Official journal of the American Society of Gene & Cell Therapy

MTOpen

Potent and Targeted Activation of Latent HIV-1 Using the CRISPR/dCas9 Activator Complex

Sheena M Saayman[1,2], Daniel C Lazar[1], Tristan A Scott[2], Jonathan R Hart[1], Mayumi Takahashi[3], John C Burnett[3], Vicente Planelles[4], Kevin V Morris[1,5] and Marc S Weinberg[1,2,6]

[1]Department of Molecular and Experimental Medicine, The Scripps Research Institute, La Jolla, California, USA; [2]HIV Pathogenesis Research Unit, Department of Molecular Medicine and Haematology, School of Pathology, University of the Witwatersrand, Johannesburg, South Africa; [3]Division of Molecular Biology, Beckman Research Institute at the City of Hope, Duarte, California, USA; [4]Division of Microbiology and Immunology, Department of Pathology, University of Utah School of Medicine, Salt Lake City, Utah, USA; [5]School of Biotechnology and Biomedical Sciences, University of New South Wales, Kensington, New South Wales, Australia; [6]Wits/SA MRC Antiviral Gene Therapy Research Unit, Department of Molecular Medicine and Haematology, University of the Witwatersrand, Johannesburg, South Africa

HIV-1 provirus integration results in a persistent latently infected reservoir that is recalcitrant to combined antiretroviral therapy (cART) with lifelong treatment being the only option. The "shock and kill" strategy aims to eradicate latent HIV by reactivating proviral gene expression in the context of cART treatment. Gene-specific transcriptional activation can be achieved using the RNA-guided CRISPR-Cas9 system comprising single guide RNAs (sgRNAs) with a nuclease-deficient Cas9 mutant (dCas9) fused to the VP64 transactivation domain (dCas9-VP64). We engineered this system to target 23 sites within the long terminal repeat promoter of HIV-1 and identified a "hotspot" for activation within the viral enhancer sequence. Activating sgRNAs transcriptionally modulated the latent proviral genome across multiple different *in vitro* latency cell models including T cells comprising a clonally integrated mCherry-IRES-Tat (LChIT) latency system. We detected consistent and effective activation of latent virus mediated by activator sgRNAs, whereas latency reversal agents produced variable activation responses. Transcriptomic analysis revealed dCas9-VP64/sgRNAs to be highly specific, while the well-characterized chemical activator TNFα induced widespread gene dysregulation. CRISPR-mediated gene activation represents a novel system which provides enhanced efficiency and specificity in a targeted latency reactivation strategy and represents a promising approach to a "functional cure" of HIV/AIDS.

Received 27 August 2015; accepted 23 October 2015; advance online publication 5 January 2016. doi:10.1038/mt.2015.202

INTRODUCTION

Combined antiretroviral therapies (cARTs) have had a marked impact on the treatment and progression of HIV/AIDS, reducing the morbidity, mortality, and transmission of HIV-related illness.[1] Despite these successes, the most significant limitation of currently available cARTs is their inability to purge latent HIV reservoirs, resulting in a persistent infection even under lifelong treatment.[2] Integrated viral copies, which mostly lie within resting memory $CD4^+$ T cells and also other long-lived "reservoir" cells, persist for as long as 73 years in patients receiving cART.[2,3] A promising strategy to eradicate latent HIV reservoirs has been to purge the pool of latently infected cells in the presence of cARTs by reactivating dormant virus: a strategy known as "shock and kill."[4] Reactivation of latent HIV purges infected cells directly via active viral replication or indirectly via the host immune system. cARTs can then act to prevent new infection from the released virus and thereby extinguish the reservoir. Therapeutic attempts to purge the reservoir of latently infected cells have thus far been unsuccessful. These approaches made use of histone deacetylase inhibitors (HDACi) such as valproic acid,[5] which has had a limited impact on reducing the size of the latent reservoir.[6,7] More recent attempts with the HDACi suberoylanilide hydroxamic acid (SAHA/vorinostat)[8–12] and panobinostat[13] showed more promising results, yet it still remains unclear how these effects will translate clinically relative to the total size of the latent reservoir. Future efforts in this space seem to be aimed more at using different cocktails of latency-reversing agents (LRAs).[14,15] A major safety concern with this approach lies with the risk of widespread and nonspecific induction of host gene expression and T-cell activation.[16] Novel LRAs that specifically target the integrated provirus and which function by inducing HIV expression, remain a sought after objective for eradicating latent infection.

