin antibody was tagged with Rhodamine (Bio-Rad #12004166) and was imaged using Rhodamine channel in Chemidoc-MP as per manufacturer's instruction.

Reverse-transcription quantitative PCR (RT-qPCR)

RNA (including pre-miRNA) was isolated using the RNeasy Plus mini kit (Qiagen #74136). 500-2000ng of RNA (quantified using Nanodrop 3000C; Thermo Fisher) was used as a template for cDNA synthesis (Bio-Rad #1725038). cDNA was diluted 10X and 4.5µL of diluted cDNA was used for each qPCR reaction in 10µL reaction volume. Real-time quantitative PCR was performed using SYBR Green Master Mix (Bio-Rad #1725275) in the CFX96 Real-Time PCR system with a C1000 Thermal Cycler (Bio-Rad). Results are represented as fold change above control after normalization to *GAPDH* in all experiments using human cells. For murine cells, 18s rRNA was used for normalization. Undetectable samples were assigned a Ct value of 45 cycles. All qPCR primers and cycling conditions are listed in **Supplementary Table 5.**

Deep Mutational Scanning Library Design

A deep mutational scanning library was designed including all single amino acid substitutions and deletions within the 1381-1420 residue positions of the p300 full-length protein (**Supplementary Table 1**). Amino acid sequences were reverse translated into DNA sequences using DNAChisel, a Python library for optimizing sequences with constraints and optimization objectives, as described previously[42,82]. During the sequence optimization step, codons were optimized, Esp3I sites were removed, and the GC content was constrained to be between 30 and 70% for every 80-nucleotide window. In total, the library consisted of 841 variants of p300 and 100 random controls.

Library Cloning

Oligonucleotides were synthesized as pooled libraries (Twist Biosciences) and PCR amplified. 50uL reactions were prepared with 5ng of template, 2.5uL of each 10mM primer, 10uL of Q5 Reaction Buffer (NEB), 10uL of Q5 High GC Enhancer (NEB), 1uL of 10nM DNTPs, and 0.5uL of Q5 High Fidelity DNA Polymerase. PCR amplification parameters were as follows: 3 minutes at 98˚C, then 29x cycles of 98˚C for 10 sec, 61˚C for 30 sec, 72˚C for 30 sec, and a final step of 72˚C for 2 minutes. Resulting DNA was gel extracted using a QIAgen gel extraction kit. Libraries were then cloned into a vector for lentiviral recruitment (rTetR-p300Entry) with 4x10uL Golden Gate reactions with each containing 75ng of predigested recruitment vector, 5ng of gel library, 0.13uL of T4 DNA Ligase (NEB), 0.75uL of Esp3I (NEB), and 1uL of T4 DNA ligase buffer. Thermocycling parameters were as follows: 30x cycles at 37˚C and 16˚C for 5 minutes each followed by a 5-minute final digestion and heat inactivation for 20 minutes at 70˚C. Reactions were pooled, purified using a DNA Clean and Concentrator-5 (Zymo), and eluted in 6uL of molecular grade water. Two tubes of 50uL Endura Electrocompetent Cells (Lucigen, Catalog #60242-2) were transformed with 2uL per tube of column-purified libraries as per manufacturer instruction. After recovery, cells were plated on 3 large 15cm^{2}LB plates with ampicillin along with 3x 10cm^2 LB plates with ampicillin with serial dilutions of the transformed culture to confirm maintenance of 30x library coverage. After overnight growth, plates were scraped and extracted with a Qiagen Maxiprep Kit.

Lentivirus Production

One day before transfection, HEK293T cells were seeded at ~40% confluency in a 10-cm plate. The next day cells were transfected at ~80–90% confluency. For each transfection, 10 µg of plasmid containing the vector of interest, 10 µg of pMD2.G (Addgene, 12259), and 15 µg of psPAX2 (Addgene, 12260) were transfected using Lipofectamine 3000 (Invitrogen, L3000015) or calcium phosphate. 14 hours post-transfection the media was changed. Supernatant was

harvested 24 and 48 h post-transfection and filtered with a 0.45-µm PVDF filter (Millipore, SLGVM33RS), and then virus was concentrated at 100X using Lenti-X™ Concentrator (Takara, 631232), aliquoted and stored at −80 °C. Lentiviral titers were measured by the Lenti-X™ qRT-PCR Titration Kit (Takara, 631232).

Measurement of Library Expression and Transcriptional Activity

5e6 reporter HEK293T cells expressing 9xTetO-SV40core were transduced with the lentiviral library containing the library of p300 variants on Day 0 by reverse transduction with 10ug/mL polybrene. On Day 2, transduced cells were selected with 10ug/mL blasticidin. Cells were maintained in 3 15cm^2 tissue culture plates. On Day 7, cells were treated with 1ug/mL doxycycline for 2 days. On Day 9, cells were collected for cell sorting. Citrine and mCherry levels were measured on a Sony MA900 cell sorter and ~100k cells were collected per bin corresponding to mChe^{lo}/Cit^{lo}, mChe^{lo}/Cit^{hi}, mChe^{hi}/Cit^{lo}, mChe^{hi}/Cit^{hi}.

Library preparation and sequencing

Genomic DNA was extracted with a DNEasy Blood and Tissue Kit (QIAgen) following manufacturer's protocol for each bin extracted from the sorting experiment (at 1e5 cells). DNA was eluted in EB and sequences were amplified by PCR. A two-step PCR process was used to append Illumina adapters as overhangs. 6x reactions were prepared with 1ug of genomic DNA, 5uL of each 10uM primer, and 25uL of NEBnext 2x Master Mix (NEB) was used for the first reaction. PCR amplification parameters were as follows: 3 minutes at 98˚C, then 21x cycles of 98˚C for 10 sec, 68˚C for 30 sec, 72˚C for 30 sec, and a final step of 72˚C for 2 minutes. Resulting DNA was gel extracted using a QIAgen Gel Extraction Kit (QIAgen). A second PCR was performed with the Illumina indexed primers with the following PCR amplification parameters: 3 minutes at 98˚C, then 11x cycles of 98˚C for 10 sec, 68˚C for 30 sec, 72˚C for 30 sec, and a final step of 72˚C for 2 minutes. Resulting DNA was again gel extracted using a QIAgen Gel Extraction Kit (QIAgen). Libraries were then quantified with a Qubit HS dsDNA High Sensitivity Kit (Thermo Fisher), pooled with 15% PhiX control (Illumina), and sequenced on an Illumina HiSeq 3000.

Deep Mutational Scanning Analysis

Sequencing reads were processed as described previously using the HT-recruit pipeline[42]. Briefly, reads from each quadrant were grouped by fluorophore expression level, mCherry$^{hi/lo}$ and Citrine$^{hi/lo}$. A custom reference was generated using Bowtie from the designed p300 library using the script 'makeIndices.py' and reads were aligned with a mismatch allowance of 3 using the script 'makeCounts.py'. The enrichments for each domain between HI and LO samples were determined using the script 'makeRhos.py'. Domains with < 5 reads were adjusted to 5 to avoid inflation of enrichment values due to low sequencing depth.

Immunoprecipitation and FLAG Pulldown

Cells were harvested with a cell scraper and spun at 300xg for 5 min. The packed cell volume was measured and 2.5 volumes of NETN buffer was added to the cell pellet. (NETN: 50mM Tris pH 7.3, 170mM NaCl, 1mM EDTA, 0.5% NP-40). Lysate was sonicated (30sec bust, 59sec off, repeat 6 times) using a Diagenode Bioruptor Pico (B01060010) and then centrifuged at 20,000xg, for 20min at 4˚C. Supernatant was harvested and then transferred to a fresh tube. 5µg of FLAG antibody (F1804-200UG) was added to the above supernatant and incubated for 1 hour at 4˚C rocking. Supernatant was spun again at 20,000xg for 20min and the supernatant was collected in a fresh tube. 20 µl of protein A bead slurry was added to the above supernatant and incubated for 1 hour at 4˚C with inverted rotation. Supernatant and beads were spun at 1000xg for 1min. Flow-through was saved for downstream QC testing. Beads were washed with 1ml NETN buffer and spun at 1000xg for 1min and supernatant was discarded. A flat-ended tip was used to remove

residual buffer solution from the beads completely. 20µl of 2X SDS loading dye was added to the beads and samples were heated for 8-10 min at 90˚C.

Mass Spectrometry

The immunoprecipitated samples were resolved on NuPAGE 10% Bis-Tris Gel (Life Technologies), each lane was excised into 6 equal pieces and combined into two peptide pools after in-gel digestion using trypsin enzyme. The peptides were dried in a speed vac and dissolved in 5% methanol containing 0.1% formic acid buffer. The LC-MS/MS analysis was carried out using the nano-LC 1000 system coupled to Orbitrap Fusion mass spectrometer (Thermo Scientific). Peptides were eluted on an analytical column (20 cm x 75 µm I.D.) filled with Reprosil-Pur Basic C18 (1.9 µm, Dr. Maisch GmbH, Germany) using 110 minutes discontinuous gradient of 90% acetonitrile buffer (B) in 0.1% formic acid at 200 nl/min (2-30% B: 86 min, 30-60% B: 6 min, 60-90% B: 8 min, 90-50% B: 10 min). The full MS scan was performed in Orbitrap analyzer in the range of 300-1400m/z at 120,000 resolution followed by an IonTrap HCD-MS2 fragmentation for a cycle time of 3 seconds with precursor isolation window of 3m/z, collision energy 30%, AGC of 50000, maximum injection time of 30 ms.

