Compact engineered human transactivation modules enable potent and versatile synthetic transcriptional control

Barun Mahata[1], Jing Li[1], Alan Cabrera[1], Daniel A. Brenner[1], Rosa Selenia Guerra-Resendez[1,2], Jacob Goell[1], Kaiyuan Wang[1], Mario Escobar[3], Yannie Guo[1], Abinand Krishna Parthasarathy[1], Isaac B. Hilton[1,2,3*]

[1] Department of Bioengineering, Rice University, Houston, TX, USA
[2] Systems, Synthetic, and Physical Biology Graduate Program, Rice University, Houston, TX, USA
[3] Department of BioSciences, Rice University, Houston, TX, USA

* To whom correspondence should be addressed. Tel: 713-348-8247; Email: isaac.hilton@rice.edu

Abstract
Engineered transactivation domains (TADs) combined with programmable DNA binding platforms have revolutionized synthetic transcriptional control. Despite recent progress in programmable CRISPR/Cas-based transactivation (CRISPRa) technologies, the TADs used in these systems often contain components from viral pathogens and/or are prohibitively large for many applications. Here we defined and optimized minimal TADs built from human mechanosensitive transcription factors (MTFs). We used these components to construct potent and compact multipartite transactivation modules (MSN, NMS, and eN3x9) and to build the CRISPR-dCas9 recruited enhanced activation module (CRISPR-DREAM) platform. We found that CRISPR-DREAM was specific, robust across mammalian cell types, and efficiently stimulated transcription from diverse regulatory loci within the human genome. We also showed that MSN and NMS were portable across Type I, II, and V CRISPR systems, TALEs, and ZF proteins, and further that these TADs permitted superior multiplexed transactivation. Finally, as a proof of concept, we used dCas9-NMS to efficiently reprogram human fibroblasts into iPSCs. Altogether, the compact human TADs, design rules, and fusion proteins we have developed here could be valuable for applications where sophisticated synthetic transactivation is needed.

Introduction
Nuclease deactivated CRISPR-Cas (dCas) systems can be used as programmable transcriptional modulators in cells and organisms[1-7]. For CRISPR/Cas based transactivation (CRISPRa) approaches, transcriptional activators can be recruited to genomic regulatory elements using direct fusions to dCas proteins[8-10], antibody-mediated recruitment in tandem with dCas proteins[11], or using engineered gRNA aptamer architectures[12, 13]. High levels of CRISPRa-driven transactivation have been achieved by shuffling[14], reengineering[15], or combining[8, 16] transactivation domains (TADs) and/or chromatin modifiers. However, many of the transactivation components used in CRISPRa systems have coding sizes that are restrictive for applications such as viral vector-based delivery. Moreover, most of the transactivation modules that display high potencies harbor components derived from viral pathogens, which could hamper clinical or *in vivo* use. Finally, there is an untapped repertoire of thousands of human transcription factors (TFs) and chromatin modifiers[17-19] that has yet to be systematically tested and optimized as

programmable transactivation components. These diverse human protein building blocks could be used to reduce the size of transactivation components, obviate the use of viral TFs, and possibly permit cell and/or pathway specific transactivation.

Mechanosensitive transcription factors (MTFs) modulate transcription in response to mechanical cues and/or external ligands[20, 21]. When stimulated, MTFs are shuttled into the nucleus where they can rapidly transactivate target genes by engaging key nuclear factors including RNA polymerase II (RNAP) and/or histone modifiers[22-25]. The dynamic shuttling of MTFs can depend upon both the nature and the intensity of stimulation. Mammalian cells encode several classes of MTFs, including serum regulated MTFs (e.g., YAP, TAZ, SRF, MRTF-A and B, and MYOCD)[21, 26], cytokine regulated/JAK-STAT family MTFs (e.g., STAT proteins)[27], and oxidative stress/antioxidant regulated MTFs (e.g., NRF2)[28]; each of which can potently activate transcription when appropriately stimulated. The robust, highly orchestrated, and relatively ubiquitous gene regulatory effects of these classes of human MTFs make them excellent potential sources of new non-viral TADs that could be leveraged as components of engineered CRISPRa systems and/or other synthetic gene activation platforms.

Here, we quantified the endogenous transactivation potency of dozens of different TADs derived from human MTFs in different combinations and across various dCas-based recruitment architectures. We used these data to design new multipartite transactivation modules, called MSN, NMS, and eN3x9. We used the MSN and NMS effectors to build the CRISPR-dCas9 recruited enhanced activation module (DREAM) platform and showed that CRISPR-DREAM potently stimulated transcription in primary human cells and cancer cell lines, as well as in murine and CHO cells. We also demonstrated that CRISPR-DREAM activated diverse classes of human RNAs from within the human genome. Further, we found that the MSN/NMS effectors were highly portable to smaller engineered dCas9 variants, natural orthologues of dCas9, dCas12a, Type I CRISPR/Cas systems, and TALE and ZF proteins. Moreover, we demonstrated that a dCas12a-NMS fusion enabled superior multiplexing transactivation capabilities compared to existing systems. Finally, we applied dCas9-NMS to efficiently reprogram human fibroblasts to induced pluripotency. Overall, the engineered transactivation modules that we have developed here are small, highly potent, devoid of viral sequences, versatile across programmable DNA binding systems, and enable robust multiplexed transactivation in human cells – important features that could be leveraged to test new biological hypotheses and engineer complex cellular functions.

Results

Select TADs from MTFs can activate transcription from diverse endogenous human loci when recruited by dCas9

We first isolated TADs from 7 different serum-responsive MTFs (YAP, YAP-S397A[29], TAZ, SRF, MRTF-A, MRTF-B, and MYOCD) and analyzed their ability to activate transcription when recruited to human promoters using either N- or C-terminal fusion to *Streptococcus pyogenes* dCas9 (dCas9), SunTag-mediated recruitment[11], or recruitment

via a gRNA aptamer and fusion to the MCP protein[12] (**Supplementary Fig. 1**). TADs derived from MRTF-A, MRTF-B, or MYOCD displayed consistent transactivation potential across recruitment architectures. We next compared the optimal recruitment strategies for MRTF-A and MRTF-B TADs because they were more potent than, or comparable to, the MYOCD TAD yet slightly smaller. Our results demonstrated that TADs from MRTF-A and B functioned best when fused to the MCP protein and recruited via gRNA aptamers (**Supplementary Fig. 2**), and further that this strategy could be used with pools or single gRNAs, and to activate enhancer RNAs (eRNAs) and long noncoding RNAs (lncRNAs).

Although the NRF2-ECH homology domains 4 and 5 (Neh4 and Neh5, respectively) within the oxidative stress/antioxidant regulated NRF2 human MTF have been shown to activate gene expression in Gal4 systems[22], we observed that neither Neh4 nor Neh5 were capable of potent human gene activation when recruited to promoters in any dCas9-based architecture (**Supplementary Fig. 3**). Therefore, we constructed an engineered TAD called eNRF2, consisting of Neh4 and Neh5 separated by an extended glycine-serine linker and found that the eNRF2 TAD stimulated high levels of transactivation in all dCas9-based recruitment configurations (**Supplementary Fig. 3**). Similar to the MRTF-A/B TADs, eNRF2 displayed optimal potency in the gRNA aptamer/MCP-based recruitment architecture and transactivated diverse human regulatory loci (**Supplementary Fig. 4**). We next tested whether TADs derived from one of 6 different cytokine regulated/JAK-STAT family MTFs (STAT1 – 6) could transactivate human genes but observed that single STAT TADs alone were incapable of potent transactivation regardless of dCas9-based recruitment context (**Supplementary Fig. 5**). Nevertheless, these data demonstrate that TADs from human MTFs can transactivate human loci when recruited via dCas9 and that these TADs are amenable to protein engineering.

Combinations of TADs from MTFs can potently activate human genes when recruited by dCas9

STAT proteins typically activate gene expression in combination with co-factors[30]. Therefore, we tested if TADs from different STAT proteins might synergize with other MTF TADs. We built 24 different bipartite fusion proteins by linking each STAT TAD to the N- or C- terminus of either the MRTF-A or MRFT-B TAD and then assayed the relative transactivation potential of each bipartite fusion when recruited to the human *OCT4* promoter using gRNA aptamer/MCP-based recruitment (**Supplementary Fig. 6**). Each of these 24 fusions markedly outperformed TADs from MRTF-A/B or STAT TADs alone, and one bipartite TAD configuration (MRFT-A/STAT1) was comparable to MCP fused to the dCas9-SAM derived bipartite p65-HSF1[12] module. We next investigated whether the eNRF2 TAD could further enhance the potency of the MRFT-A/STAT1 module by building tripartite fusions consisting of MRTF-A/STAT1/eNRF (MSN) or eNRF/MRTF-A/STAT1 (NMS) TADs. Both MSN and NMS stimulated *OCT4* mRNA synthesis to levels comparable to the state-of-the-art CRISPRa platforms (**Supplementary Fig. 7a, 7b**) when recruited to the *OCT4* promoter using gRNA aptamers/MCP-based targeting. Surprisingly, this potency was not further enhanced by the direct fusion of other TADs to the C-terminus of dCas9 (**Supplementary Fig. 7c**). Collectively, our data show that gRNA aptamer/MCP-based recruitment of the MSN or NMS modules – termed the CRISPR-dCas9 recruited enhanced activation module (DREAM) platform – can efficiently stimulate

transcription without viral components. Our results also demonstrate that natural and engineered human TADs can have non-obvious interactions when combinatorially recruited in bi- and tripartite fashions.

CRISPR-DREAM displays potent activation of endogenous promoters, is specific, and is robust across diverse mammalian cell types

To assess the relative transactivation potential of CRISPR-DREAM, we first targeted the DREAM or SAM[12] systems (**Fig. 1a, 1b**), to different human promoters in HEK293T cells. All components for both the DREAM and SAM systems were well-expressed in HEK293T cells (**Fig. 1c**). At all promoters targeted using pools of gRNAs (n = 15), DREAM was superior or comparable or to the SAM system (**Fig. 1d and Supplementary Fig. 8**). Similarly, when human promoters were targeted using only single gRNAs (n = 11), DREAM remained superior or comparable to the SAM system in all experiments (**Fig. 1e and Supplementary Fig. 9**). Interestingly, this trend extended throughout ~1kb upstream of the transcription start sites (TSSs) surrounding human genes (**Supplementary Fig. 10**). Collectively, these data demonstrate that, although the DREAM system is smaller than the SAM system, and is devoid of viral TADs, it displays superior or comparable transactivation potency in human cells.

To test the transcriptome-wide specificity of CRISPR-DREAM, we used 4 gRNAs to target the DREAM or the SAM system to the *HBG1*/*HBG2* locus in HEK293T cells and then performed RNA-seq (**Fig. 1f**). *HBG1*/*HBG2* gene activation was specific and potent for both the CRISPR-DREAM and SAM systems relative to dCas9 + MCP-mCherry control treated cells. However, DREAM activated substantially more H*BG1*/*HBG2* transcription than the SAM system or dCas9-VPR[8] (**Fig. 1f and Supplementary Fig. 11**). We also found that the DREAM system was significantly ($P < 0.05$) more potent than the SAM system at all targeted genes when each system was combined with a pool of six gRNAs, each targeting a different gene (**Fig. 1g**). Additionally, we evaluated the efficacy of the DREAM system across a battery of different human cell types, including a diverse panel of cancer cell lines (**Fig. 1h and Supplementary Fig. 12**) as well as primary and/or karyotypically normal human cells (**Fig. 1i and Supplementary Fig. 13**). Finally, we tested the transactivation potency of the DREAM system in mammalian cell types widely used for disease modeling/biocompatibility applications and therapeutic production pipelines (NIH3T3 and CHO-K1 cells, respectively; **Supplementary Fig. 14**). Across all experiments the DREAM system displayed highly potent transactivation. Overall, our data demonstrate that CRISPR-DREAM is robust, broadly potent, specific, and functionally compatible with diverse human and mammalian cell types.