The clustered regularly interspaced palindromic repeats (CRISPR)/Cas9 gene-editing system has been exploited as a novel tool for both gene editing and gene regulation. CRISPR/Cas9, in its native function, is an acquired form of immunity in prokaryotes[17,18]; however, the basic elements of the system have been modified into a minimal system that can be engineered to target and modulate virtually any DNA sequence.[19–21] The engineered CRISPR/Cas9 system comprising a short chimeric single guide RNA (sgRNA) and a separate Cas9 nuclease has transformed the field of biology and medicine.[22,23] Mutations

The first two authors are co-first authors.

Correspondence: Marc S. Weinberg, Department of Molecular and Haematology, School of Pathology, University of the Witwatersrand, Parktown, South Africa. E-mail: marcow@scripps.edu

in both Cas9 endonuclease domains result in deactivated or "dead" Cas9 (dCas9), enabling the development of programmable RNA-dependent DNA-binding proteins.[24,25] Gene-specific transcriptional activation has been achieved with fusion variants of dCas9 with activation domains (ADs) that function as a transcriptional cellular programming tool.[24,26–29] Initial systems made use of dCas9 fused to several copies of the C-terminal herpes virus transcriptional activation domain 16 (VP16) and coupled with a sgRNA to induce highly specific and enhanced expression of targeted genes.[24,26–29] In some cases, activation can be enhanced with several nonoverlapping sgRNAs, which recruit multiple copies of the dCas9-VP64 to a targeted promoter.[26–29] More recent activation modules include a hybrid fusion of VP64, p65 (RelA), and the Epstein–Barr virus R transactivator (Rta) to dCas9 (dCas9-VPR)[30]; or include the catalytic histone acetyltransferase core domain of the human E1A-associated protein p300 (dCas9-p300).[31] Additionally, engineered systems have been generated for recruiting multiple ADs to a single dCas9 molecule.[32,33] The SunTag system uses a polypeptide scaffold fused to dCas9 to recruit multiple antibody-fusion proteins to link a promoter-localized dCas9 to multiple VP64 domains.[32] The synergistic activation mediator (SAM) system uses modifications to the loop regions in the sgRNA scaffold to allow for embedded RNA aptamers, such as those that bind to the dimerized bacteriophage coat protein MS2, to recruit MS2-bound ADs. SAM recruits multiple ADs for enhanced activation at a single sgRNA-targeted dCas9-VP64 bound locus.[33]

In this study, we identify a targeted hotspot for CRISPR activation within the long terminal repeat (LTR) enhancer region, at the junction between two nuclear factor (NF)-κB transcription factor–binding sites. We demonstrate that RNA-guided dCas9-activator targeting at this locus is highly effective at inducing potent transcriptional activation of latent HIV-1 infection in human T cells. Unlike chemical activating agents, RNA-guided dCas9-VP64 reactivates HIV-1 in all latency models tested and does not lead to nonspecific global gene expression or adverse cellular toxicities. RNA-guided CRISPR activation provides an exciting new avenue for targeted reactivation of latent HIV.