The MS raw data were searched using Proteome Discoverer 2.0 software (Thermo Scientific) with Mascot algorithm against human NCBI RefSeq database updated 2020_0324. The precursor ion tolerance and product ion tolerance were set to 20 ppm and 0.5 Da, respectively. Maximum cleavage of 2 with Trypsin enzyme, dynamic modification of oxidation (M), protein N-term acetylation, deamidation (N/Q) and destreak (C) was allowed. The peptides identified from the mascot result file were validated with a 5% false discovery rate (FDR). The gene product inference and quantification were done with label-free iBAQ approach using the 'gpGrouper' algorithm. For statistical assessment, missing value imputation was employed through sampling a normal distribution N (µ-1.8 σ, 0.8σ), where µ, σ are the mean and standard deviation of the quantified values. For differential analysis, we used the moderated t-test and log2 fold changes as implemented in the R package limma[83] and multiple-hypothesis testing correction was performed with the Benjamini–Hochberg procedure. To filter out background contaminants, we removed from further analysis proteins that were recovered in more than 50% of IP analyses in HEK293T cells according to the contaminant repository for affinity purification (CRAPome[56]). To quantify shared hits in the STRING database, all interactors of p300 with a combined score > 0.15 were extracted and compared against hits found in this study's IP-MS dataset.

Genotyping Prime Editing Efficiency

Genomic DNA was extracted with a DNEasy Blood and Tissue Kit (QIAgen) following manufacturer's recommendation. DNA was eluted in water and sequences were amplified by PCR using Q5 DNA polymerase and 32 cycles of amplification. Primers used in this study were described previously[84]. Resulting DNA was gel extracted using a QIAgen Gel Extraction Kit (QIAgen). Amplicons were sequenced using Sanger sequencing and the percentage of prime edited alleles was determined using BEAT[85].

RNA-Sequencing

RNA sequencing (RNA-seq) was performed in duplicate for each experimental condition. RNA was isolated from transfected cells using the RNeasy Plus mini kit (Qiagen, 74136). RNA-seq libraries were constructed using the TruSeq Stranded Total RNA Gold (Illumina, RS-122-2303). The qualities of RNA-seq libraries were verified using the Tape Station D1000 assay (Tape Station 2200, Agilent Technologies), and the quantities of RNA-seq libraries were checked again using real-time PCR (QuantStudio 6 Flex Real time PCR System, Applied Biosystem). Libraries were normalized and pooled and then 75 bp paired-end reads were sequenced on the HiSeq3000 platform (Illumina). Sequencing reads were aligned to the human genome (GRCh38.p13) using

the STAR (Spliced Transcripts Alignment to a Reference, version 2.7.9a) against the Gencode Release 36 primary assembly annotation[86]. Sequence alignment was conducted on an Amazon Web Services Elastic Compute Cloud Instance (EC2) with 64 GB of memory and 256 GB of storage. Read quantification and differential expression analysis were conducted on R (version 4.2.2). Read quantification was carried out with featureCounts[87]. Differential expression analysis was performed using DESeq2[88]. Volcano plots, PCA plots, and heat maps were created on DESeq2 using the analyzed data.

CUT&RUN

CUT&RUN was performed using the Epicypher CUTANA ChIC/CUT&RUN Kit (Epicypher, #14-1048). Briefly, 500k HEK293T cells were detached and harvested using 0.5mM EDTA (Fisher, #BP2482-500), washed once with 1X PBS and then resuspended in 300ul of wash buffer. Next each of 3 100ul aliquots (~1/3 of each 24 well) of cells were processed for H3K4me3 antibody (Epicypher, #13-0041), H3K27ac antibody (Epicypher, #13-0045), H3K27me3 antibody (Epicypher, #13-0055), H2BK20ac antibody (Abcam, ab177430), or input DNA, respectively. Cells were first immobilized on concanavalin A beads, and then incubated with respective antibody (0.5ug/sample) overnight at 4°C in antibody dilution buffer (cell permeabilization buffer + EDTA). On the following day, cells were washed twice with cell permeabilization buffer. After washing the beads, pAG-MNase was added to the immobilized cells and then incubated for 2 hours at 4°C to digest and release DNA. Libraries were prepared using the CUT&RUN Library Prep Kit (Epicypher, #14-1001) and DNA was quantified using the Qubit HS dsDNA High Sensitivity Kit (Thermo Fisher) and sequenced at a depth of 15M paired end reads per sample on a NextSeq 2000 (High Output 2x100). Samples were normalized using an E. Coli spike-in DNA (Epicypher, #18-1401).

For CUT&RUN-qPCR assays, purified DNA from either H3K4me3, H3K27ac, H3K27me3, or H2BK20ac antibody incubated samples were then assayed by qPCR. Relative enrichment of H3K4me3 and H3K27ac is expressed as fold change above control cells transfected with dCas9 plasmid and after normalization to purified input DNA. qPCR primers used for CUT&RUN are shown in **Supplementary Table 5**.

CUT&RUN Sequencing Analysis

Sequencing reads were aligned to Human hg38 using Bowtie2, with multi-aligned, duplicate, and blacklist region reads removed. RPKM-normalized bigWigs were generated using deepTools with –binsize set to 20. ComputeMatrix was used to calculate target enrichment relative to transcription start sites (reference-point mode) and annotated gene bodies (scale-regions mode), and heatmaps were generated with plotheatmap. Smoothed Bigwigs were generated for visualization in the Integrative Genomics Viewer (IGV), using the –smoothLength option set to 100.

scCRISPRa transduction

Monoclonal K562 cell lines expressing the construct of interest were generated by transducing cells (8ug/mL) at a range of %v/v and performing flow cytometry 2 days later. Wells receiving dilutions with ~40% mCherry+ were selected and plated for monoclonal lines by limiting dilution. Monoclonal lines were transduced (10ug/mL) with varying titers to assess viral copy number. Library transduction with a 1.5%v/v was selected for the scCRISPRa experiment. For this experiment, 500k cells were transduced with virus containing the gRNA library. Cells were spun out of polybrene-containing media and replaced with standard K562 culture media. At 2 days post-transduction, 1ug/mL puromycin was added to the culture. 9 days after transduction, cells were collected for scRNA-seq.

10X Genomics scRNA-sequencing with gRNA Capture

Cells were harvested and prepared as per the 10X Genomics Single Cell Protocols Cell Preparation Guide. ~10,000 cells were captured per lane using a 10X Chromium device. One lane was used per construct (dCas9-p300 WT and dCas9-p300 I1417N). Cells were captured using a 10X Chromium chip using the Chromium Next GEM Single Cell 3' Reagents Kit v.3 with Feature Barcoding Technology for CRISPR screening. Final libraries were sequenced using a NovaSeq 6000 for each p300 screen. Gene expression and CRISPR Guide Capture transcript libraries were pooled at a 4:1 ratio for sequencing.

Transcriptome and gRNA data processing and QC

Cell Ranger Count v7.1.0 was used to perform count matrix generation using default parameters. Human (GRCh38) 2020-A was used for the transcriptome reference. Cells with less than 10% mitochondrial reads and between 750 and 7000 genes detected were kept for analysis. Resulting processed matrices were used for all downstream analyses.

Gene Expression Quantification

Each of the sequencing datasets was processed using Seurat[65,89]. Cells with greater than 10% mitochondrial transcripts, or less than 4000 total RNA transcripts were removed. For each gRNA in the library, expression levels were gathered for all genes within 1Mb of the gRNA target site. For each gRNA-target gene pairing, the FindMarkers() function was used to calculate the log2() fold changes in expression using the following parameters: ident.1 = gRNA_Cells, ident.2 = Control_Cells, min.pct = 0, min.cells.feature = 0, min.cells.group = 0, features = target_gene, logfc.threshold = 0. Violin plots were created using Seurat on the top few hits by adjusted p-value for each dataset analyzed. To make the violin plots, cells were divided into those with the gRNA present and those without, and the resulting expression levels for the target gene for each set of cells were visualized.

High-MOI gRNA Assignment and Differential Expression Testing using SCEPTRE

SCEPTRE was implemented as described previously{Citation}. Our processed count matrices were used along with single-cell metadata as covariates to fit the SCEPTRE model. We tested gRNA-response pairs for differential expression for all genes within 1MB upstream and 1Mb downstream of the gRNA of interest. SCEPTRE p-values were adjusted with the Benjamini-Hochberg procedure and target genes were identified if they were significantly below a threshold value of 10% FDR.

Quantification and Statistical Analysis

Statistical information for all experiments is found in the figure legends. Illustrations were created using BioRender and Adobe Illustrator.

DATA AVAILABILITY

The raw and processed P300 screening data, RNA-seq, CUT&RUN, and scRNA-seq data are publicly available at GEO (GSE255610). The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD050050. Accession number for publicly available data analyzed in this paper are as follows: GSE96035 for H3K18cr and ENCBS847SOB for all else in Fig. S1A. Additional information required for reanalysis reported herein is available from the corresponding author upon request.

ACKNOWLEDGEMENTS

The authors thank all members of the Hilton lab for helpful discussions and insights. This work was supported by a Cancer Prevention & Research Institute of Texas (CPRIT) Award (RR170030) and NIH Awards (R35GM143532 and R56HG012206) to I.B.H. J.G. was supported by NSF NRT

Program 1828869. M.D.E. was supported by the American Heart Association predoctoral fellowship program (917025).

This project was supported in part by the Baylor College of Medicine (BCM) Genomic and RNA Profiling Core with funding from NIH NCI (P30CA125123), NIH S10 grant (1S10OD023469), CPRIT (RP200504). This project was also supported by the BCM Single Cell Genomics Core with funding from the NIH (S10OD018033, S10OD023469, S10OD025240) and P30EY002520. The project also received support from the BCM Mass Spectrometry Proteomics Core. The BCM Mass Spectrometry Proteomics Core is supported by the Dan L. Duncan Comprehensive Cancer Center NIH award (P30 CA125123), CPRIT Core Facility Award (RP210227).

AUTHOR CONTRIBUTIONS

J.G. and I.B.H. conceived the project and designed experiments with input from all authors. J.G. performed most experiments and analyzed the data with the assistance of all authors. J.G. and I.B.H. wrote the manuscript with input from all authors.

DECLARATION OF INTERESTS

J.G. and I.B.H are inventors on patents related to this work. J.G., J.L., B.M., M.E., and I.B.H. are inventors on patents related to genome and epigenome editing technologies. J.G. and I.B.H. are founders of Mercator Biosciences.