CRISPR-DREAM efficiently catalyzes RNA synthesis from noncoding genomic regulatory elements

Since CRISPR-DREAM efficiently and robustly activated mRNAs when targeted to promoter regions, we next tested whether the DREAM system could also activate transcription from distal human regulatory elements (i.e., enhancers) and other non-coding transcripts (i.e., enhancer RNAs; eRNAs, long noncoding RNAs; lncRNAs, and microRNAs; miRNAs). We first targeted the DREAM or SAM systems to the *OCT4* distal enhancer (DE)[31] and found that the DREAM system significantly ($P < 0.05$) upregulated

OCT4 expression relative to the SAM system when targeted to the DE (**Fig. 2a**). Similar results were observed when targeting the DREAM system to the DRR enhancer[32] upstream of the *MYOD* gene (**Supplementary Fig. 15a**). We also targeted the DREAM system to the human HS2 enhancer[33, 34] and observed that the DREAM system induced expression from the downstream *HBE*, *HBG*, and *HBD* genes (**Fig. 2b**). We further observed transactivation of the *SOCS1* gene when the DREAM system was targeted to either of two different intragenic *SOCS1* enhancers; one located ~15kb, and the other ~50kb downstream of the *SOCS1* TSS (**Fig. 2c**). Together these data demonstrate that CRISPR-DREAM can stimulate human gene expression when targeted to different classes of enhancers (those regulating a single-gene, multiple genes, or intragenic enhancers) embedded within native chromatin.

We next tested whether CRISPR-DREAM could activate eRNAs when targeted to endogenous human enhancers. When targeted to the *NET1* enhancer, the DREAM system activated eRNA transcription (**Fig. 2d**), consistent with other reports[35]. Moreover, when the DREAM system was targeted to the bidirectionally transcribed *KLK3* and *TFF1* enhancers, we observed substantial upregulation of eRNAs in both the sense and antisense directions (**Figs. 2e, 2f**). Similar results were obtained when targeting the human *FKBP5* and *GREB1* enhancers (**Supplementary Figs. 15b, 15c**). CRISPR-DREAM also stimulated the production of endogenous lncRNAs when targeted to the *CCAT1, GRASLND*, *HOTAIR*, or *MALAT1* loci (**Figs. 2g and 2h, Supplementary Figs. 15d, 15e**). Finally, we found that the DREAM system activated *miRNA-146a* expression when targeted to the *miRNA-146a* promoter (**Fig. 2i**). Taken together, these data show that CRISPR-DREAM can robustly transactivate regulatory regions spanning diverse classes of the human transcriptome.

Smaller, orthogonal CRISPR-DREAM platforms enable expanded genomic targeting beyond NGG PAM sites

To enhance the versatility of CRISPR-DREAM beyond SpdCas9 and to expand targeting to non-NGG PAM sites, we selected the two smallest naturally occurring orthogonal Cas9 proteins; SadCas9 (1,096aa) and CjdCas9 (1,027aa) for further analyses (**Fig. 3a, 3d**). We used SaCas9-specific gRNAs harboring MS2 loops[36] to compare the potency between the SadCas9-DREAM and SAM systems in HEK293T cells. SadCas9-DREAM was significantly ($P < 0.05$) more potent than SadCas9-SAM when targeted to either the *HBG1* or *TTN* promoters (**Fig. 3b**). We also found that SadCas9-DREAM outperformed or was comparable to SadCas9-VPR when targeted to these loci (**Fig. 3c**). CjdCas9-based transcriptional activation platforms have also recently been developed using viral TADs (miniCAFE)[37]; however, gRNA-based recruitment of transcriptional modulators using CjdCas9 has not been described. Therefore, we engineered the CjCas9 gRNA scaffold to incorporate an MS2 loop within the tetraloop of the CjCas9 gRNA scaffold (**Supplementary Fig. 16c**). We used this MS2-modified CjCas9 gRNA to generate CjdCas9-DREAM and compared the potency between CjdCas9-DREAM, CjdCas9-SAM, and the miniCAFE systems at the *HBG1* or *TTN* promoters (**Figs. 3e, 3f**) in HEK293T cells. At all targeted sites, CjdCas9-DREAM outperformed or was comparable to the CjdCas9-SAM or miniCAFE systems. We also observed high levels of transactivation using SadCas9-DREAM and CjdCas9-DREAM in a different human cell line

(**Supplementary Fig. 16a, 16b, 16d, and 16e**). These data demonstrate that DREAM is not only compatible with other orthogonal dCas9 targeting systems, but that it displays superior performance at most tested promoters.

Generation and validation of a compact mini-DREAM system

We next sought to reduce the sizes of the CRISPR-DREAM components. We first investigated whether individual TADs could be minimized while still retaining the transactivation potency when recruited by dCas9. We focused on individual TADs from MTFs that displayed transactivation potential (i.e., MRTF-A, MRTF-B, and MYCOD proteins, **Supplementary Figs. 1, 2**). 9aa TADs have been shown to synthetically activate transcription previously using GAL4 systems[38, 39]. Therefore, we used predictive software[38] to identify 9aa TADs in MRTF-A, MRTF-B, and MYCOD proteins, and recruited these TADs to human loci using dCas9 and MCP-MS2 fusions in single, bipartite, and tripartite formats (**Supplementary Note 1; Supplementary Fig. 17**). Interestingly, we observed that only tripartite combinations of 9aa TADs were able to robustly activate endogenous gene expression, and to varying degrees (**Supplementary Fig. 17f**). We selected one tripartite 9aa combination (3x 9aa TAD; MRTF-B.3 + MYOCD.1 + MYOCD.3) for further analysis (**Fig. 3g**). This 3x 9aa TAD activated *HBG1*, *TTN*, and *CD34* gene expression when recruited to corresponding promoters using dCas9 (**Fig 3h; Supplementary Fig. 17g**). We also found that this 3x 9aa TAD combination could activate gene expression via a single gRNA, and moreover could transactivate other endogenous regulatory loci (**Supplementary Fig. 17h-j**). These results suggest that combinations of 9aa TADs can be used as minimal functional units to transactivate endogenous human loci when recruited via dCas9.

We next combined the 3x 9aa TAD with the engineered NRF2 TAD (eNRF2) in four different combinations to generate a small, yet potent transactivation module called eN3x9 (**Supplementary Fig. 18**). Notably, minimized Cas9 proteins that retain DNA binding activity have also been recently created[40, 41]. Therefore, we next evaluated the relative transactivation capabilities among a panel of minimized, HNH-deleted, dCas9 variants in tandem with MCP-MSN and found that an HNH-deleted variant without a linker between two RuvC domains was optimal, albeit with slight protein expression decreases (**Supplementary Fig. 19a, 19b**). We further validated this linker-less, HNH-deleted CRISPR-DREAM variant at multiple human promoters and other regulatory elements (**Supplementary Fig. 19c-h**) and then combined this minimized dCas9 with MCP-eN3x9 to generate the mini-DREAM system (**Fig. 3i**). The mini-DREAM system transactivated *HBG1*, *TTN*, and *IL1RN* gene expression when recruited to corresponding promoters (**Fig 3j; Supplementary Fig. 20a**). We also found that the mini-DREAM system could activate endogenous promoters via a single gRNA (**Supplementary Figs. 20b, c**), and could activate downstream gene expression when targeted to an upstream enhancer (**Supplementary Fig. 20d**). Finally, we evaluated whether the minimized components of the mini-DREAM system were functional when delivered within a single vector (**Fig. 3k**) and found that this compact, single vector mini-DREAM system retained transactivation potential when targeted to human promoters using pooled (**Fig 3l; Supplementary Fig. 20e-g**), or a single gRNA (**Supplementary Fig. 20h**). Overall, these data show that the

components of the CRISPR DREAM system can be minimized to fit within a single vector delivery framework while retaining functionality.

The MSN and NMS effector domains are robust across programmable DNA binding platforms

We next tested the potency of tripartite MSN and NMS effectors when fused the to dCas9 in different architectures and observed that both effectors could activate gene expression when fused to the N- or C-terminus of dCas9 (**Supplementary Note 2; Supplementary Fig. 21**) or when recruited via the Sun-Tag[11] architecture (**Supplementary Fig. 22**). Interestingly, in contrast to MCP-mediated recruitment (**Supplementary Fig. 7**), additional TADs were observed to improve performance in direct fusion architectures (**Supplementary Figs. 21a, c; Supplementary Note 2**). In the SunTag architecture, the NMS domain was superior to other benchmarked effector domains, such as VP64[11], VPR[42], and p65-HSF1[43] (**Supplementary Figs. 22a – c**). To maximize the potential use of the MSN/NMS effector domains and explore their versatility, we next tested whether each was capable of gene activation when fused to TALE or ZF scaffolds (**Figs. 4a, 4d**). Both effectors strongly transactivated *IL1RN* using a single TALE fusion protein (**Supplementary Fig. 23**) or a pool of 4 TALE fusion proteins targeted to the *IL1RN* promoter (**Fig 4a**). Similarly, both effectors activated *ICAM1* expression using a single synthetic ZF fusion protein targeted to the *ICAM1* promoter (**Fig. 4b**). These data demonstrate that the MSN and NMS effectors are compatible with diverse programmable DNA binding scaffolds beyond Type II CRISPR/Cas systems.

Transcriptional activators have recently been shown to modulate the expression of endogenous human loci when recruited by Type I CRISPR systems[44]. Therefore, to evaluate whether MSN and/or NMS were functional beyond Type II CRISPR systems, we fused each to the Cas6 component of the *E. coli* Type I CRISPR Cascade (Eco-Cascade) system (**Fig. 4c**). Our data showed that Cas6-MSN (or NMS) performed comparably to the Cas6-p300 system when targeted to a spectrum of human promoters (**Fig. 4d; Supplementary Fig. 24a – d**). We also observed that the Cas6-MSN (or NMS) systems could activate eRNAs from when targeted to the endogenous *NET1* enhancer (**Supplementary Figs. 24e**). One advantage of CRISPR Cascade is that the system can process its own crRNA arrays, which can enable multiplexed targeting to the human genome. Previous reports have leveraged this capability to simultaneously activate two human genes[44]. We found that when Cas6 was fused to MSN, the CRISPR Cascade system could simultaneously activate up to six human genes when corresponding crRNAs were co-delivered in an arrayed format (**Fig. 4e; Supplementary Fig. 24f**). We also found that these transactivation capabilities were extensible to another Type I CRISPR system; *Pae*-Cascade[45] (**Supplementary Fig. 24g – i**). In sum, these data show that the MSN and NMS effectors are robust and directly compatible with programmable DNA binding platforms beyond Type II CRISPR systems without any additional engineering.

The NMS effector enables superior multiplexed gene activation when fused to dCas12a

The CRISPR/Cas12a system has attracted significant attention because the platform is smaller than SpCas9, and because Cas12a can process its own crRNA arrays in human cells[46]. This feature has been leveraged for both multiplexed genome editing and multiplexed transcriptional control[15]. Therefore, we next investigated the potency of the tripartite MSN and NMS effectors when they were directly fused to dCas12a (**Fig. 4f**). We selected the AsdCas12a variant for this analysis because AsdCas12a (hereafter dCas12a) has been shown to activate human genes when fused to transcriptional effectors[15]. Our results demonstrated that both dCas12a-MSN and dCas12a-NMS were able to induce gene expression when targeted to different human promoters using pooled or single crRNAs (**Fig. 4g, 4h, Supplementary Fig. 25a – 25e**). dCas12a-NMS was generally superior to dCas12a-MSN and to the previously described dCas12a-Activ system[15] at the loci tested here. These data demonstrate that the NMS and MSN effectors domains are potent transactivation modules when combined with the dCas12a targeting system in human cells.