RESULTS

Targeting the HIV-1 5′ LTR with sgRNAs and dCas9-VP64 for activation of HIV-1 transcription

Twenty-three possible NGG *Streptococcus pyogenes* (Sp) Cas9 photospacer adjacent motif (PAM) sites for targeting of sgRNAs were identified in the U3 region of the LTR, upstream of the HIV transcriptional start site (TSS) (−450 to 0 bp; HIV genome HXB2) within the enhancer and modulatory region of the promoter. All sgRNAs were named according to the 3′ adjacent sense strand nucleotide cleavage site catalyzed by nuclease-active Cas9 (**Figure 1a**) and were screened together with dCas9-VP64 or dCas9-VP160 for their activation properties by transient transfection with the reporter NL4-3.Luc.R$^-$E$^-$, a full-length HIV molecular clone where luciferase is driven by the viral LTR. Variable activation efficiencies were observed; however, several candidate sgRNAs induced a 10–20-fold increase in gene expression (**Figure 1b**). The viral enhancer region emerged as a "hotspot" region favorable for targeted activation. The guide sg362F induced the highest observed activation and overlaps the NF-κB–binding site doublet.

The activation efficacy of dCas9-VP64 and sgRNAs targeted to the HIV-1 5′ LTR in *in vitro* latency models

To assess the ability of the LTR-targeting sgRNAs to activate viral transcription in a robust model of latency, we made use of the LTR-mCherry-IRES-Tat (LChIT) reporter system (**Figure 2a**)[34] that allows for exquisite sensitivity to alterations in viral gene expression. CD4$^+$ T cells transduced with LChIT are in a TAT-mediated positive feedback loop that drives expression of TAT and mCherry from a full-length LTR. Screening of single-copy integrants from clonal CEM T-cell populations enabled us to select LChIT clone 3.2, which is in a bimodal state, expressing mCherry in ~20% of the population (**Figure 2a**). We selected five effective sgRNA candidates targeting different regions of the LTR, including sg78F, sg125F, sg221F, sg223R, and sg362F, identified as activators in our initial screen. Nucleofected sgRNAs and dCas9-VP64 (ref. 35) led to a range of ~3–40-fold activation of gene expression (**Figure 2b**). Again, sg362F proved to be the most potent activator, with a 37.5-fold increase in reporter gene levels (and 85% of the cells expressing mCherry). In parallel, cells were treated with known LRAs.[36] Both tumor necrosis factor-alpha (TNFα) and prostratin activate NF-κB, with the latter acting via the protein kinase C pathway. SAHA was the HDACi used. TNFα produced the most robust signal. Modest activation was observed for prostratin and SAHA, with levels significantly lower than sg362F-mediated activation (**Figure 2c**).

One of the limiting factors for study of HIV latency is the inability to truly recapitulate the latent infection *in vivo*. Latent reservoirs reside in multiple different cell types and include different quasi-species of HIV genomic integration sites[37] and cellular signaling pathways. While multiple *in vitro* latency models exist, no single experimental system of latency adequately depicts the cellular environment *in vivo.* For example, *in vitro* cell models show different levels of sensitivity to HIV reactivation when using specific LRA classes,[36] and none consistently replicate the characteristics of latently infected T cells from patients. In order to mitigate against these heterogeneous effects, we screened our sgRNAs across seven different *in vitro* latency cell models. Each cell line was transfected with the dCas9-VP64 together with individual sgRNA-expressing plasmids. We screened a selection of effective sgRNA candidates, including sg78F, sg125F, sg221F, sg223R, and sg362F, identified in our initial screen (**Figure 1b**).

The ACH2 cell line is a chronically infected T cell line harboring a mutation in TAT, which impairs its transactivation function and is commonly used as a latency cell model.[38] Supporting previous observations in initial screens and CEM LChIT 3.2 cells, sg362F again exhibited the highest activating efficacy, increasing HIV expression almost 100-fold. Single guide sg362F showed more potent activating properties than both prostratin and TNFα (**Figure 3**). The J-Lat model is based on Jurkat-derived clonal cell lines infected with lentiviruses containing the TAT and GFP open-reading frames under the control of the viral LTR promoter. The latent phenotype of this model is attributed to the integration of HIV in heterochromatic regions. These cell lines therefore show little to no GFP expression without stimulation and are often