MAIN TEXT FIGURES

Figure 1

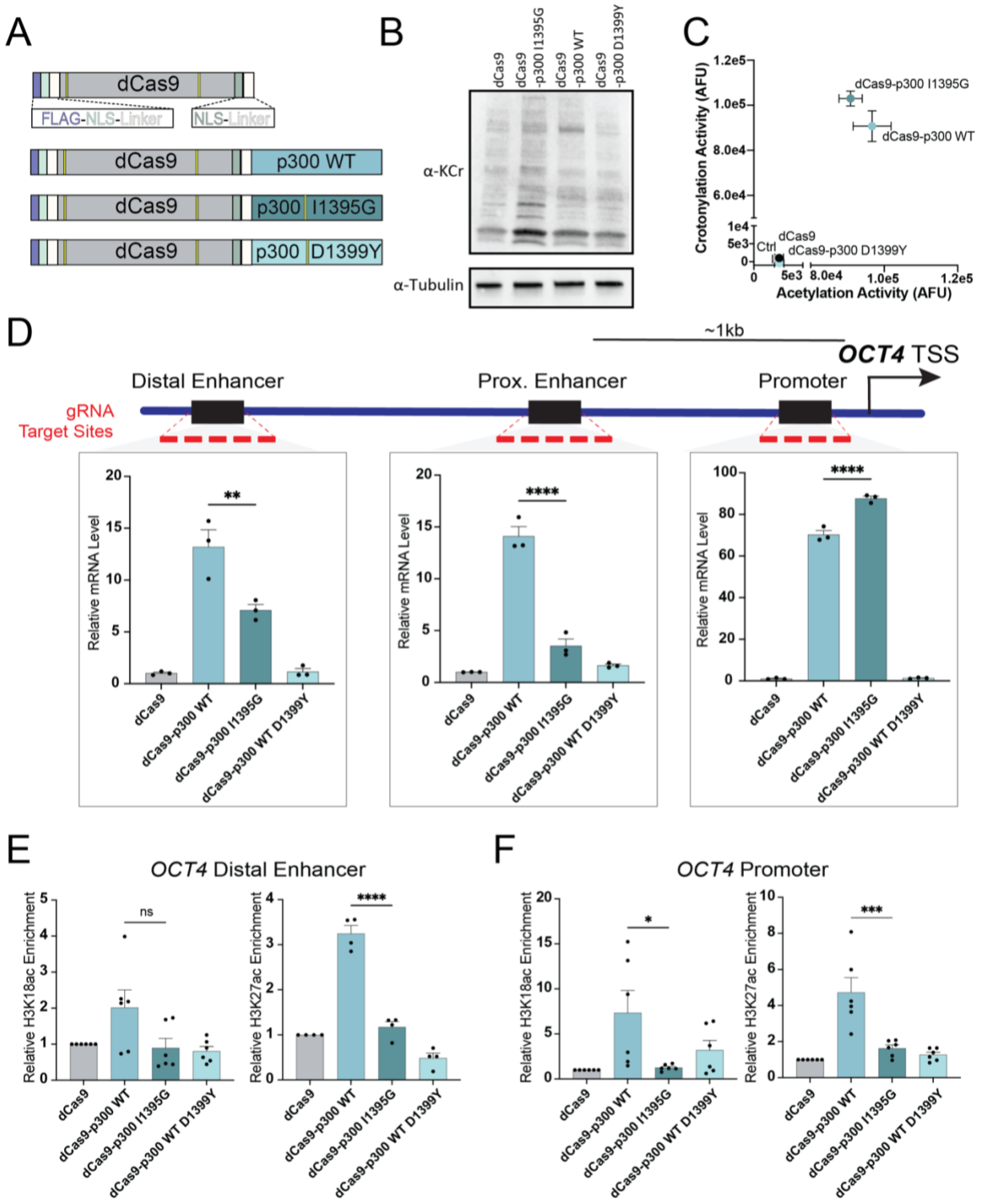


Figure 2

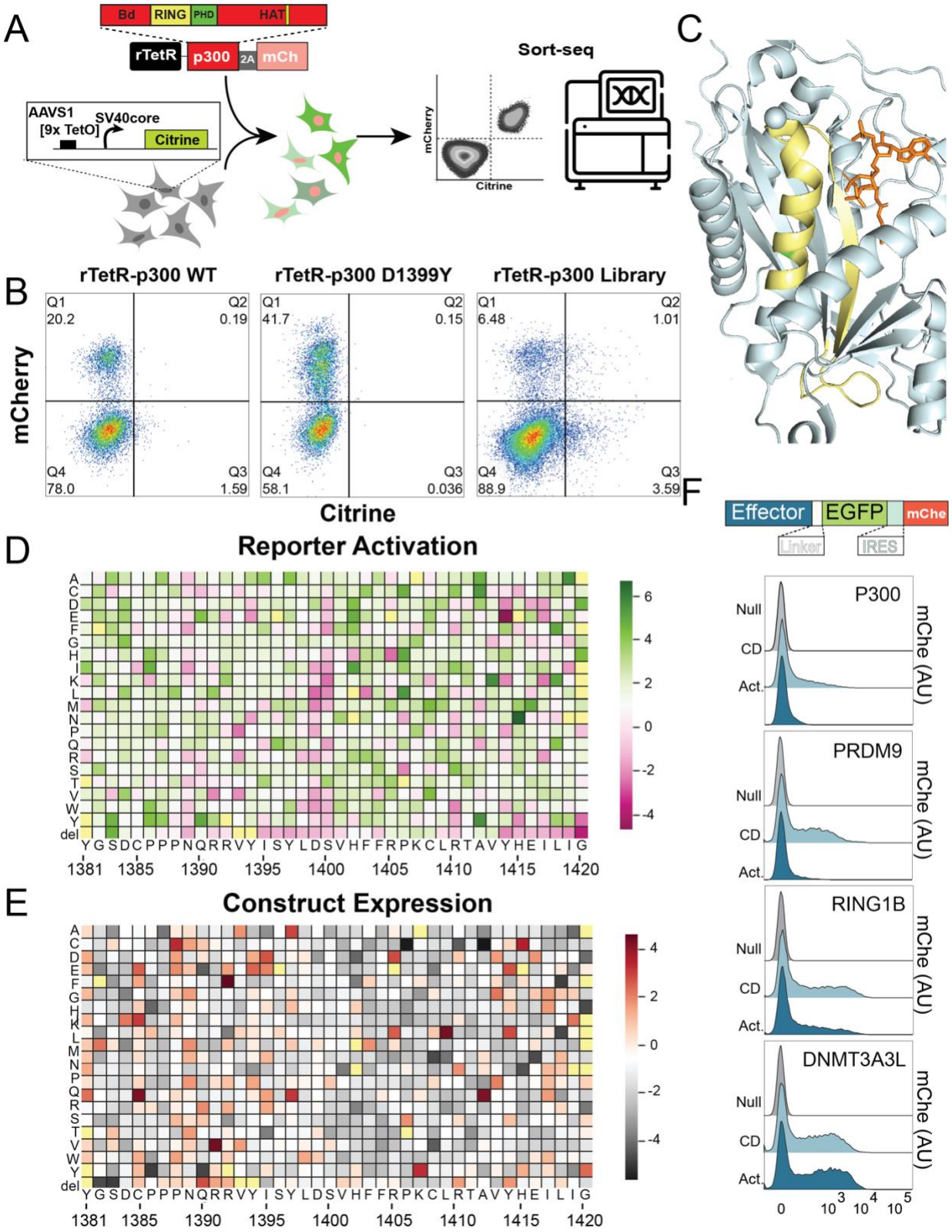

A
Bd
RING
PHD
HAT
rTetR
p300
2A
mCh
AAVS1
[9x TetO]
SV40core
Citrine
Sort-seq
mCherry
Citrine
C
B
rTetR-p300 WT
rTetR-p300 D1399Y
rTetR-p300 Library
Q1 20.2
Q2 0.19
Q4 78.0
Q3 1.59
Q1 41.7
Q2 0.15
Q4 58.1
Q3 0.036
Q1 6.48
Q2 1.01
Q4 88.9
Q3 3.59
mCherry
Citrine
F
Effector
EGFP
mChe
Linker
IRES
D
Reporter Activation
E
Construct Expression
P300
PRDM9
RING1B
DNMT3A3L
Null
CD
Act.
mChe (AU)
YGSDCPPPNQRRVYISYLDSVHFFRPKCLRTAVYHEILIG
1381 1385 1390 1395 1400 1405 1410 1415 1420