We next tested the extent to which dCas12a-MSN/NMS could be used in conjunction with crRNA arrays for multiplexed endogenous gene activation. We cloned 8 previously described crRNAs[15] (targeting the *ASCL1*, *IL1R2*, *IL1B* or *ZFP42* promoters) into a single plasmid in an array format and then transfected this vector into HEK293T cells with either dCas12a control, dCas12a-MSN, dCas12a-NMS, or the dCas12a-Activ system. Again, our data demonstrated that dCas12a-NMS was superior or comparable to dCas12a-Activ, even in multiplex settings (**Supplementary Fig. 25f**). Finally, to evaluate if dCas12-NMS could simultaneously activate multiple genes on a larger scale, we cloned 20 full-length (20bp) crRNAs targeting 16 different loci into a single array (**Supplementary Fig. 25g**). This array was designed to enable simultaneous targeting of several classes of human regulatory elements; including 13 different promoters, 2 different enhancers (one intrageneric; *SOCS1,* and one driving eRNA output; *NET1*), and one lncRNA (*GRASLND*). When this crRNA array was transfected into HEK293T cells along with dCas12a-NMS, RNA synthesis was robustly stimulated from all 16 loci (**Fig. 4i**). To our knowledge this is the most loci that have been targeted simultaneously using CRISPR systems, demonstrating the versatility and utility of the engineered NMS effector in combination with dCas12a.

dCas9-NMS permits efficient reprogramming of human fibroblasts *in vitro*

CRISPRa systems using repeated portions of the alpha herpesvirus VP16 TAD (dCas9-VP192) have been used to efficiently reprogram human foreskin fibroblasts (HFFs) into induced pluripotent stem cells (iPSCs)[14]. To evaluate the functional capabilities of our engineered human transactivation modules, we fused the NMS domain directly to the C-terminus of dCas9 (dCas9-NMS) and tested its ability to reprogram HFFs. We used a direct dCas9 fusion architecture so that we could leverage gRNAs previously optimized for this reprogramming strategy and to better compare dCas9-NMS with the corresponding state of the art (dCas9-VP192)[14]. We used the NMS effector as opposed to MSN, as NMS displayed more potency than MSN when directly fused to dCas9 (**Supplementary Fig. 21a**). We targeted dCas9-NMS (or dCas9-VP192) to endogenous loci using the 15 gRNAs previously optimized to reprogram HFFs to pluripotency with the dCas9-VP192 system. Using this approach, we observed morphological changes

beginning by 8 days post-nucleofection (**Fig. 5a**) and efficient reprogramming by 16 days post-nucleofection, although to a lesser extent than when using dCas9-VP192 (**Supplementary Fig. 26a**).

We picked and expanded iPSC colonies and then measured the expression of pluripotency and mesenchymal genes ~40 days post-nucleofection. We found that genes typically associated with pluripotency (*OCT4, SOX2, NANOG, LIN28A, REX1, CDH1,* and *FGF4*)[47, 48] were highly expressed in colonies derived from HFFs nucleofected with the gRNA cocktail and dCas9-NMS or dCas-VP192 (**Fig. 5b; Supplementary Fig. 26b-f**). Conversely, we observed that genes typically associated with fibroblast/mesenchymal cell identity (*THY1, ZEB1, ZEB2, TWIST,* and *SNAIL2*)[47, 48] were poorly expressed in colonies derived from HFFs nucleofected with the gRNA cocktail and dCas9-NMS or dCas-VP192 (**Fig. 5c; Supplementary Fig. 26g-i**). Finally, we assessed the expression of pluripotency associated markers (SSEA-4, TRA-1-81 and TRA-1-60)[49] and found that all were highly expressed in iPSC colonies derived from HFFs nucleofected with the gRNA cocktail and either dCas9-NMS or dCas-VP192 (**Fig. 5d, 5e; Supplementary Fig. 26j**). These data show that engineered transactivation modules sourced from human MTFs can be used to efficiently reprogram complex cell phenotypes, including cell lineage.

Discussion

Here we harnessed the programmability and versatility of different dCas9-based recruitment architectures (direct fusion, gRNA-aptamer, and SunTag-based) to define and optimize the transcriptional output of TADs derived from natural human TFs. We leveraged these insights to build superior and widely applicable transactivation modules that are portable across all modern synthetic DNA binding platforms, and that can activate the expression of diverse classes endogenous RNAs. We selected mechanosensitive TFs (MTFs) for molecular building blocks because they naturally display rapid and potent gene activation at target loci, can interact with diverse transcriptional co-factors across different human cell types, and because their corresponding TADs are relatively small[50-52]. We not only identified and validated the transactivation potential of TADs sourced from individual MTFs, but we also established the optimal TAD sequence compositions and combinations for use across different synthetic DNA binding platforms, including Type I, II and V CRISPR systems, TALE proteins, and ZF proteins.

Our study also revealed that for MTFs, tripartite fusions using TADs from MRTA-A (M), STAT1 (S), and NRF2 (N) in one of two different combinations (either MSN or NMS) consistently resulted in the most potent human gene activation across different DNA binding platforms. Interestingly, each of these components has been shown to interact with key transcriptional co-factors. For example, individual TADs from MRTF-A, STAT1, NRF2 can directly interact with endogenous p300[24, 53]. Moreover, the Neh4 and Neh5 TADs from NRF2 can also cooperatively recruit endogenous CBP for transcriptional activity[22, 54]. Therefore, we suspect that the potency of the MSN and NMS tripartite effector proteins is likely related to their robust capacity to recruit the powerful and

ubiquitous endogenous transcriptional modulators p300 and/or CBP, which is likely positively impacted by their direct tripartite fusion.

Additionally, our study demonstrated that the superior transactivation capabilities of the CRISPR/dCas9-recruited enhanced activation module (DREAM) system – consisting of dCas9 and a gRNA-aptamer recruited MCP-MSN fusion – are not reliant upon the direct fusion(s) of any other proteins (viral or otherwise) to dCas9, in contrast to the SAM system which relies upon dCas9-VP64[12]. We used this advantage to combine the MCP-MSN module with HNH domain deleted dCas9 variants[40, 41], which exhibited similar potencies to full-size dCas9 variants. To further reduce the size of CRISPR-DREAM, we built a minimal transactivation module (eN3x9; 96aa) by evaluating the potency of a suite of 9aa TADs and by then combining the most potent variants with the small eNRF2 TAD. We then combined the minimized eN3x9 transactivation module with an HNH domain deleted dCas9 variant in two-vector (mini-DREAM) and single-vector (mini-DREAM compact) delivery architectures, which retained potent transactivation capabilities.

We also integrated the MSN and NMS effectors with the Type I CRISPR/Cascade and Type II dCas12a platforms to enable superior multiplexed endogenous activation of human genes. This superior multiplexing capability holds tremendous promise for reshaping endogenous cellular pathways and/or engineering complex transcriptional networks. dCas9-based transcription factors harboring viral TADs have also been used for directed differentiation and cellular reprogramming[8, 14, 55, 56]. Here, we showed that we could reprogram human fibroblasts into iPSCs using dCas9 directly fused to the NMS transcriptional effector with similar gene expression profiles, times to conversion, and morphological characteristics compared to iPSCs derived using dCas9 fused to viral TADs[14]. However, dCas9-NMS resulted in slightly fewer iPSC colonies than dCas9-VP192, which we attribute to the reprogramming framework tested here being optimized for use with dCas9-VP192.

In summary, we have used the rational redesign of natural human TADs to build synthetic transactivation modules that enable consistent and potent performance across programmable DNA binding platforms, mammalian cell types, and genomic regulatory loci embedded within human chromatin. Although we used MTFs as molecular building blocks here, our work establishes a framework that could be used with practically any natural or engineered TF and/or chromatin modifier in future efforts. We anticipate that the potency, small size, versatility, capacity for multiplexing, and the lack of viral components associated with the newly engineered TADs (MSN, NMS, and eN3x9) and CRISPR-DREAM systems developed here will be valuable tools for fundamental and biomedical applications requiring potent and predictable activation of endogenous transcription.

Methods

Cell Culture

All experiments were performed within 10 passages of cell thaws. HEK293T (ATCC, CRL-11268), HeLa (ATCC, CCL-2), A549 (ATCC, CCL-185), SK-BR-3 (ATCC, HTB-30), U2OS (ATCC, HTB-96), HCT116 (ATCC, CRL-247), K562 (ATCC, CRL-243), CHO-K1 (ATCC, CCL-61), ARPE-19 (ATCC, CRL-2302), HFF (ATCC, CRL-2429), Jurkat-T (ATCC, TIB-152), and hTERT-MSC (ATCC, SCRC-4000) 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 (100 units/ mL penicillin, 100 µg / mL streptomycin; Gibco) at 37° C and 5% CO_2. NIH3T3 cells were a kind gift from Dr. Caleb Bashor's lab and were cultured in DMEM supplemented with 10% FBS (Sigma-Aldrich) and 1% pen/strep (100 units/ mL penicillin, 100 µg/ mL streptomycin) at 37° C and 5% CO_2.

Plasmid Transfection and Nucleofection

HEK293T cell transfections were performed in 24-well plates using 375ng of dCas9 expression plasmid and 125ng of equimolar pooled or individual gRNAs/crRNAs. 1.25×10^5 HEK293T cells were plated the day before transfection and then transfected using Lipofectamine 3000 (Invitrogen, USA) as per manufacturer's instruction. For two component systems (dCas9 + MCP or dCas9 + scFv systems) 187.5ng of each plasmid was used. For multiplex gene activation experiments using DREAM, 25ng of each gRNA encoding plasmid targeting each respective gene was used. Transfections in HeLa, A549, SK-BR-3, U2OS, HCT-116, HFF, NIH3T3, and CHO-K1 were performed in 12-well plates using Lipofectamine 3000 and 375ng dCas9 plasmid, 375ng of MCP-effector fusion proteins, and 250ng DNA of MS2-modifed gRNA encoding plasmid. For transfections using dCas12a fusion proteins where single genes were targeted, 375ng of dCas12a-effector fusion plasmids and 125ng of crRNA plasmids were transfected using lipofectamine 3000 per manufacturer's instruction. For multiplex gene activation experiments using dCas12a, 375ng of dCas12a-effector fusion encoding plasmid and 250ng of multiplex crRNA expression plasmids were used. For experiments using *E. coli* and *P. aeruginosa* Type I CRISPR systems, we followed the same stoichiometries used in previous studies[44, 45]. For transfection of *ICAM1*-ZF effectors, 500ng of each *ICAM1* targeting ZF fusion was transfected. Transfections using *IL1RN*-TALE fusion proteins were performed using 500ng of either single TALE or a pool of 4 TALEs using 125ng of each TALE fusion. All ZF and TALE transfections were performed in HEK293T cells in 24-well format using Lipofectamine 3000 as per manufacturers instruction. For K562 cells, 1×10^6 cells were nucleofected using the Lonza SF Cell Line 4D-Nucleofector Kit (Lonza V4XC-2012) and a Lonza 4D Nucleofector (Lonza, AAF1002X) using the FF-120 program. 2000ng of total plasmids were nucleofected in each condition using 1×10^6 K562 cells and 667ng each of; dCas9 plasmid, MCP fusion plasmid, and pooled MS2-sgRNA expression plasmid was nucleofected per condition. Immediately after nucleofection, K562 cells were transferred to prewarmed media containing 6-well plates. hTERT-MSCs were electroporated with using the Neon transfection system (Thermo Fisher Scientific) using the 100µL kit. 5×10^5 hTERT-MSCs were resuspended in 100µL resuspension buffer R and 10µg total DNA (3.75µg dCas9, 3.75µg MCP-fusion effector plasmid, and 2.5µg

MS2-modifed gRNA encoding plasmid). Electroporation was performed using the settings recommended by the manufacturers for mesenchymal stem cells: Voltage: 990V, Pulse width: 40ms, Pulse number: 1. For fibroblast reprogramming experiments, we used the Neon transfection system using the amounts of endotoxin free DNA described previously[14] and below.