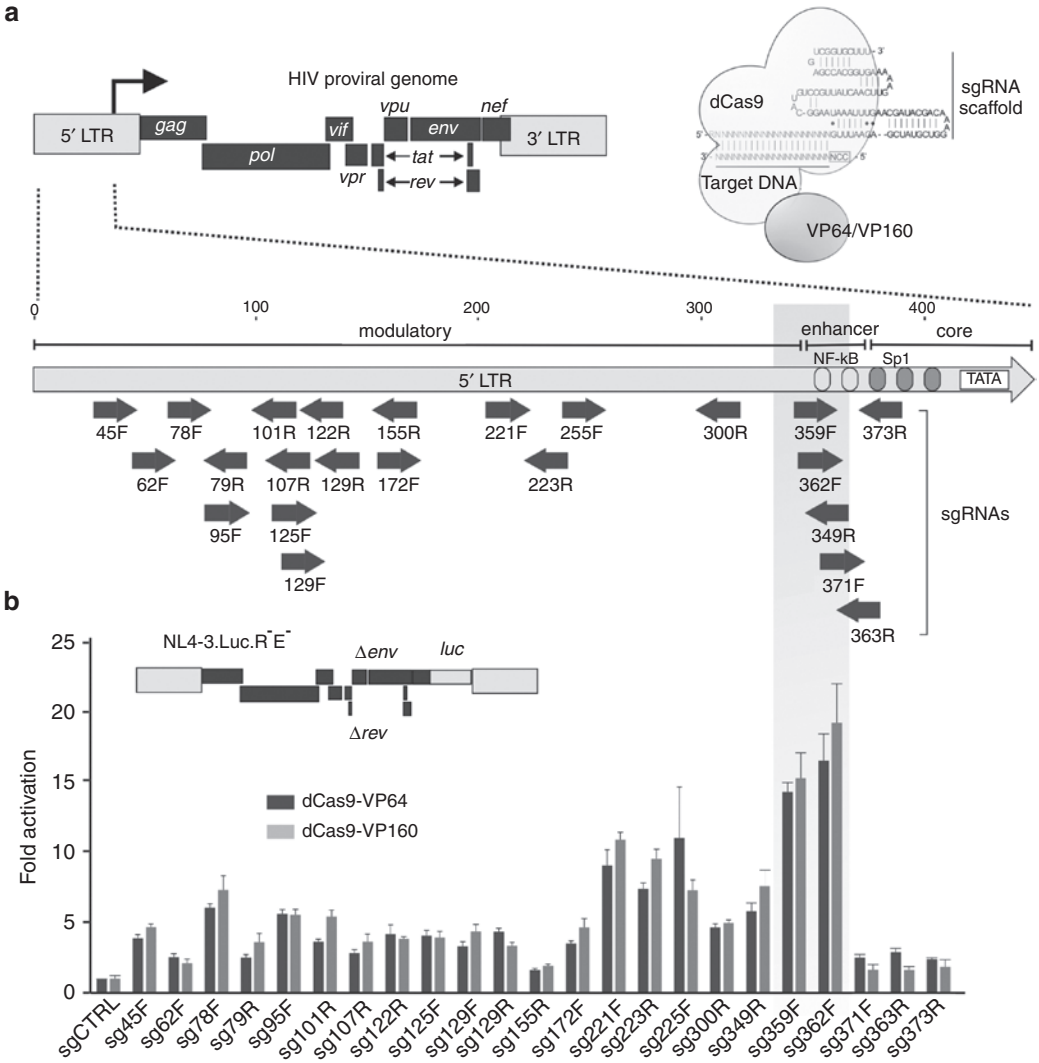

Figure 1 CRISPR activation sgRNA screen targeted to the HIV-1 5′LTR. (**a**) A representation of the HIV-1 proviral genome (HXB2) depicting 23 potential sgRNA target sites within the 5′ LTR from positions 0 to 450 (transcription start site). The NF-κB doublet (yellow) and Sp1 triplet (red) binding sites are included. A basic CRISPR/dCas9 transcriptional activation system comprises a nuclease-deficient Cas9 (dCas9) fused to a C-terminal VP64 transactivation domain. This system, coupled with a single effective sgRNA, is capable of binding to a target sequence and inducing transcriptional gene activation. (**b**) LTR-mediated activation screen by sgRNA together with dCas-VP64 or dCas9-VP160. HEK293T cells were transiently co-transfected with individual LTR-targeting sgRNAs or a nonspecific control sgRNA (sgCTRL), dCas9-VP64 or dCas9-VP160 and the pNL4-3.Luc.R$^-$E$^-$ reporter plasmid. Fold activation denotes relative luciferase activity normalized to sgCTRL. Values represent mean + SD of cells treated in triplicate.