Figure 3

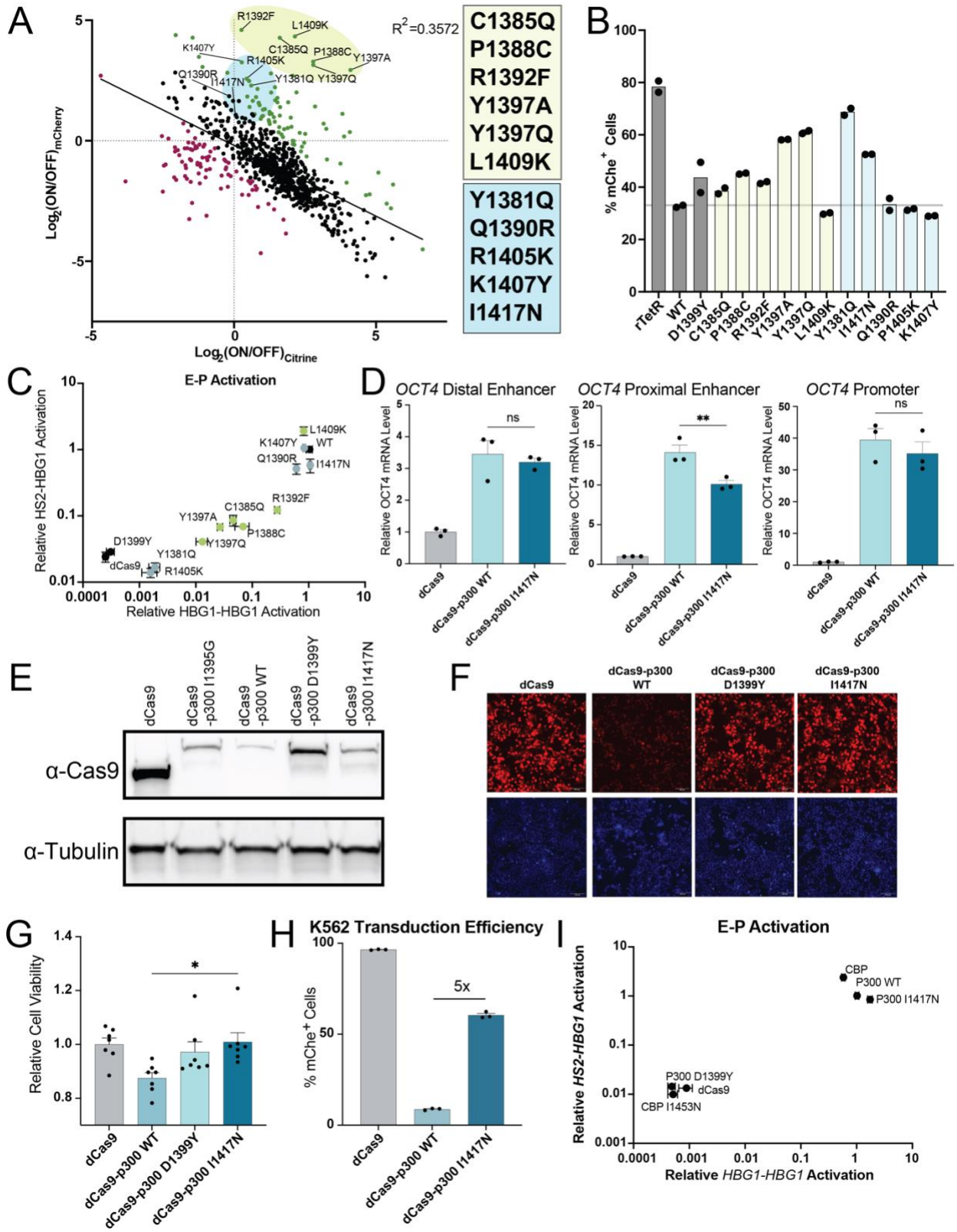

A
R1392F
L1409K
K1407Y
C1385Q
P1388C
Y1397A
R1405K
Q1390R
I1417N
Y1381Q
Y1397Q
R^2=0.3572
C1385Q
P1388C
R1392F
Y1397A
Y1397Q
L1409K
Y1381Q
Q1390R
R1405K
K1407Y
I1417N
$Log_2(ON/OFF)_{mCherry}$
$Log_2(ON/OFF)_{Citrine}$
B
% mChe+ Cells
rTetR
WT
D1399Y
C1385Q
P1388C
R1392F
Y1397A
Y1397Q
L1409K
Y1381Q
I1417N
Q1390R
P1405K
K1407Y
C
E-P Activation
Relative HS2-HBG1 Activation
Relative HBG1-HBG1 Activation
D
OCT4 Distal Enhancer
OCT4 Proximal Enhancer
OCT4 Promoter
Relative OCT4 mRNA Level
ns
**
ns
dCas9
dCas9-p300 WT
dCas9-p300 I1417N
E
dCas9
dCas9 -p300 I1395G
dCas9 -p300 WT
dCas9 -p300 D1399Y
dCas9 -p300 I1417N
α-Cas9
α-Tubulin
F
dCas9
dCas9-p300 WT
dCas9-p300 D1399Y
dCas9-p300 I1417N
G
Relative Cell Viability
*
dCas9
dCas9-p300 WT
dCas9-p300 D1399Y
dCas9-p300 I1417N
H
K562 Transduction Efficiency
% mChe+ Cells
5x
dCas9
dCas9-p300 WT
dCas9-p300 I1417N
I
E-P Activation
Relative HS2-HBG1 Activation
Relative HBG1-HBG1 Activation
CBP
P300 WT
P300 I1417N
P300 D1399Y
dCas9
CBP I1453N

Figure 4

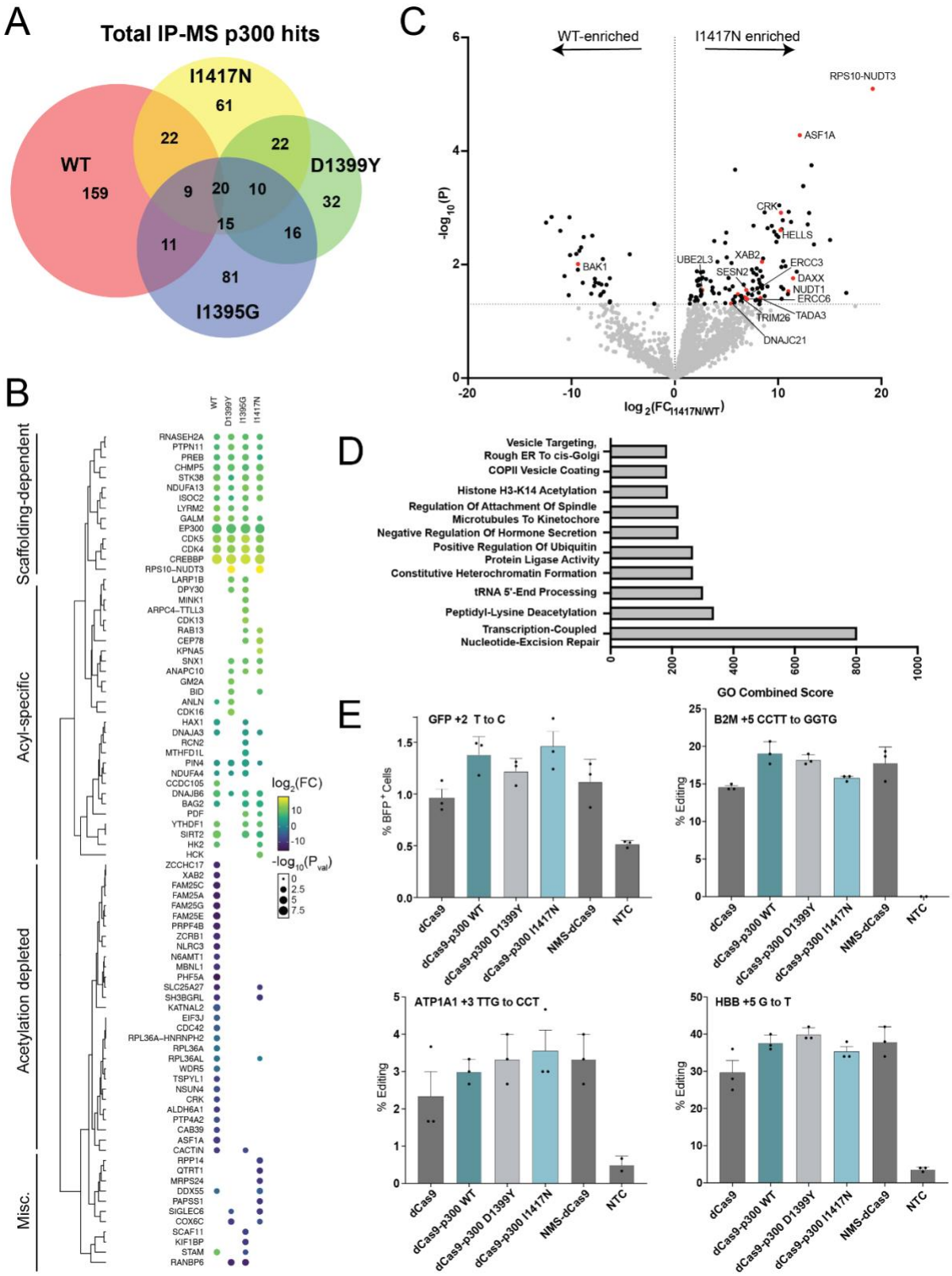

A
Total IP-MS p300 hits
I1417N
61
22
22
WT
159
D1399Y
32
9
20
10
15
16
11
81
I1395G
B
WT
D1399Y
I1395G
I1417N
Scaffolding-dependent
Acyl-specific
Acetylation depleted
Misc.
RNASEH2A
PTPN11
PREB
CHMP5
STK38
NDUFA13
ISOC2
LYRM2
GALM
EP300
CDK5
CDK4
CREBBP
RPS10–NUDT3
LARP1B
DPY30
MINK1
ARPC4–TTLL3
CDK13
RAB13
CEP78
KPNA5
SNX1
ANAPC10
GM2A
BID
ANLN
CDK16
HAX1
DNAJA3
RCN2
MTHFD1L
PIN4
NDUFA4
CCDC105
DNAJB6
BAG2
PDF
YTHDF1
SIRT2
HK2
HCK
ZCCHC17
XAB2
FAM25C
FAM25A
FAM25G
FAM25E
PRPF4B
ZCRB1
NLRC3
N6AMT1
MBNL1
PHF5A
SLC25A27
SH3BGRL
KATNAL2
EIF3J
CDC42
RPL36A–HNRNPH2
RPL36A
RPL36AL
WDR5
TSPYL1
NSUN4
CRK
ALDH6A1
PTP4A2
CAB39
ASF1A
CACTIN
RPP14
QTRT1
MRPS24
DDX55
PAPSS1
SIGLEC6
COX6C
SCAF11
KIF1BP
STAM
RANBP6
log2(FC)
10
0
-10
-log10(Pval)
0
2.5
5
7.5
C
WT-enriched
I1417N enriched
RPS10-NUDT3
ASF1A
CRK
HELLS
XAB2
UBE2L3
SESN2
BAK1
ERCC3
DAXX
NUDT1
ERCC6
TRIM26
TADA3
DNAJC21
-log10(P)
log2(FCI1417N/WT)
D
Vesicle Targeting, Rough ER To cis-Golgi
COPII Vesicle Coating
Histone H3-K14 Acetylation
Regulation Of Attachment Of Spindle Microtubules To Kinetochore
Negative Regulation Of Hormone Secretion
Positive Regulation Of Ubiquitin Protein Ligase Activity
Constitutive Heterochromatin Formation
tRNA 5'-End Processing
Peptidyl-Lysine Deacetylation
Transcription-Coupled Nucleotide-Excision Repair
GO Combined Score
E
GFP +2 T to C
% BFP+ Cells
B2M +5 CCTT to GGTG
% Editing
ATP1A1 +3 TTG to CCT
HBB +5 G to T
dCas9
dCas9-p300 WT
dCas9-p300 D1399Y
dCas9-p300 I1417N
NMS-dCas9
NTC