PBMC isolation, Culture, and Nucleofection

De-identified white blood cell concentrates (buffy coats) were obtained from the Gulf Coast Regional Blood Center in Houston, Texas. PBMCs were isolated from buffy coats using Ficoll gradient separation and cryopreserved in liquid nitrogen until later use. $1x10^6$ PBMCs per well were stimulated for 48h in a CD3 (Tonbo Biosciences, 700037U100) /CD28 (Tonbo Biosciences, 70289U100)-coated 24-well plate containing RPMI media supplemented with 10% FBS (Sigma-Aldrich), 1% Pen/Strep (Gibco), 10ng/mL IL-15 (Tonbo Biosciences, 218157U002), and 10ng/mL IL-7 (Tonbo Biosciences, 218079U002). Stimulated PBMCs were electroporated using the Neon transfection system (Thermo Fisher Scientific) 100µL kit per manufacturer protocol. Briefly, PBMCs were centrifuged at 300g for 5min and resuspended in Neon Resuspension Buffer T to a final density of $1x10^7$ cells/mL. 100µL of the resuspended cells ($1x10^6$ cells) were then mixed with 12µg total plasmid DNA (4.5µg of dCas9 fusion encoding plasmids, 4.5µg of MCP fusion encoding plasmids, and 3µg of four equimolar pooled MS2-modifed gRNA encoding plasmids) and electroporated with the following program specifications using a 100µL Neon Tip: pulse voltage 2,150v, pulse width 20ms, pulse number 1. Endotoxin free plasmids were used in all experiments. After electroporation, PBMCs were incubated in prewarmed 6-well plates containing RPMI media supplemented with 10% FBS (Sigma-Aldrich), 1% Pen/Strep (Gibco), 10ng/mL IL-15, and 10ng/mL IL-7. PBMCs were maintained at 37°C, 5% CO_2 for 48h before RNA isolation and QPCR.

Plasmid Cloning

Lenti-dCas9-VP64 (Addgene #61425), dCas9-VPR (Addgene #63798), dCas9-p300 (Addgene #83889), MCP-p65-HSF1 (Addgene #61423), scFv-VP64 (Addgene #60904), SpgRNA expression plasmid (Addgene #47108), MS2-modified gRNA expression plasmid (Addgene #61424), AsCas12a (Addgene #128136), *E. Coli* Type I Cascade system (Addgene #106270-106275) and *Pae* Type I Cascade System (Addgene #153942 and 153943), YAP-S5A (Addgene #33093) have been described previously. The eNRF2 TAD fusion was synthetically designed and ordered as a gBlock from IDT. To generate an isogenic C-terminal effector domain cloning backbone, the dCas9-p300 plasmid (Addgene #83889) was digested with BamHI and then a synthetic double-stranded ultramer (IDT) was incorporated using NEBuilder HiFi DNA Assembly (NEB, E2621) to generate a dCas9-NLS-linker-BamHI-NLS-FLAG expressing plasmid. This plasmid was further digested with AfeI and then a synthetic double-stranded ultramer (IDT) was incorporated using NEBuilder HiFi DNA Assembly to generate a FLAG-NLS-MCS-linker-dCas9 expressing Plasmid for N-terminal effector domain cloning. For fusion of effector domains to MCP, the MCP-p65-HSF1 plasmid (Addgene #61423) was digested with BamHI and NheI and respective effector domains were cloned using NEBuilder HiFi DNA Assembly. For SunTag components, the scFv-GCN4-linker-VP16-GB1-Rex NLS sequence was PCR amplified from pHRdSV40-scFv-GCN4-sfGFP-VP64-GB1-NLS (Addgene #60904) and cloned into a lentiviral backbone containing an EF1-alpha promoter.

Then VP64 domain was removed and an AfeI restriction site was generated and used for cloning TADs using NEBuilder HiFi DNA Assembly. The pHRdSV40-dCas9-10xGCN4_v4-P2A-BFP (Addgene #60903) vector was used for dCas9-based scFv fusion protein recruitment to target loci. All MTF TADs were isolated using PCR amplified from a pooled cDNA library from HEK293T, HeLa, U2OS and Jurkat-T cells. TADs were cloned into the MCP, dCas9 C-terminus, dCas9 N-terminus, and scFv backbones described above using NEBuilder HiFi DNA Assembly. Bipartite N-terminal fusions between MCP-MRTF-A or MCP-MRTF-B TADs and STAT 1-6 TADs were generated by digesting the appropriate MCP-fusion plasmid (MCP-MRTF-A or MCP-MRTF-B) with BamHI and then subcloning PCR-amplified STAT 1-6 TADs using NEBuilder HiFi DNA Assembly. Bipartite C-terminal fusions between MCP-MRTF-A or MCP-MRTF-B TADs and STAT 1-6 TADs were generated by digesting the appropriate MCP-fusion plasmid (MCP-MRTF-A or MCP-MRTF-B) with NheI and then subcloning PCR-amplified STAT 1-6 TADs using NEBuilder HiFi DNA Assembly. Similarly, eNRF2 was fused to the N- or C-terminus of the bipartite MRTF-A-STAT1 TAD in the MCP-fusion backbone using either BamHI (N-terminal; MCP-eNRF2-MRTF-A-STAT1 TAD) or NheI (C-terminal; MCP-MRTF-A-STAT1-eNRF2 TAD) digestion and NEBuilder HiFi DNA Assembly to generate the MCP-NMS or MCP-MSN tripartite TAD fusions, respectively. SadCas9 (with D10A and N580A mutations derived using PCR) was PCR amplified and then cloned into the SpdCas9 expression plasmid backbone created in this study digested with BamHI and XbaI. This SadCas9 expression plasmid was digested with BamHI and then PCR-amplified VP64 or VPR TADs were cloned in using NEBuilder HiFi DNA Assembly. CjCas9 was PCR-amplified from pAAV-EFS-CjCas9-eGFP-HIF1a (Addgene #137929) as two overlapping fragments using primers to create D8A and H559A mutations. These two CjdCas9 PCR fragments were then cloned into the SpdCas9 expression plasmid digested with BamHI and XbaI using NEBuilder HiFi DNA Assembly. This CjdCas9 expression plasmid was digested with BamHI and the PCR-amplified VP64 or VPR TADs were cloned in using NEBuilder HiFi DNA Assembly. HNH domain deleted SpdCas9 plasmids were generated using different primer sets designed to amplify the N-terminal and C-terminal portions of dCas9 excluding the HNH domain and resulting in either: no linker, a glycine-serine linker, or an XTEN16 linker, between HNH-deleted SpdCas9 fragments. These different PCR-amplified regions were cloned into the SpdCas9 expression plasmid digested with BamHI and XbaI using NEBuilder HiFi DNA Assembly. MCP-mCherry, MCP-MSN and MCP-p65-HSF1 were digested with NheI and a single strand oligonucleotide encoding the FLAG sequence was cloned onto the C-terminus of each respective fusion protein using NEBuilder HiFi DNA Assembly to enable facile detection via Western blotting. 1x 9aa TADs were designed and annealed as double strand oligos and then cloned into the BamHI/NheI-digested MCP-p65-HSF1 backbone plasmid (Addgene #61423) using T4 ligase (NEB). Heterotypic 2x 9aa TADs were generated by digesting MCP-1x 9aa TAD plasmids with either BamHI or NheI and then cloning single strand DNA encoding 1x 9aa TADs to the N- or C-termini using NEBuilder HiFi DNA Assembly. Heterotypic MCP-3x 9aa TADs were generated similarly by digesting MCP-2x 9aa TAD containing plasmids either with BamHI or NheI and then single strand DNA encoding 1x 9aa TADs were cloned to the N- or C-termini using NEBuilder HiFi DNA Assembly. Selected fusions between 3x 9aa TADs and eNRF2 were generated using gBlock (IDT) fragments and cloned into the BamHI/NheI-digested MCP-p65-HSF1 backbone plasmid (Addgene #61423) using NEBuilder HiFi DNA Assembly. To generate mini-DREAM compact single plasmid

system, SpdCas9-HNH (no linker) deleted plasmid was digested with BamHI and then PCR amplified P2A self-cleaving sequence and MCP-eNRF2-3x 9aa TAD (eN3x9) was cloned using NEBuilder HiFi DNA Assembly. For dCas12a fusion proteins, SiT-Cas12a-Activ (Addgene #128136) was used. First, we generated a nuclease dead (E993A) SiT-Cas12a backbone using PCR amplification and we used this plasmid for subsequent C-terminal effector cloning using BamHI digestion and NEBuilder HiFi DNA Assembly. For *E. coli* Type I CRISPR systems, the Cas6-p300 plasmid (Addgene #106275) was digested with BamHI and then MSN and NMS domains were cloned in using NEBuilder HiFi DNA Assembly. *Pae* Type I Cascade plasmids encoding Csy1-Csy2 (Addgene #153942) and Csy3-VPR-Csy4 (Addgene #153943) were obtained from Addgene. The Csy3-VPR-Csy4 plasmid was digested with MluI (NEB) and BamHI (to remove the VPR TAD) and then the nucleoplasmin NLS followed by a linker sequence was added using NEBuilder HiFi DNA Assembly. Next, this Csy3-Csy4 plasmid was digested with AscI and either the MSN or NMS TADs were cloned onto the N-terminus of Csy3 NEBuilder HiFi DNA Assembly. ZF fusion proteins were generated by cloning PCR-amplified MSN, NMS, or VPR domains into the BsiWI and AscI digested *ICAM1* targeting ZF-p300 plasmid[9] using NEBuilder HiFi DNA Assembly. Similarly, TALE fusion proteins were created by cloning PCR-amplified MSN, NMS, or VPR domains into the BsiwI and AscI digested *IL1RN* targeting TALE plasmid backbone[9] using NEBuilder HiFi DNA Assembly. pCXLE-dCas9VP192-T2A-EGFP-shP53 (Addgene #69535), GG-EBNA-OSK2M2L1-PP (Addgene #102898) and GG-EBNA-EEA-5guides-PGK-Puro (Addgene #102898) used for reprogramming experiments have been described previously[14, 57]. The PCR-amplified NMS domain was cloned into the sequentially digested (XhoI then SgrDI; to remove the VP192 domain) pCXLE-dCas9VP192-T2A-EGFP-shP53 backbone using NEBuilder HiFi DNA Assembly. TADs were directly fused to the C-terminus of dCas9 by digesting the dCas9-NLS-linker-BamHI-NLS-FLAG plasmid with BamHI and then cloning in PCR-amplified TADs using NEBuilder HiFi DNA Assembly. TADs were directly fused to the N-terminus to dCas9 by digesting the FLAG-NLS-MCS-linker-dCas9 plasmid with AgeI (NEB) and then cloning in PCR-amplified TADs using NEBuilder HiFi DNA Assembly. For constructs harboring both N- and C-terminal fusions, respective plasmids with TADs fused to the C-terminus of dCas9 were digested with AgeI and then PCR-amplified TADs were cloned onto the N-terminus of dCas9 using NEBuilder HiFi DNA Assembly.