used to study latency and reactivation.[36] We screened our activating sgRNAs across multiple J-Lat cell lines including clones 6.3, 10.6, 15.4 (Figure 3), 8.4, and 9.2 (**Supplementary Figure S1**). In J-Lat 6.3 cells, sg362F once again proved to be the most potent activator exhibiting 16-fold activation, more than five times that of prostratin, the most responsive LRA in this cell line, while in the J-Lat 15.4 cell line, sg362F increased GFP expression 48-fold, more than 10 times that of TNFα (Figure 3). It should be noted that sg362F retained its activation function in latent cell models exhibiting higher base levels of HIV gene expression such as the J1.1 chronically infected latent cell line (**Supplementary Figure S1**), and even in some cases where cells were refractory to activation by LRAs as observed in J-Lat 8.4 cells (**Supplementary Figure S1**). Furthermore, as noted previously, the panel of latent cells all demonstrated variable responses to the LRAs,[36] while sg362F showed consistent and efficient activation in all the different latency models.

The effects of spatial and sequence modifications on sg362F activation efficacy

The LTR sequences of different viral subtypes are diverse, and the region targeted by sg362F has several distinct differences in the sequence and composition of NF-κB–binding sites.[39] The sg362F guide RNA matches 19 bp spanning the junction of two identical NF-κB–binding sites. The GGG PAM (position −88 to −90) sequence forms part of the NF-κB (I) site in the NL4-3 (subtype B) LTR (Figure 4a). The "seed region" of the sg362F matches 4 bp adjacent to the PAM site, between NF-κB (I) and NF-κB (II). This likely provides sg362F with target-binding specificity despite matching fully to the ubiquitous NF-κB sequence. To determine