Figure 5

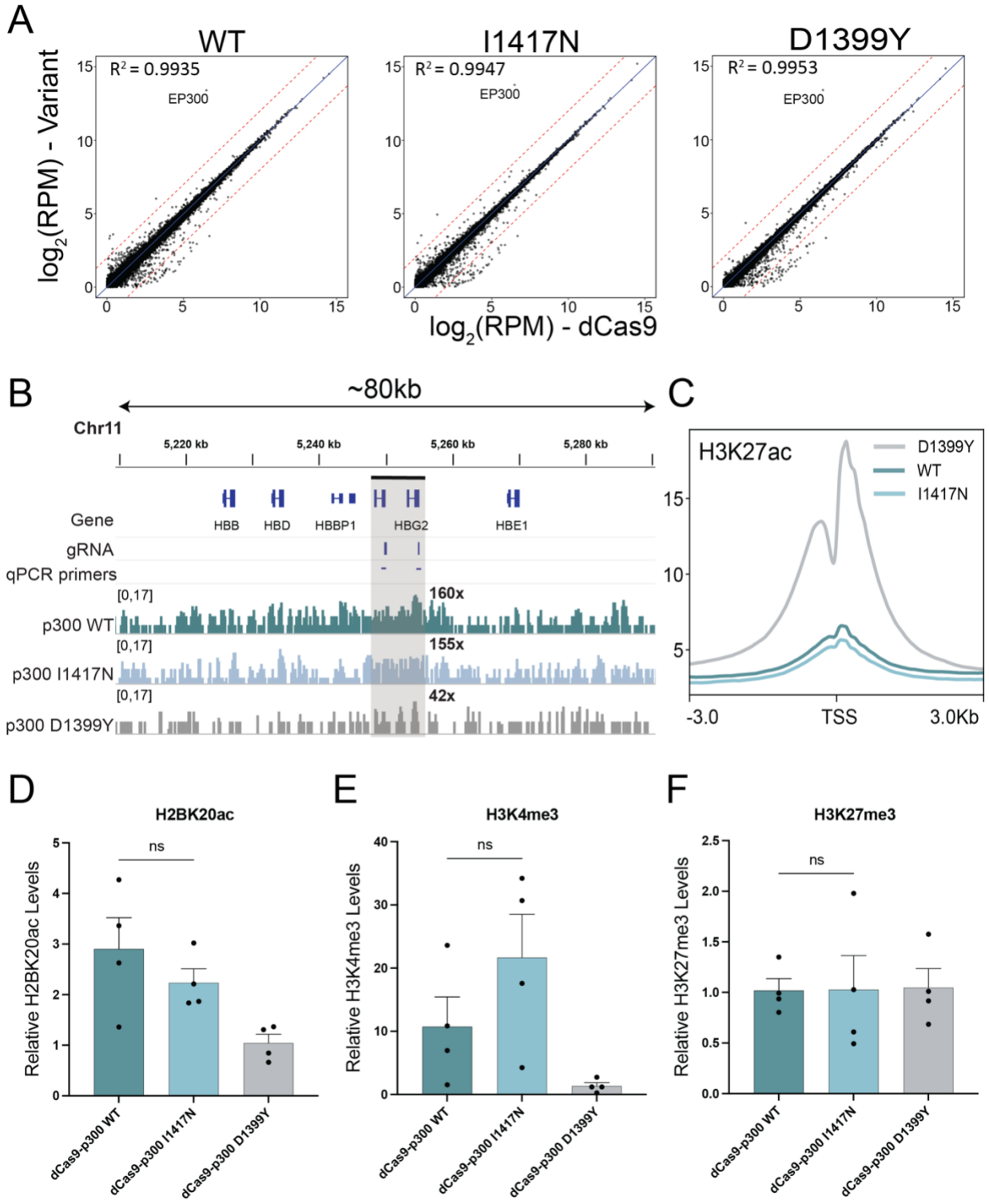
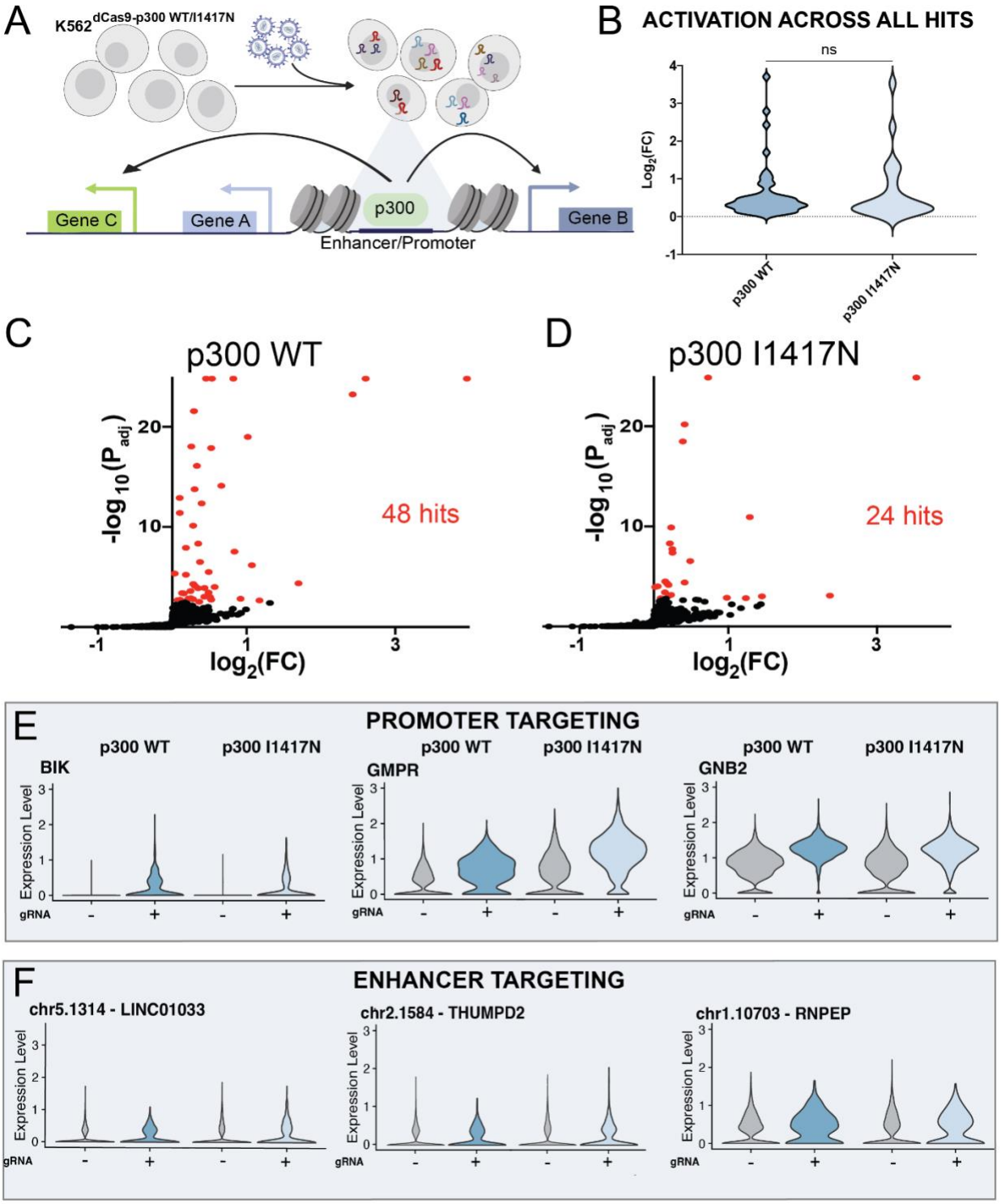

A
WT
I1417N
D1399Y
R² = 0.9935
R² = 0.9947
R² = 0.9953
EP300
log2(RPM) - Variant
log2(RPM) - dCas9
B
~80kb
Chr11
5,220 kb
5,240 kb
5,260 kb
5,280 kb
Gene
HBB
HBD
HBBP1
HBG2
HBE1
gRNA
qPCR primers
p300 WT
p300 I1417N
p300 D1399Y
[0,17]
160x
155x
42x
C
H3K27ac
D1399Y
WT
I1417N
-3.0
TSS
3.0Kb
D
H2BK20ac
Relative H2BK20ac Levels
ns
E
H3K4me3
Relative H3K4me3 Levels
F
H3K27me3
Relative H3K27me3 Levels
dCas9-p300 WT
dCas9-p300 I1417N
dCas9-p300 D1399Y

Figure 6


A
K562 dCas9-p300 WT/I1417N
Gene C
Gene A
p300
Enhancer/Promoter
Gene B
B
ACTIVATION ACROSS ALL HITS
ns
Log2(FC)
p300 WT
p300 I1417N
C
p300 WT
-log10(Padj)
48 hits
log2(FC)
D
p300 I1417N
24 hits
E
PROMOTER TARGETING
BIK
GMPR
GNB2
Expression Level
gRNA
F
ENHANCER TARGETING
chr5.1314 - LINC01033
chr2.1584 - THUMPD2
chr1.10703 - RNPEP