gRNA Design and Construction

All protospacer sequences for SpCas9 systems were designed using the Custom Alt-R® CRISPR-Cas9 guide RNA design tool (IDT). All gRNA protospacers were then phosphorylated, annealed, and cloned into chimeric U6 promoter containing sgRNA cloning plasmid (Addgene #47108) and/or an MS2 loop containing plasmid backbone (Addgene #61424) digested with Bbs1 and treated with alkaline phosphatase (Thermo) using T4 DNA ligase (NEB). The SaCas9 gRNA expression plasmid (pIBH072) was a kind gift from Charles Gersbach and was digested with BbsI or BpiI (NEB or Thermo, respectively) and treated with alkaline phosphatase and then annealed protospacer sequences were cloned in using T4 DNA ligase (NEB). gRNAs were cloned into the pU6-Cj-sgRNA expression plasmid (Addgene #89753) by digesting the vector backbone with BsmBI or Esp3I (NEB or Thermo, respectively), and then treating the digested plasmid with alkaline phosphatase, annealing phosphorylated gRNAs, and then cloning annealed

gRNAs into the backbone using T4 DNA ligase. MS2-stem loop containing plasmids for SaCas9 and CjCas9 were designed as gBlocks (IDT) with an MS2-stem loop incorporated into the tetraloop region for both respective gRNA tracr sequences. crRNA expression plasmids for the Type I *Eco* Cascade system were generated by annealing synthetic DNA ultramers (IDT) containing direct repeats (DRs) and cloning these ultramers into the BbsI and SacI-digested SpCas9 sgRNA cloning plasmid (Addgene #47108) using NEBuilder HiFi DNA Assembly. crRNA expression plasmids for *Pae* Type I Cascade system were generated by annealing and then PCR-extending overlapping oligos (that also harbored a BsmBI or Esp3I cut site for facile crRNA array incorporation) into the sequentially BbsI (or BpiI) and SacI-digested SpCas9 sgRNA cloning plasmid (Addgene #47108) using NEBuilder HiFi DNA Assembly. crRNA expression plasmids for Cas12a systems were generated by annealing and then PCR-extending overlapping oligos (that also harbored a BsmBI or Esp3I cut site for facile crRNA array incorporation) into the sequentially BbsI (or BpiI) and SacI-digested SpCas9 sgRNA cloning plasmid (Addgene #47108) using NEBuilder HiFi DNA Assembly.

crRNA Array Cloning

crRNA arrays for AsCas12a and Type I CRISPR systems were designed in fragments as overlapping ssDNA oligos (IDT) and 2-4 oligo pairs were annealed. Oligos were designed with an Esp3I cut site at 3' of the array for subsequent cloning steps. Equimolar amounts of oligos were mixed, phosphorylated, and annealed similar to the standardized gRNA/crRNA assembly protocol above. Phosphorylated and annealed arrays were then cloned into the respective Esp3I-digested and alkaline phosphatase treated crRNA cloning backbone (described above) using T4 DNA ligase (NEB). crRNA arrays were verified by Sanger sequencing. Correctly assembled 4-8 crRNA array expressing plasmids were then digested again with Esp3I and alkaline phosphatase treated to enable incorporation of subsequent arrays up to 20 crRNAs.

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).15-30µg of lysate were separated using precast 7.5% or 10% 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; Diagenode #C15200216, Anti-FLAG; Sigma-Aldrich #F1804, anti-β-Tubulin; Bio-Rad #12004166). Then membranes were washed with 1X TBST 3 times (10mins each wash) and incubated with respective HRP-tagged secondary antibodies for 1hr. Next membranes were washed with 1X TBST 3 times (10mins 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.

Quantitative Reverse-transcriptase PCR (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 mastermix (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. For CHO-K1 cells, *Gnb1* was used for normalization. Undetectable samples were assigned a Ct value of 45 cycles.

Mature miRNA isolation and QPCR for miRNAs

Mature miRNA (miRNA) was isolated using the miRNA isolation kit (Qiagen #217084). 500ng of isolated miRNA was polyadenylated using poly A polymerase (Quantabio #95107) in 10µL reactions per sample and then used for cDNA synthesis using qScript Reverse Transcriptase and oligo-dT primers attached to unique adapter sequences to allow specific amplification of mature miRNA using QPCR in a total 20µL reaction (Quantabio #95107). cDNA was diluted and 10ng of miRNA cDNA was used for QPCR in a 25µL reaction volume. PerfeCTa SYBR Green SuperMix (Quantabio #95053), *miR-146a* specific forward primer, and PerfeCTa universal reverse primer was used to perform QPCR. U6 snRNA was used for normalization.

Immunofluorescence Microscopy

Human foreskin fibroblasts (HFFs; CRL-2429, ATCC) and HFF-derived iPSCs were grown in Geltrex (Gibco, A1413302) coated 12-well plates and were fixed with 3.7% formaldehyde and then blocked with 3% BSA in 1X PBS for 1hr at Room Temperature prior to imaging. Primary antibodies for SSEA-4 (CST #43782), TRA1-60 (CST #61220) and TRA1-81 (CST #83321) were diluted in 1% BSA in 1X PBS and incubated overnight at 4°C. The next day, cells were washed with 1X PBS, incubated with appropriate Alexaflour-488 conjugated secondary antibodies for 1hr at Room Temperature and then washed again with 1X PBS. Cells were then incubated with DAPI (Invitrogen #D1306) containing PBS for 10m, washed with 1X PBS, and then imaged using a Nikon ECLIPSE Ti2 fluorescent microscope.

Fibroblast Reprogramming

HFFs were cultured in 1X DMEM supplemented with 1X Glutamax (Gibco, 35050061) for two passages before transfection with respective components. Cells were grown in 15cm dishes (Corning), and detached using TrypLE select (Gibco, #12563011). Single cell suspensions were washed with complete media and then with 1X PBS. For each 1 x 10^6 cells, a total of 6 µg of endotoxin free plasmids (Macherey-Nagel, 740424; 2 µg CRISPR activator plasmid, 2 µg of pluripotency factor targeting gRNA plasmid, and 2 µg of EEA-motif targeting gRNA expression plasmids) were nucleofected using a 100µL Neon transfection tip in R buffer using the following settings: 1650V, 10ms, and 3 pulses. Nucleofected fibroblasts were then immediately transferred to Geltrex (Gibco) coated 10cm cell culture dishes in prewarmed media. The next day media was exchanged. 4 days later, media was replaced with iPSC induction media[14]. Induction media was then

exchanged every other day for 18 days. After 18 days iPSC colonies were counted, and colonies picked using sterile forceps and then transferred to Geltrex coated 12-well plates. iPSC colonies were maintained in complete E8 media and passaged as necessary using ReLeSR passaging reagent (Stem Cell Technology, #05872). RNA was isolated from iPSC clones using the RNeasy Plus mini kit (Qiagen # 74136) and colonies were immunostained using indicated antibodies and counterstained with DAPI (Invitrogen) for nuclear visualization.

RNA Sequencing (RNA-seq)

RNA-seq was performed in duplicate for each experimental condition. 72hrs post-transfection RNA was isolated using the RNeasy Plus mini kit (Qiagen). RNA integrity was first assessed using a Bioanalyzer 2200 (Agilent) and then 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 concentration 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 prior to sequencing. Sequencing was performed using an Illumina Hiseq 3000 with paired end 75 base pair reads. Reads were aligned to the human genome (hg38) Gencode Release 36 reference using STAR aligner (v2.7.3a). Transcript levels were quantified to the reference genome using a Bayesian approach. Normalization was done using counts per million (CPM) method. Differential expression was done using DESeq2 (v3.5) with default parameters. Genes were considered significantly differentially expressed based upon a fold change >2 or <-2 and an FDR <0.05.

9aa TAD Prediction

9aa TADs were predicted using previously described software (http://www.at.embnet.org/toolbox/9aatad/.)[38] using the “moderately stringent pattern” criteria and all “refinement criteria” and only TADs with 100% matches were then selected for evaluation in MCP fusion proteins.

Data Analysis

All data used for statistical analysis had a minimum 3 biological replicates. Data are presented as mean ± SEM Gene expression analyses were conducted using Student’s t-tests (Two-tailed pair or multiple unpaired). Results were considered statistically significant when the *P*-value was <0.05. All bar graphs, error bars, and statistics were generated using GraphPad Prism v 9.0.

References:

1. Qi, L.S. et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**, 1173-1183 (2013).
2. Gilbert, L.A. et al. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154**, 442-451 (2013).
3. Perez-Pinera, P. et al. RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat Methods* **10**, 973-976 (2013).
4. 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).
5. Liao, H.K. et al. In Vivo Target Gene Activation via CRISPR/Cas9-Mediated Trans-epigenetic Modulation. *Cell* **171**, 1495-1507 e1415 (2017).
6. Goell, J.H. & Hilton, I.B. CRISPR/Cas-Based Epigenome Editing: Advances, Applications, and Clinical Utility. *Trends Biotechnol* **39**, 678-691 (2021).
7. Gemberling, M.P. et al. Transgenic mice for in vivo epigenome editing with CRISPR-based systems. *Nat Methods* **18**, 965-974 (2021).
8. Chavez, A. et al. Highly efficient Cas9-mediated transcriptional programming. *Nat Methods* **12**, 326-328 (2015).
9. 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).
10. Li, J. et al. Programmable human histone phosphorylation and gene activation using a CRISPR/Cas9-based chromatin kinase. *Nat Commun* **12**, 896 (2021).
11. Tanenbaum, M.E., Gilbert, L.A., Qi, L.S., Weissman, J.S. & Vale, R.D. A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell* **159**, 635-646 (2014).
12. Konermann, S. et al. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* **517**, 583-588 (2015).
13. Zalatan, J.G. et al. Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. *Cell* **160**, 339-350 (2015).
14. Weltner, J. et al. Human pluripotent reprogramming with CRISPR activators. *Nat Commun* **9**, 2643 (2018).
15. 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).
16. Li, K. et al. Interrogation of enhancer function by enhancer-targeting CRISPR epigenetic editing. *Nat Commun* **11**, 485 (2020).
17. Lambert, S.A. et al. The Human Transcription Factors. *Cell* **175**, 598-599 (2018).
18. Soto, L.F. et al. Compendium of human transcription factor effector domains. *Mol Cell* (2021).
19. Tycko, J. et al. High-Throughput Discovery and Characterization of Human Transcriptional Effectors. *Cell* **183**, 2020-2035 e2016 (2020).
20. Mammoto, A., Mammoto, T. & Ingber, D.E. Mechanosensitive mechanisms in transcriptional regulation. *J Cell Sci* **125**, 3061-3073 (2012).
21. Wagh, K. et al. Mechanical Regulation of Transcription: Recent Advances. *Trends Cell Biol* **31**, 457-472 (2021).