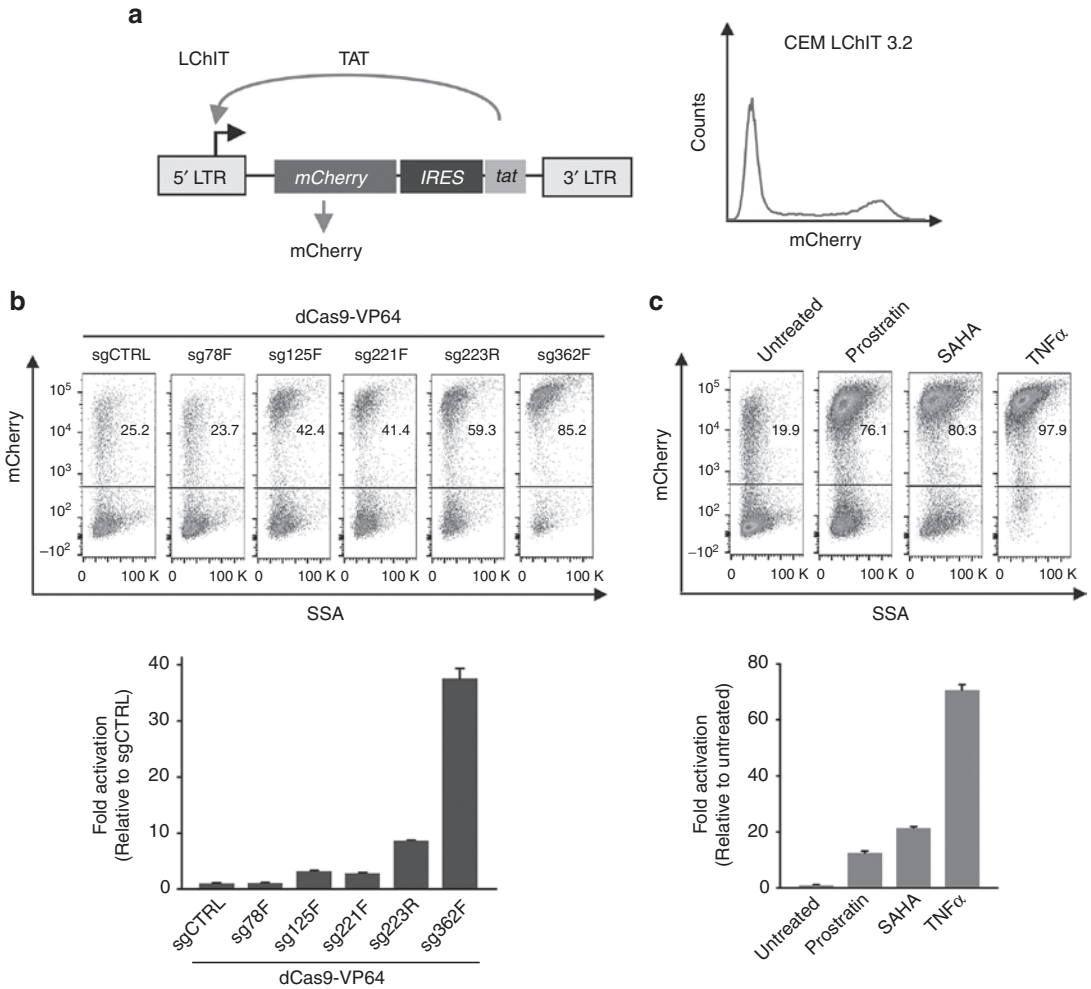

Figure 2 Activation of LTR-driven transcription by selected candidate sgRNAs, and dCas9-VP64, in the CEM T-cell-based LChIT latency model. (**a**) The LTR-mCherry-IRES-Tat (LChIT) reporter system comprises a Tat-mediated positive feedback loop that drives expression of Tat and mCherry from a full-length LTR. The CEM T-cell LChIT clone 3.2 encodes a single LChIT copy and is in a bimodal state, expressing mCherry in ~20% of the population. (**b**) The LTR-activating properties of a selection of effective sgRNA candidates including sg78F, sg125F, sg221F, sg223R, and sg362F, together with dCas9-VP64, were screened in the CEM LChIT 3.2 cell line. Cells were co-nucleofected with individual LTR-targeting sgRNAs or a nonspecific control sgCTRL, dCas9-VP64 and a GFP-expressing plasmid. Transfected CEM LChIT 3.2 cells were gated on GFP expression and mCherry expression was measured by FACS 72 hours post transfection. Graphed values represent the geometric mean of mCherry expression normalized to sgCTRL + SD in cells treated in triplicate. (**c**) HIV-1 transcription activation was analyzed in CEM LChIT 3.2 cells treated with latency reactivation agents (LRAs) known to activate latent HIV-1. These include tumor necrosis factor-alpha (TNFα), the prostratin and suberoylanilide hydroxamic acid (SAHA). LTR-driven expression (mCherry) was measured in the total treated cell population by FACS 72 hours posttreatment. Graphed values represent the geometric mean of mCherry expression normalized to untreated cells + SD in cells treated in triplicate. FACS, fluorescence activated cell sorting.