SUPPLEMENTAL FIGURES

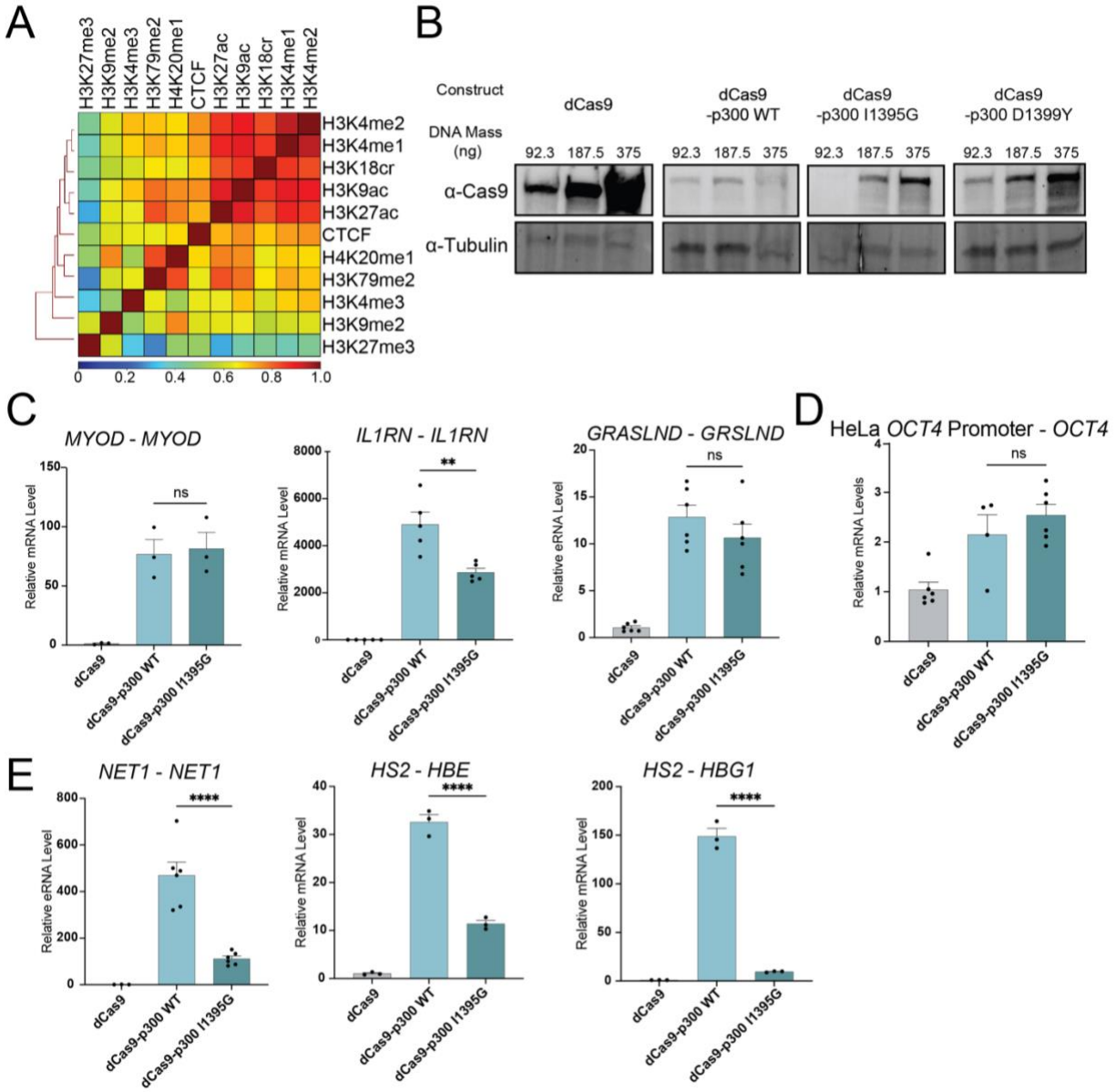


Figure S1: Targeting of p300 variants to OCT4 regulatory elements results in divergent transcriptional responses, related to Figure 1.

(A) Jaccard similarity scores of ChIP-seq data indicating the correlation of specific histone modifications and transcription factors at specific sites in HCT116 cells. Data was accessed via GSE96035 for H3K18cr and ENCBS847SOB for the rest of the groups.

(B) Western blot analysis of construct expression probed using an anti-Cas9 antibody with an anti-Tubulin loading control in HEK293T cells transfected with different amounts of dCas9-p300 variant plasmid DNA.

(C) RT-qPCR of promoter gene expression upon dCas9-p300 variant targeting in HEK293T cells (n = 4-6, mean ± sem).

(D) RT-qPCR of OCT4 gene expression upon dCas9-p300 variant targeting in HeLa cells (n = 4-6, mean ± sem).
(E) RT-qPCR of enhancer gene expression upon dCas9-p300 variant targeting in HEK293T cells (n = 3-6, mean ± sem).

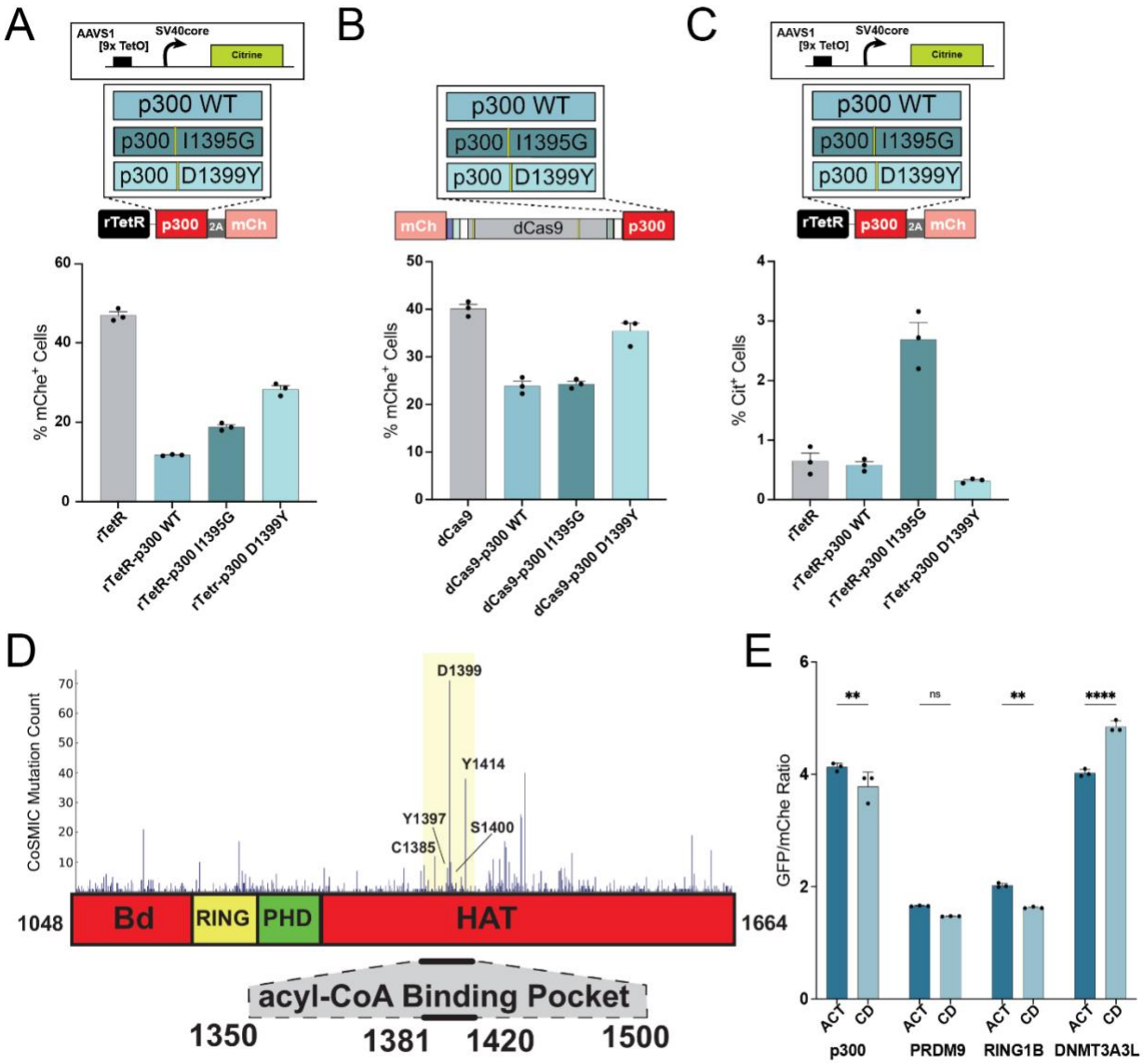


Figure S2: Developing a high throughput screen to assess protein expression and transcriptional activity in mammalian cells, related to Figure 2.

(A) Top: Construct design for testing rTetr-p300 variant activity/expression levels in transient transfection experiments. Bottom: Percentage of mCherry+ cells in transient transfection experiments in HEK293T reporter cells without doxycycline addition (n = 3, mean ± sem).

(B) Top: Construct design for testing mChe-dCas9-p300 variant expression levels in transient transfection experiments. Bottom: Percentage of mCherry+ cells in transient transfection experiments in HEK293T cells (n = 3, mean ± sem).

(C) Top: Construct design for testing rTetr-p300 variant activity levels in transient transfection experiments. Bottom: Percentage of mCherry+ cells in transient transfection validation experiments in HEK293T reporter cells 2 days after doxycycline addition (n = 3, mean ± sem).

(D) Domain structure of p300 core with Catalogue of Somatic Mutations in Cancer (COSMIC) database mutation occurrence overlaid to correspond to location in p300 core.

(E) Ratios of GFP to mCherry expression (geometric mean, AU) across 4 different epigenome editor effector domains with (ACT: active) and without catalytic activity (CD: catalytically deactivated) (n = 3, mean ± sem).