22. Katoh, Y. et al. Two domains of Nrf2 cooperatively bind CBP, a CREB binding protein, and synergistically activate transcription. *Genes Cells* **6**, 857-868 (2001).
23. Galli, G.G. et al. YAP Drives Growth by Controlling Transcriptional Pause Release from Dynamic Enhancers. *Mol Cell* **60**, 328-337 (2015).
24. He, H. et al. Transcriptional factors p300 and MRTF-A synergistically enhance the expression of migration-related genes in MCF-7 breast cancer cells. *Biochem Biophys Res Commun* **467**, 813-820 (2015).
25. Zanconato, F. et al. Transcriptional addiction in cancer cells is mediated by YAP/TAZ through BRD4. *Nat Med* **24**, 1599-1610 (2018).
26. Dasgupta, I. & McCollum, D. Control of cellular responses to mechanical cues through YAP/TAZ regulation. *J Biol Chem* **294**, 17693-17706 (2019).
27. Zhao, J. et al. Chemokines protect vascular smooth muscle cells from cell death induced by cyclic mechanical stretch. *Sci Rep* **7**, 16128 (2017).
28. McSweeney, S.R., Warabi, E. & Siow, R.C. Nrf2 as an Endothelial Mechanosensitive Transcription Factor: Going With the Flow. *Hypertension* **67**, 20-29 (2016).
29. Zhao, B. et al. Inactivation of YAP oncoprotein by the Hippo pathway is involved in cell contact inhibition and tissue growth control. *Genes Dev* **21**, 2747-2761 (2007).
30. Bromberg, J. & Darnell, J.E., Jr. The role of STATs in transcriptional control and their impact on cellular function. *Oncogene* **19**, 2468-2473 (2000).
31. Nordhoff, V. et al. Comparative analysis of human, bovine, and murine Oct-4 upstream promoter sequences. *Mamm Genome* **12**, 309-317 (2001).
32. Chen, J.C., Love, C.M. & Goldhamer, D.J. Two upstream enhancers collaborate to regulate the spatial patterning and timing of MyoD transcription during mouse development. *Dev Dyn* **221**, 274-288 (2001).
33. Tolhuis, B., Palstra, R.J., Splinter, E., Grosveld, F. & de Laat, W. Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol Cell* **10**, 1453-1465 (2002).
34. Carter, D., Chakalova, L., Osborne, C.S., Dai, Y.F. & Fraser, P. Long-range chromatin regulatory interactions in vivo. *Nat Genet* **32**, 623-626 (2002).
35. Zhang, Z. et al. Transcriptional landscape and clinical utility of enhancer RNAs for eRNA-targeted therapy in cancer. *Nat Commun* **10**, 4562 (2019).
36. Nishimasu, H. et al. Crystal Structure of Staphylococcus aureus Cas9. *Cell* **162**, 1113-1126 (2015).
37. Zhang, X. et al. MiniCAFE, a CRISPR/Cas9-based compact and potent transcriptional activator, elicits gene expression in vivo. *Nucleic Acids Res* **49**, 4171-4185 (2021).
38. Piskacek, S. et al. Nine-amino-acid transactivation domain: establishment and prediction utilities. *Genomics* **89**, 756-768 (2007).
39. Piskacek, M., Vasku, A., Hajek, R. & Knight, A. Shared structural features of the 9aaTAD family in complex with CBP. *Mol Biosyst* **11**, 844-851 (2015).
40. Sternberg, S.H., LaFrance, B., Kaplan, M. & Doudna, J.A. Conformational control of DNA target cleavage by CRISPR-Cas9. *Nature* **527**, 110-113 (2015).
41. Shams, A. et al. Comprehensive deletion landscape of CRISPR-Cas9 identifies minimal RNA-guided DNA-binding modules. *Nat Commun* **12**, 5664 (2021).

42. Kunii, A. et al. Three-Component Repurposed Technology for Enhanced Expression: Highly Accumulable Transcriptional Activators via Branched Tag Arrays. *CRISPR J* **1**, 337-347 (2018).
43. Zhou, H. et al. In vivo simultaneous transcriptional activation of multiple genes in the brain using CRISPR-dCas9-activator transgenic mice. *Nat Neurosci* **21**, 440-446 (2018).
44. Pickar-Oliver, A. et al. Targeted transcriptional modulation with type I CRISPR-Cas systems in human cells. *Nat Biotechnol* **37**, 1493-1501 (2019).
45. Chen, Y. et al. Repurposing type I-F CRISPR-Cas system as a transcriptional activation tool in human cells. *Nat Commun* **11**, 3136 (2020).
46. Zetsche, B. et al. Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. *Nat Biotechnol* **35**, 31-34 (2017).
47. Polo, J.M. et al. A molecular roadmap of reprogramming somatic cells into iPS cells. *Cell* **151**, 1617-1632 (2012).
48. Nishimura, K. et al. Manipulation of KLF4 expression generates iPSCs paused at successive stages of reprogramming. *Stem Cell Reports* **3**, 915-929 (2014).
49. Takahashi, K. et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861-872 (2007).
50. Ramana, C.V., Chatterjee-Kishore, M., Nguyen, H. & Stark, G.R. Complex roles of Stat1 in regulating gene expression. *Oncogene* **19**, 2619-2627 (2000).
51. Esnault, C. et al. Rho-actin signaling to the MRTF coactivators dominates the immediate transcriptional response to serum in fibroblasts. *Genes Dev* **28**, 943-958 (2014).
52. Tonelli, C., Chio, I.I.C. & Tuveson, D.A. Transcriptional Regulation by Nrf2. *Antioxid Redox Signal* **29**, 1727-1745 (2018).
53. Wojciak, J.M., Martinez-Yamout, M.A., Dyson, H.J. & Wright, P.E. Structural basis for recruitment of CBP/p300 coactivators by STAT1 and STAT2 transactivation domains. *EMBO J* **28**, 948-958 (2009).
54. Sun, Z., Chin, Y.E. & Zhang, D.D. Acetylation of Nrf2 by p300/CBP augments promoter-specific DNA binding of Nrf2 during the antioxidant response. *Mol Cell Biol* **29**, 2658-2672 (2009).
55. Black, J.B. et al. Master Regulators and Cofactors of Human Neuronal Cell Fate Specification Identified by CRISPR Gene Activation Screens. *Cell Rep* **33**, 108460 (2020).
56. Liu, Y. et al. CRISPR Activation Screens Systematically Identify Factors that Drive Neuronal Fate and Reprogramming. *Cell Stem Cell* **23**, 758-771 e758 (2018).
57. Weltner, J. & Trokovic, R. Reprogramming of Fibroblasts to Human iPSCs by CRISPR Activators. *Methods Mol Biol* **2239**, 175-198 (2021).

Data Availability

Selected CRISPR/Cas constructs and fusions are available through Addgene, and all reagents are available from the authors upon request. All RNA sequencing data are available through the NCBI.

Acknowledgement
The authors thank Jacopo de Rossi and Dhiraj Jain for their assistance with 9aa TADs prediction and cloning. The authors thank Arijita Sarkar for her critical help with NGS data analysis and the selection of miRNA target loci in HEK293T cells. The authors also thank Adrian Picker-Oliver and Charles A. Gersbach for providing *E. coli* Type I CRISPR Cascade plasmids. The authors thank all members of the Hilton lab for helpful discussions and insights.

Author Contributions
B.M. and I.B.H. conceived the project and designed experiments. B.M performed experiments and analyzed the data with the assistance of J.L., A.C., D.A.B., R.S.G-R., J.G., K.W., M.E., Y.G., and A.K.P. B.M. and I.B.H. wrote the manuscript with input from all authors.

Funding
This work was supported by a Cancer Prevention & Research Institute of Texas (CPRIT) Award (RR170030) and NIH R35 Award (R35GM143532) to I.B.H.

Conflict of Interest
B.M., J.G., and I.B.H. have filed a provisional patent related to this work. I.B.H. has filed patent applications related to other CRISPR technologies for genome engineering.

Figures and Figure Legends

Figure 1

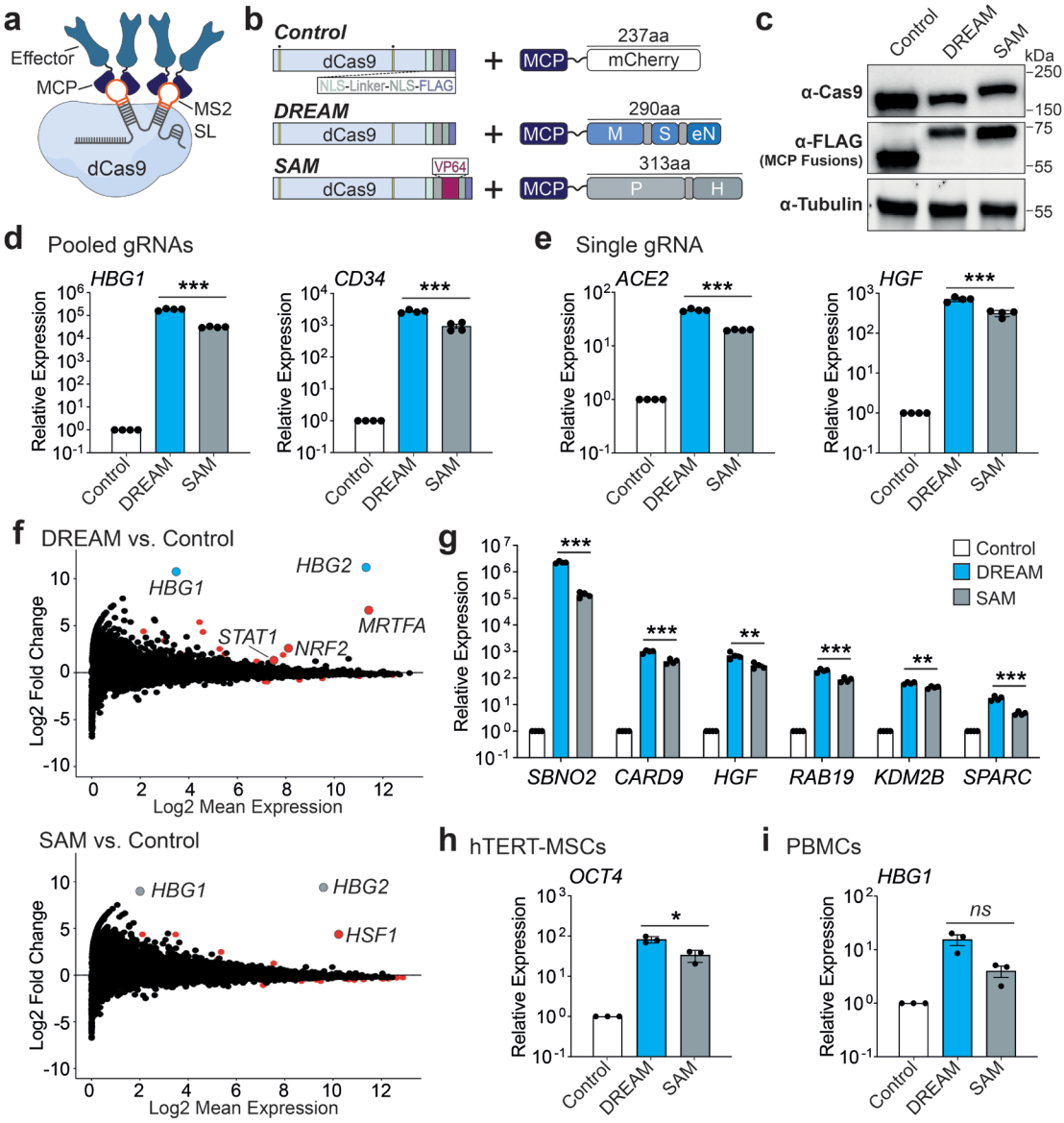

Fig.1. CRISPR-DREAM displays potent activation at human promoters, has high specificity, and is robust across cell types.

a. Nuclease inactivated *Streptococcus pyogenes* dCas9 (dCas9), a gRNA containing two engineered MS2 stem-loops (MS2 SLs) and MS2 binding Cap Protein (MCP)-fused transcriptional effector proteins are schematically depicted. Nuclease-inactivating mutations (D10A and H840A) are indicated by yellow bars with dots above. **b**. dCas9 and MCP-fusion proteins, including an MCP-mCherry fusion (Control; top), the engineered tripartite MCP-MSN domain fusion (DREAM system; middle), and dCas9-VP64 and the MCP-p65-HSF1 fusion protein (SAM system; bottom) are schematically depicted. **c**. The

expression levels of dCas9 and dCas9-VP64 (top), FLAG tagged MCP-mCherry, FLAG tagged MCP-MSN, FLAG tagged MCP-p65-HSF1 (middle), and β-Tubulin (loading control; bottom) are shown as detected by Western blotting in HEK293T cells 72hrs post-transfection. **d and e**. Relative expression levels of endogenous human genes 72hrs after Control, DREAM, or SAM systems were targeted to their respective promoters using pools of 4 or 3 gRNAs (*HBG1* and *CD34*, respectively; **panel d**), or using single gRNAs (*ACE2* and *HGF*, respectively; **panel e**) as measured by QPCR. **f.** Transcriptome wide RNA-seq data generated 72hrs after the DREAM (top) or SAM (bottom) systems were targeted to the *HBG1/HBG2* promoter using 4 pooled gRNAs. mRNAs identified as significantly differentially expressed (fold change >2 or <-2 and FDR <0.05) are shown as red dots in both MA plots. In the top MA plot (CRISPR-DREAM), mRNAs corresponding to *HBG1/HBG2* (target genes) are highlighted in light blue. mRNAs encoding components of the MSN tripartite fusion protein (*MRTF-A/STAT1/NRF2*; red), were also significantly differentially expressed (fold change >2 and FDR <0.05). In the bottom MA plot (SAM system), mRNAs corresponding to *HBG1/HBG2* (target genes) are highlighted in light gray. *HSF1* mRNA (a component of the p65-HSF1 bipartite fusion protein; red), was also significantly differentially expressed (fold change >2 and FDR <0.05). **g.** 6 endogenous genes were activated by DREAM or SAM using a pool of gRNAs (1 gRNA/gene) in HEK293T cells. **h and i.** *OCT4* (**panel h**) or *HBG1* (**panel i**) gene activation by DREAM or SAM systems when corresponding promoters were targeted by 4 gRNAs per promoter in hTERT-MSC or PMBC cells, respectively. All QPCR samples were processed 72hrs post-transfection and are the result of at least 3 biological replicates. Error bars; SEM. *; $P < 0.05$, **; $P < 0.01$, ***; $P < 0.001$. *ns*; not significant.

Figure 2

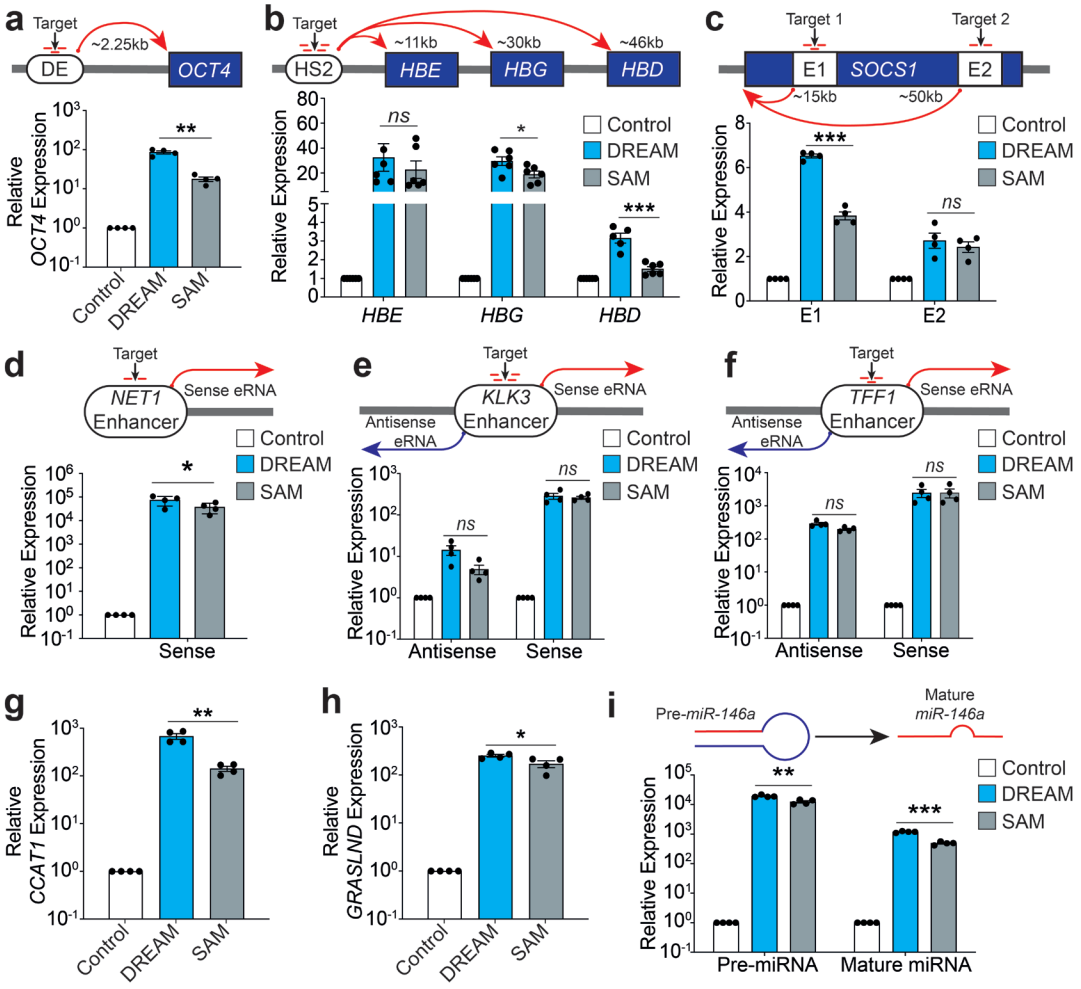

Fig. 2. CRISPR-DREAM efficiently activates transcription from diverse human regulatory elements.

a-c. CRISPR-DREAM and the SAM system activated downstream mRNA expression from *OCT4* (**panel a**), *HBE*, *HBG,* and *HBD* (**panel b**)*, and SOCS1* (**panel c**), when targeted to the *OCT4* distal enhancer (DE), HS2 enhancer, or one of two intragenic *SOCS1* enhancers, using pools of 3 (*OCT4* DE), 4 (*HS2*), 3 (*SOCS1* +15kb), or 2 (*SOCS1* + 50kb) gRNAs respectively. **d.** CRISPR-DREAM and the SAM system activated sense eRNA expression when targeted to the *NET1* enhancer using 2 gRNAs. **e and f.** CRISPR-DREAM and the SAM system bidirectionally activated eRNA expression when targeted to the *KLK3* (**panel e**) or *TFF1* (**panel f**) enhancers using pools of 4 or 3 gRNAs, respectively. **g and h.** CRISPR-DREAM and the SAM system activated the expression of long noncoding RNA when targeted to the *CCAT1* (**panel g**) or *GRASLND* (**panel h**) promoters using pools of 4 gRNAs, respectively. **i.** CRISPR-DREAM and the SAM system activated the expression of pre and mature *miR-146a* when targeted to the *miR-146a* promoter using a pool of 4 gRNAs. All samples were processed for QPCR 72hrs post-transfection. Data are the result of at least 4 biological replicates. Error bars; SEM. *; $P < 0.05$, **; $P < 0.01$, ***; $P < 0.001$. *ns*; not significant.

Figure 3

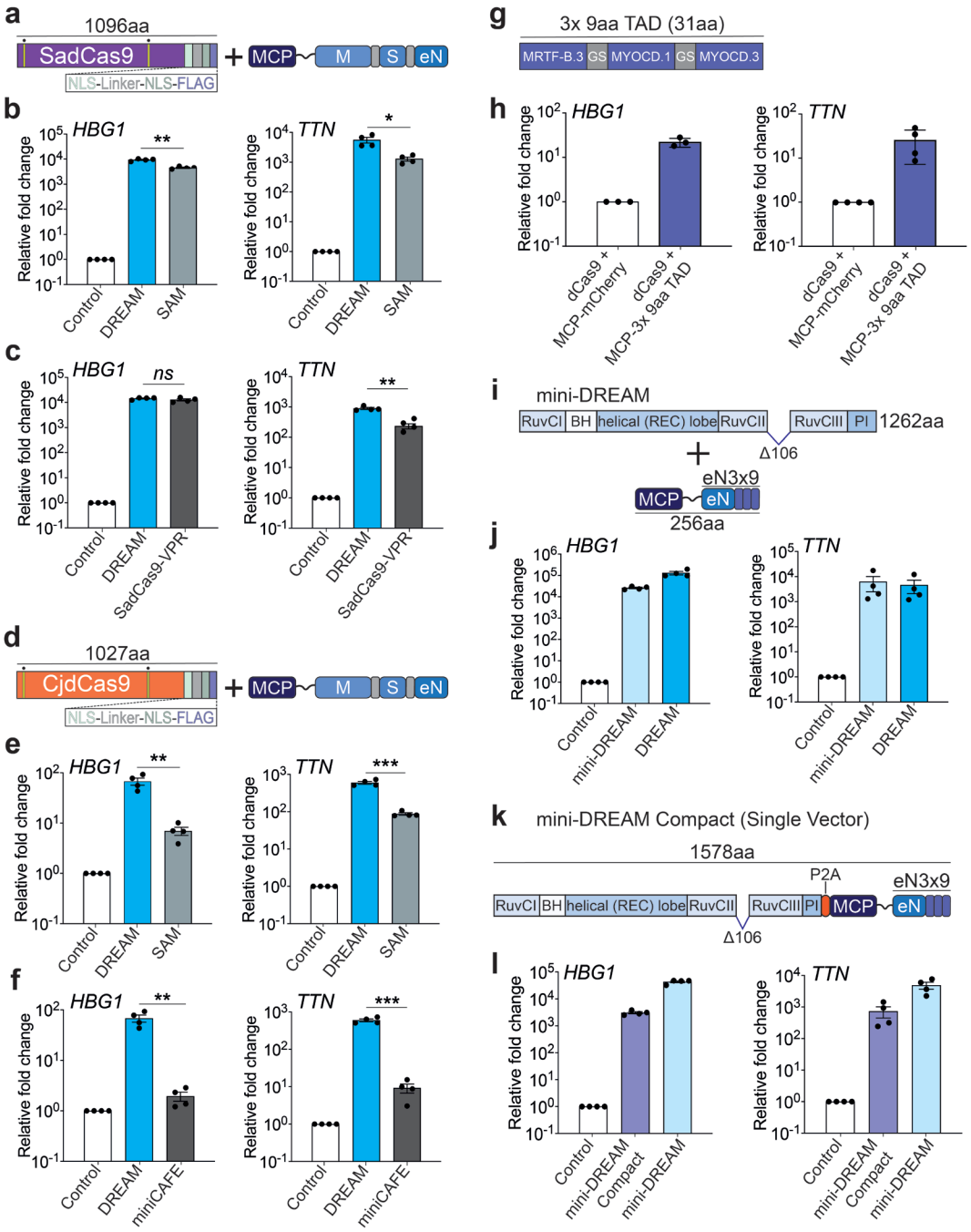

Fig. 3. CRISPR-DREAM is portable to orthogonal dCas9 proteins and amenable to miniaturization.

a. The SadCas9-DREAM system is schematically depicted, and nuclease-inactivating mutations (D10A and N580A) are indicated by yellow bars with dots above. **b.** *HBG1* (left) or *TTN* (right) gene activation using the SadCas9-DREAM or SadCas9-SAM systems, when targeted to each corresponding promoter using pools of 4 gRNAs, respectively. **c.** *HBG1* (left) or *TTN* (right) gene activation using the SadCas9-DREAM or SadCas9-VPR

systems, when targeted to each corresponding promoter using pools of 4 MS2-modifed (SadCas9-DREAM) or standard gRNAs (SadCas9-VPR), respectively. **d.** The CjdCas9-DREAM system is schematically depicted, and nuclease-inactivating mutations (D8A and H559A) are indicated by yellow bars with dots above. **e.** *HBG1* (left) or *TTN* (right) gene activation using the CjdCas9-DREAM or CjdCas9-SAM systems, when targeted to each corresponding promoter using pools of 3 MS2-modified gRNAs, respectively. **f.** *HBG1* (left) or *TTN* (right) gene activation using the CjdCas9-DREAM or MiniCAFE systems, when targeted to each corresponding promoter using pools of 3 MS2-modifed (SadCas9-DREAM) or standard gRNAs (miniCAFE), respectively. **g.** A 3x 9aa TAD derived from MYOCD and MRTF-B TADs is schematically depicted, GS; glycine-serine linker. **h.** *HBG1* (left) or *TTN* (right) gene activation when the 3x 9aa TAD was fused to MCP and recruited to each corresponding promoter using dCas9 and a pool of 4 MS2-modified gRNAs, respectively. **i.** The mini-DREAM system is schematically depicted. MCP-eN3x9 is a fusion protein consisting of MCP, eNRF2, and the 3x 9aa TAD derived from MYOCD and MRTF-B TADs. **j.** *HBG1* (left) or *TTN* (right) gene activation when either the mini-DREAM or CRISPR-DREAM system was targeted to each corresponding promoter using a pool of 4 MS2-modified gRNAs, respectively. **k.** The mini-DREAM Compact system is schematically depicted, P2A; self-cleaving peptide. **l.** *HBG1* (left) or *TTN* (right) gene activation when either the mini-DREAM Compact or mini-DREAM system was targeted to each corresponding promoter using a pool of 4 MS2-modified gRNAs, respectively. All samples were processed for QPCR 72hrs post-transfection. Data are the result of at least 3 biological replicates. Error bars; SEM. *; $P < 0.05$, **; $P < 0.01$, ***; $P < 0.001$. *ns*; not significant.