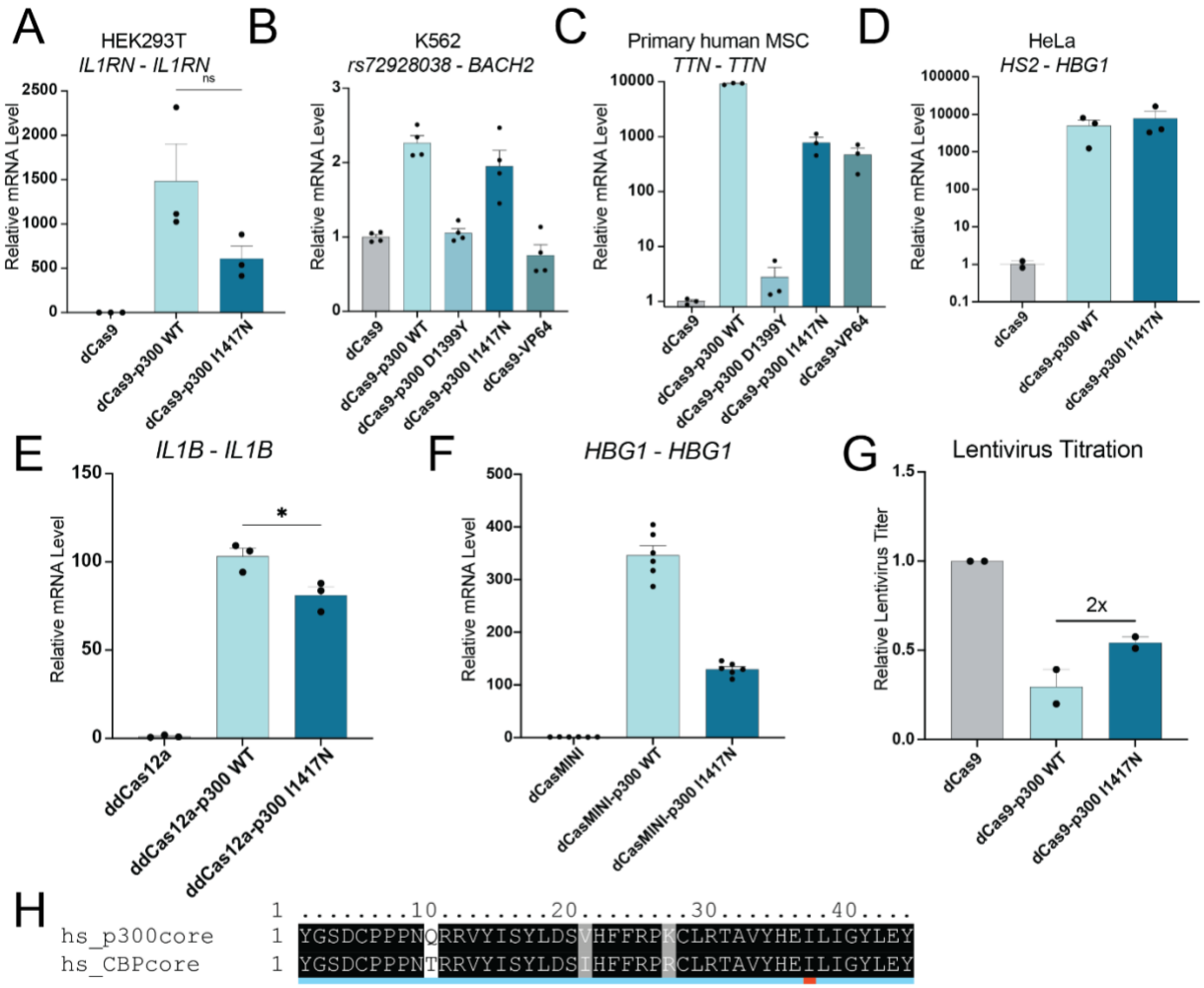


Figure S3: I1417N mutation reduces p300 cytotoxicity and improves viral delivery while preserving activity, related to Figure 3.

(A) RT-qPCR of IL1RN gene expression upon dCas9-p300 variant targeting IL1RN promoter in HEK293T cells (n = 3, mean ± sem).

(B) RT-qPCR of BACH2 gene expression upon dCas9-p300 variant targeting rs72928038 risk locus in K562 cells (n = 4, mean ± sem).

(C) RT-qPCR of TTN gene expression upon dCas9-p300 variant targeting TTN promoter in primary mesenchymal stromal cells (n = 3, mean ± sem).

(D) RT-qPCR of HBG1 gene expression upon dCas9-p300 variant targeting HS2 enhancer in HeLa cells (n = 3, mean ± sem).

(E) RT-qPCR of IL1B gene expression upon ddCas12-p300 variant targeting IL1B promoter in HEK293T cells (n = 3, mean ± sem).

(F) RT-qPCR of HBG1 gene expression upon dCasMINI-p300 variant targeting HBG1 promoter in HEK293T cells (n = 3, mean ± sem).

(G) Relative physical lentivirus titration following lentivirus transduction of dCas9-p300 variant in HEK293T cells (n = 3, mean ± sem).

(H) Sequence alignment of p300 and CBP flanking window of I to N mutation (I1417N in p300 and I1453N in CBP). The mutation site is demarcated below the sequence alignment in red.

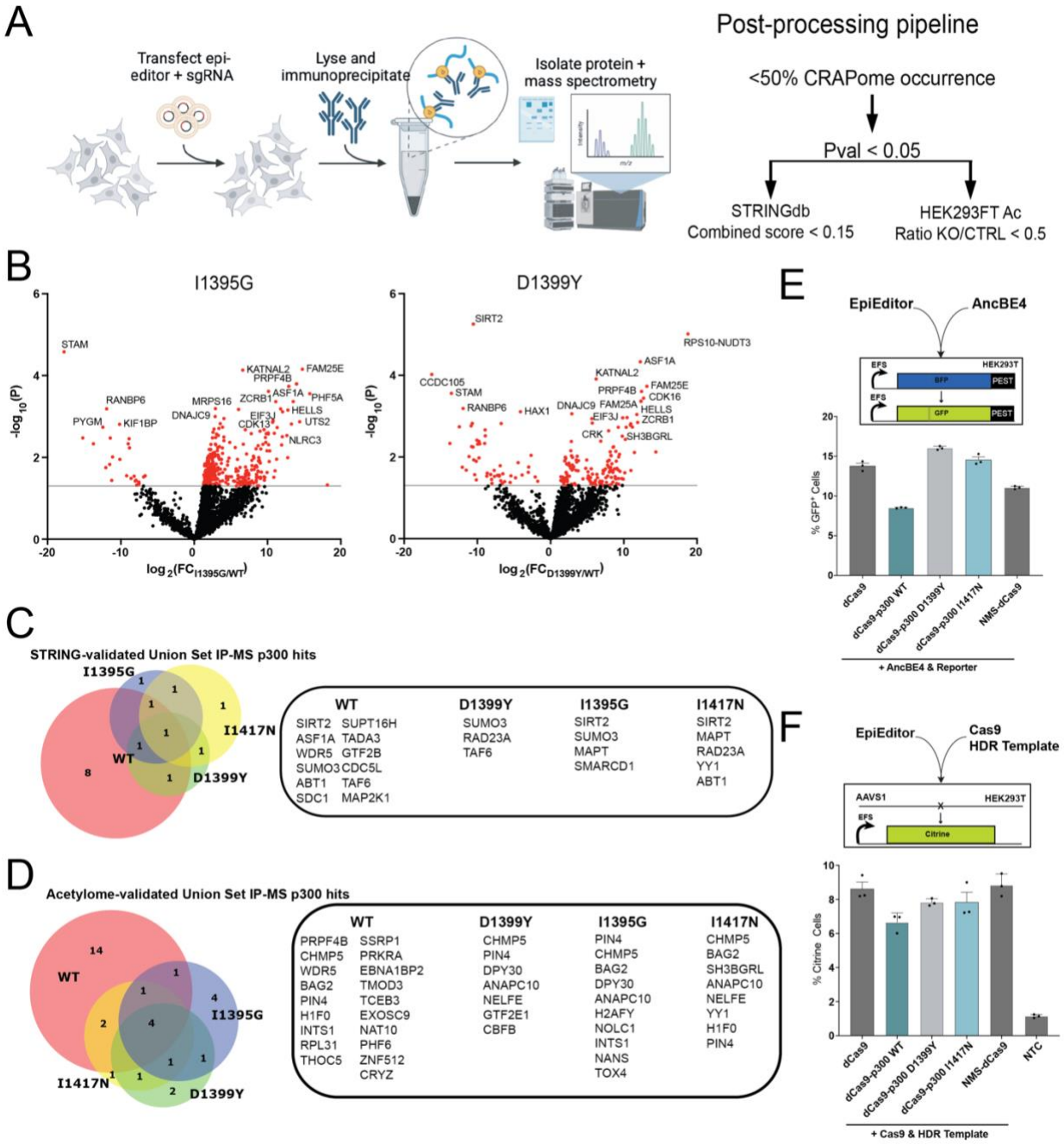


Figure S4: Immunoprecipitation-mass spectrometry (IP-MS) reveals different classes of protein interactions within different p300 variants, related to Figure 4.

(A) Schematic representation of the experimental design for IP-MS study of dCas9-p300 variants.

(B) Volcano plot of the protein interactome of the p300 I1395G variant (left) and p300 D1399Y variant (right) relative to p300 WT. Red highlights correspond to $P < 0.05$ ($n = 2$).

(C) Venn diagram containing union set of shared hits between IP-MS experiment and STRING database. Names of individual hit genes are listed to the right of the Venn diagram.

(D) Venn diagram containing union set of shared hits between IP-MS experiment and p300 acetylome database. Names of individual hit genes are listed to the right of the Venn diagram.
(E) Percentage base editing as measured by GFP+ cells in BFP-GFP conversion assay via AncBE4 co-transfected with epigenome editor targeting the same site in HEK293T cells (n = 3, mean ± sem).
(F) Percentage HDR integration as measured by Citrine+ cells in integration assay at AAVS1 locus via Cas9 co-transfected with HDR template and epigenome editor targeting the same site in HEK293T cells (n = 3, mean ± sem).