Figure 4

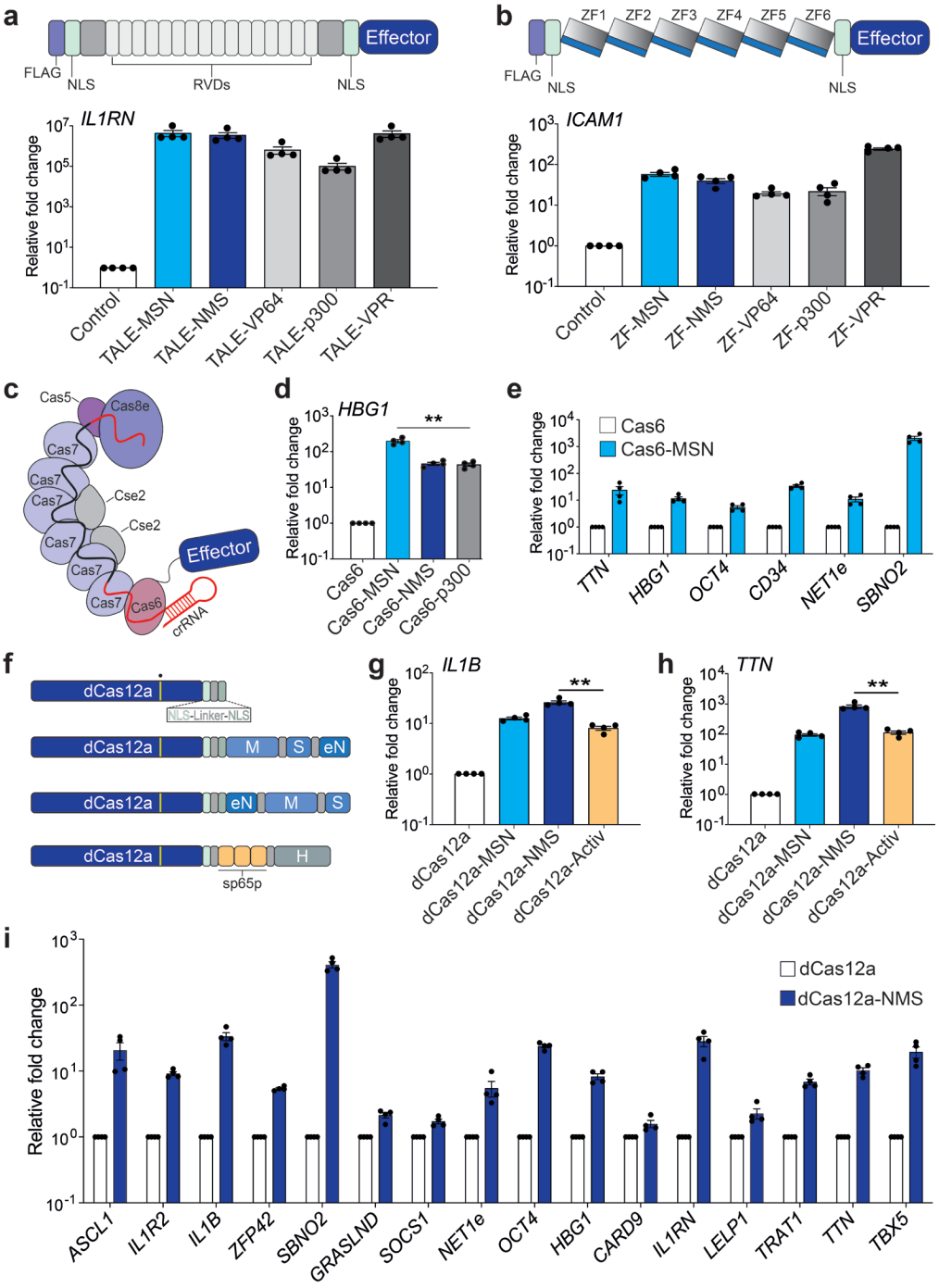

Fig. 4. The MSN and NMS effector domains are portable to diverse DNA binding platforms and enable superior multiplexing when fused to dCas12a.
a. Synthetic transcription activator-like effector (TALE) proteins harboring indicated effector domains were designed to target the human *IL1RN* promoter. Repeat variable di-residues, RVDs. Relative *IL1RN* expression (bottom) 72hrs after indicated TALE fusion protein encoding plasmids were transfected. **b**. Synthetic zinc finger (ZF) proteins

harboring indicated effector domains were designed to target the human *ICAM1* promoter. Relative *ICAM1* expression (bottom) 72hrs after indicated ZF fusion protein encoding plasmids were transfected. **c**. The Type I CRISPR system derived from *E. Coli* K-12 (Eco-cascade) is schematically depicted along with an effector fused to the Cas6 protein subunit. **d**. *HBG1* gene activation when either the MSN, NMS, or p300 effector domains were fused to Cas6 and the respective engineered Eco-Cascade complexes were targeted to the *HBG1* promoter using a single crRNA. **e**. Multiplexed activation of 6 endogenous genes 72hrs after co-transfection of Eco-cascade complexes when MSN was fused to Cas6 and targeted using a single crRNA array expression plasmid (1 crRNA/promoter). **f.** The dCas12a protein and indicated fusions are schematically depicted along with the G993A DNase-inactivating mutation indicated by a yellow bar with a dot above. **g and h**. *IL1B* (**panel g**) or *TTN* (**panel h**) gene activation using the indicated dCas12a fusion proteins when targeted to each corresponding promoter using a pool of 2 crRNAs (for *IL1B*) or a single array encoding 3 crRNAs (*TTN*), respectively. **i**. Multiplexed activation of 16 indicated endogenous genes 72hrs after co-transfection of dCas12a-NMS and a single crRNA array expression plasmid encoding 20 crRNAs. All samples were processed for QPCR 72hrs post-transfection in HEK293T cells. Data are the result of at least 4 biological replicates. Error bars; SEM. **; $P < 0.01$.

Figure 5

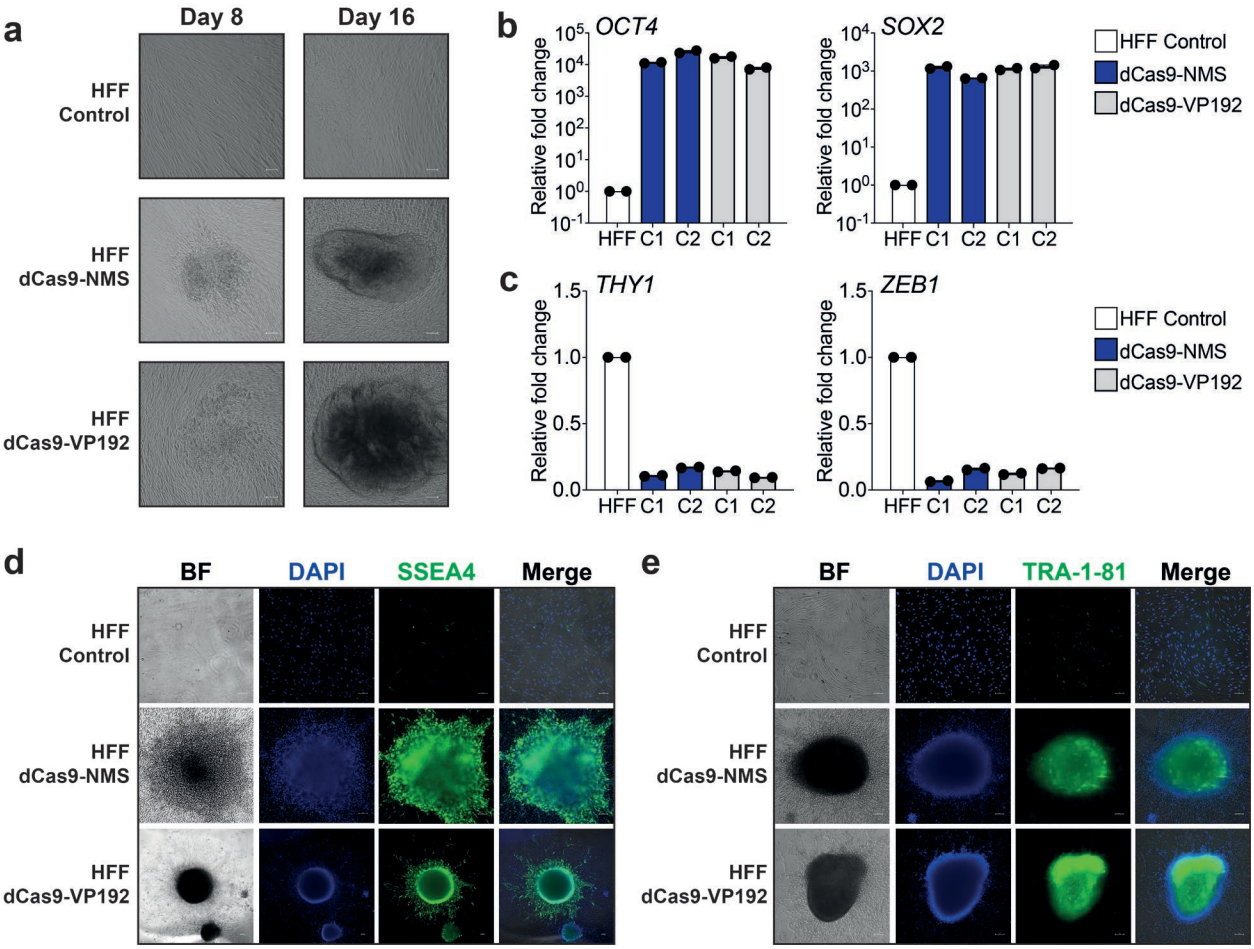

Fig.5 dCas9-NMS permits efficient *in vitro* reprogramming of human fibroblasts. a. Primary human foreskin fibroblasts (HFFs) were nucleofected with plasmids encoding 15 multiplexed gRNAs targeting the *OCT4*, *SOX2*, *KLF4*, c-*MYC*, and *LIN28A* promoter and EEA motifs (as in previous reports[14]), and either dCas9-NMS (middle row) or dCas9-VP192 (bottom row). HFF morphology was analyzed 8 and 16 days later (white scale bars, 100μm). **b.** Relative expression of pluripotency-associated genes *OCT4* (left) and *SOX2* (right) in representative iPSC colonies (C1 or C2) approximately 40 days after nucleofection of either dCas9-NMS (blue) or dCas9-VP192 (gray) and multiplexed gRNAs compared to untreated HFF controls. **c.** Relative expression of mesenchymal-associated genes *THY1* (left) and *ZEB1* (right) in representative iPSC colonies (C1 or C2) approximately 40 days after nucleofection of either dCas9-NMS (blue) or dCas9-VP192 (gray) and multiplexed gRNAs compared to untreated HFF controls. **d and e.** Immunofluorescence microscopy of HFFs approximately 40 days after nucleofection of either dCas9-NMS or dCas9-VP192 and multiplexed gRNAs compared to untreated HFF controls (white scale bars, 100μm). Cells were stained for the expression of pluripotency-associated cell surface markers SSEA4 (**panel d**, green) or TRA-1-81 (**panel e**, green). All cells were counterstained with DAPI for nuclear visualization.