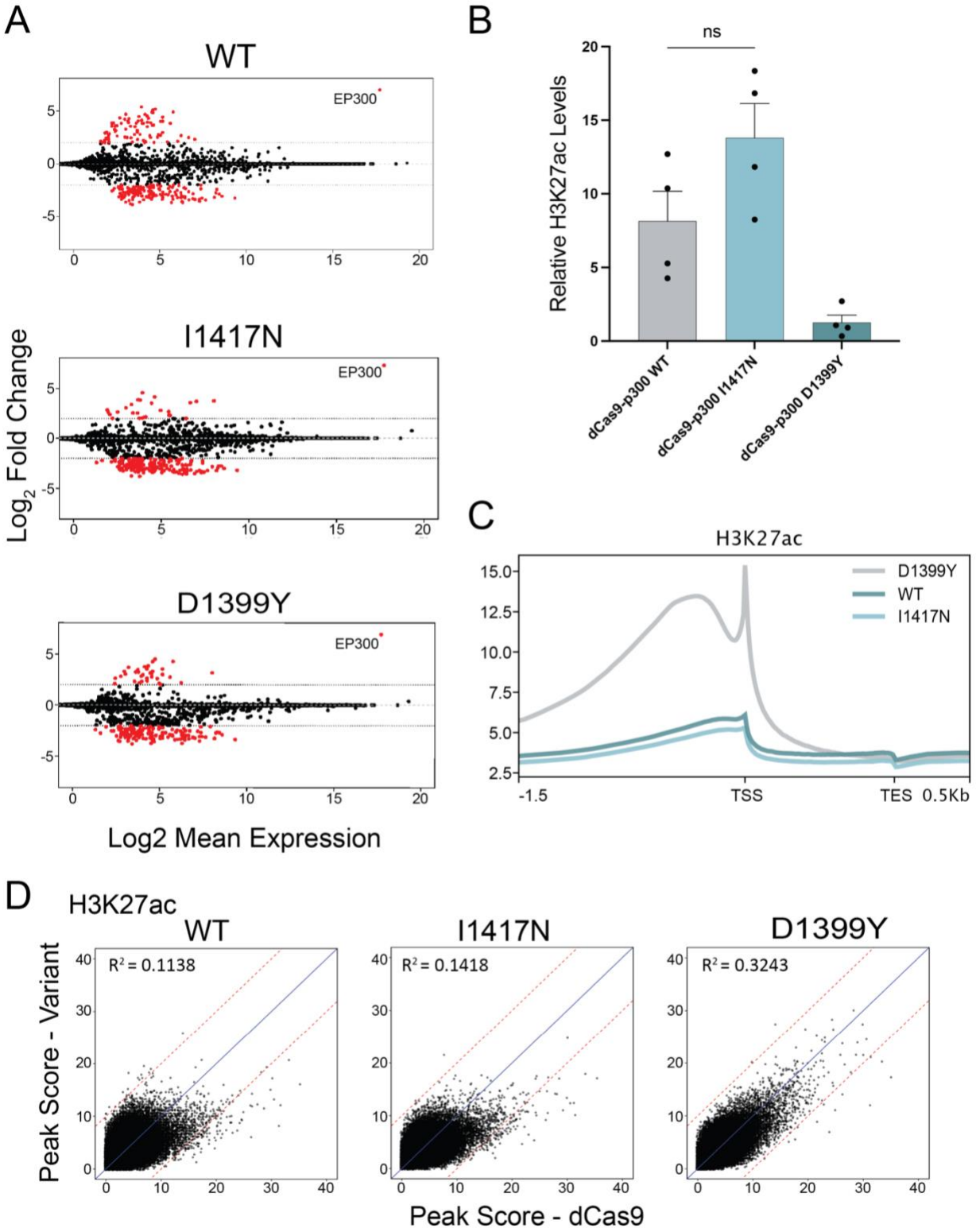


Figure S5. dCas9-p300 I1417N has reduced off-targeting activity, related to Figure 5. (A) MA plots of HEK293T cell transcriptomes transiently expressing dCas9-p300 WT compared against dCas9 (n = 2 biological replicates). Each black dot indicates the expression level of an

individual gene. Red dots indicate genes with greater or less than 2-fold expression than dCas9 using ashr shrinkage estimator.

(B) Relative enrichment of H3K27ac measured by CUT&RUN-qPCR following delivery of dCas9-p300 variant and HBG1 promoter targeting gRNAs in HEK293T cells (n = 4, mean ± sem).

(C) Metagene profile of H3K27ac enrichment.

(D) H3K27ac peak signal in HEK293T cells transiently expressing dCas9-p300 WT compared against dCas9 (n = 2 biological replicates). Each black dot indicates the normalized signal at a genomic window. Red dashed lines indicate the 2x difference between sample groups.

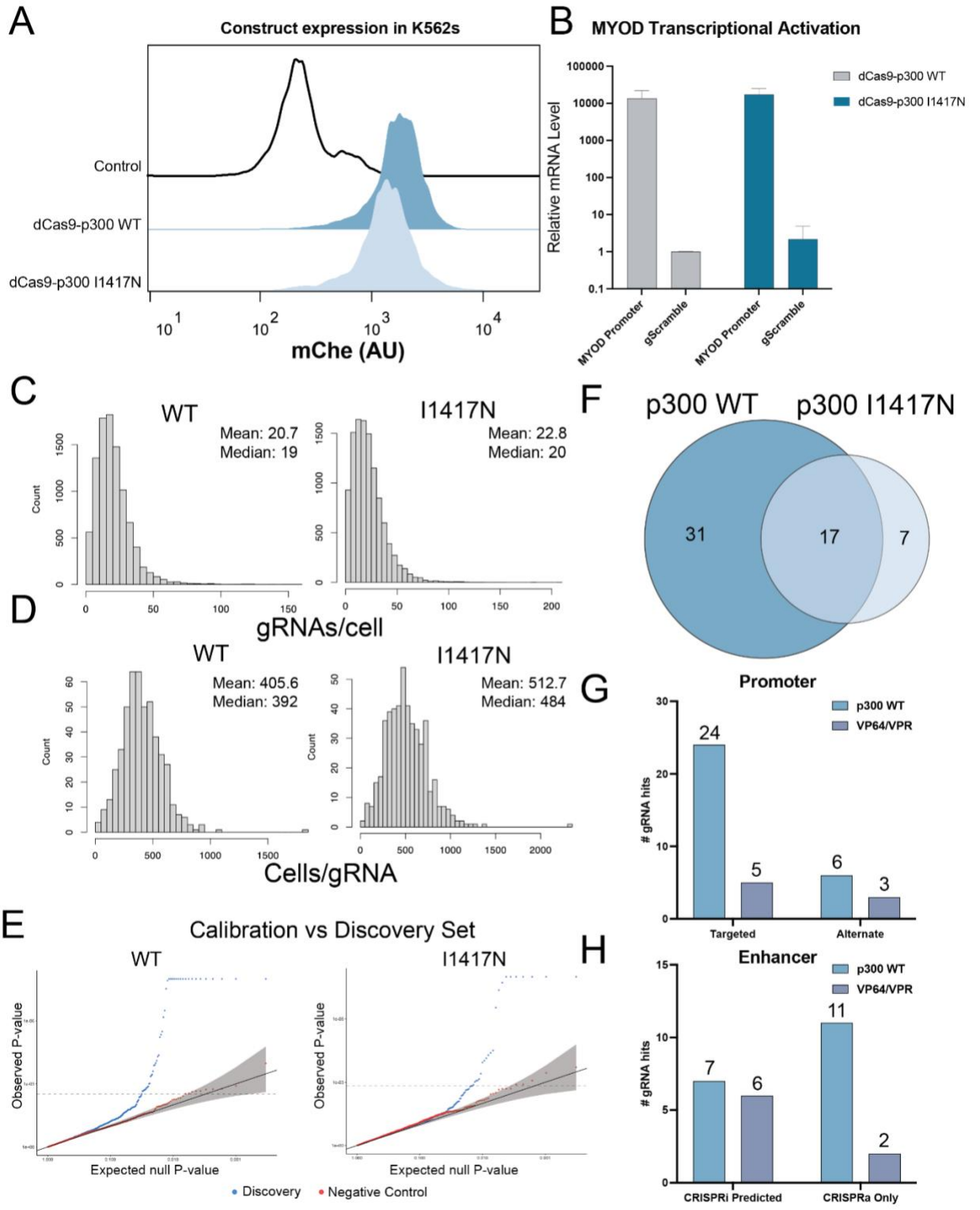


Figure S6. A high MOI Perturb-seq screen to identify effector-specific cis regulatory elements, related to Figure 6.

(A) Representative flow cytometry histogram of mCherry bicistronically expressed with dCas9-p300 variant.
(B) RT-qPCR of MYOD gene expression upon targeting MYOD promoter in monoclonal K562 cells stably expressing dCas9-p300 variant (n = 3, mean ± sem).
(C) Histogram of gRNAs identified per cell for p300 WT (left) and p300 I1417N (right).
(D) Histogram of the number of cells identified bearing a specific gRNA perturbation for p300 WT (left) and p300 I1417N (right).
(E) Quantile-quantile plots of *cis* effects within 1Mb (500kb upstream and downstream) of each gRNA. Negative control target-response pairs were generated by randomly sampling all responses. Negative control p-values are shown in red and discovery p-values in blue. The grey band indicates confidence intervals. The horizontal dashed line indicates the multiple testing threshold and points above this line are called significant.
(F) Bar graph of the number of gRNA hits that activate its cognate promoter (Targeted) or a different gene (Alternate) for p300 WT and VP64/VPR. VP64/VPR was derived from Chardon et al, BioRxiv 2023.
(G) Bar graph of the number of gRNA hits that activate a previously CRISPRi validated gene (CRISPRi Predicted) or a novel gRNA-enhancer pair (CRISPRa only) for p300 WT and VP64/VPR. VP64/VPR was derived from Chardon et al, BioRxiv 2023.
(H) Venn diagram of hits called between p300 WT and p300 I1417N using SCEPTRE processing pipeline.