the role of sequence heterogeneity surrounding the sg362F target site, we introduced variants of sg362F that abolish or reconstitute sequence specificity at this locus using representative sequences of three HIV subtypes: B, A, and C (**Figure 4b**). LTR-driven mCherry reporters were developed for each subtype. For subtype B LTR activity, sg362F and sg362F.v1 were the most robust, with activity greater than fourfold even though sg362F.v1 has only 16 bp of complementary sequence. Guide sg362F.v2 was only twofold active with 14 bp contiguous matching bps. Variant guides sg362F.v3 and sg362F.v4 are offset by −1 and recognize GGG PAM (position −89 to −91) for subtype B and A. Only sg362F.v3 with 15 bp of contiguous sequence complementarity showed any activity. For subtype A, sg362F.v2 was capable of restoring full activity with 19 bp of matched sequence. HIV subtype C is different in that it contains a third NF-κB–like motif. As such, the subtype C sequence includes an extra GGG PAM (position −98 to −100) ~10 bp upstream of the sg362F-targeted PAM site (position −88 to −90, but missing in Subtype C). Variants sg362F.v3 and sg362F.v4, which targeted both up and downstream PAM sequences, respectively, fully reconstituted LTR-mediated activation function. However, sequence matches less than 16 bp, as well as even one mismatch in the sgRNA seed sequence, mostly abolished activation. Therefore, restoring targeted guide specificity between −90 and −120 bp from the TSS largely reconstitutes sgRNA activation, irrespective of sequence differences in this enhancer region. These data indicate that distance from the TSS is likely a key determinant for dCas9-activator function and is in agreement with previously reported optimal distances of −100 to 200 bp from the TSS.[24,30,33]

Single guide RNA sg362F functions in different CRISPR/dCas9 activation modules

To determine whether sg362F can function within the context of different CRISPR/dCas9 activation modules, we tested multiple activation complexes using our previously described LChIT model (**Figure 5a**). We tested dCas9 fusions: dCas9-VP64 (refs. 26,27) and dCas9-VPR (containing VP64-p65-Rta).[30] A modified version of sg362F was developed, sg362F (SAM), consisting of two MS2 RNA aptamer-binding motifs embedded into the stemloop and loop 2 of the sgRNA. For the SAM activation complex to function,[33] the activation helper protein MS2-p65-HSF1 was included in parallel. All combinations tested containing the dCas9-VPR module showed robust 1.5- to 2-fold activation when compared to sg362F/dCas9-VP64, which was the set point of the experiment (**Figure 5b**). The sg362 (SAM) guide provided a weaker signal compared to sg362F, but full activation was restored with the MS2-p65-HSF1 helper protein. Nevertheless, the fully reconstituted dCas9-VPR/sg362F(SAM)/MS2-p65-HSF1 complex showed no improved activation relative to sg362F/dCas9-VPR alone, suggesting that dCas9-VPR is functioning at already optimal levels for this target. No activation was observed at all for experiments involving sg362F or sg362F (SAM) with the activator dCas9-p300 (ref. 31; data not shown).

CRISPR/dCas9-mediated activation of latent HIV-1 is specific and functions independently of NF-κB activity

In order to establish the specificity of sg362F for its target, we applied sg362F and sg359R together with nuclease-active Cas9 against a LTR target in transfected HEK293 cells. Double-stranded breaks were induced by sg362F and sg359F, but not sgCTRL, as determined with the T7 Endonuclease I assay (**Figure 6a**). Cleavage was not observed in control transfections using an LTR comprising a 26-bp deletion of the sg359F/sg362F target region (ΔLTR). Since CEM LChIT 3.2 cells are bimodal, we determined whether sg362F and an upstream target guide sg223R would affect TAT-mediated feedback in the presence of a nuclease-defective dCas9 (**Figure 6b**). Guide sg362F, but not sg223R, reduced native mCherry reporter levels indicating that the sg362F/dCas9 complex prevents processive transcription. Since sg362F fully occupies the sequence across both NF-κB–binding sites, we next determined if CRISPR activation via the sg362F target site was coupled to NF-κB function. To establish this, we transfected sg362F, sg223R, and sgCTRL together with dCas9 lacking a transactivation domain in the presence of increasing concentrations of TNFα in CEM LChIT 3.2 cells (**Figure 6c**). Binding of sg362F/dCas9 abrogated TNFα-mediated mCherry activation, suggesting that the bound sg362F/dCas9 complex is capable of preventing NF-κB–mediated transactivation. The sg223R/dCas9 complex, which binds upstream of any NF-κB–binding site, did not block TNFα-mediated activation. Since sg223R/dCas9-VP64 was capable of inducing modest activation (**Figures 6c** and **2b**), these data strongly support a conclusion where sg362F/dCas9-VP64 functions independently of NF-κB as a stand-alone transactivator of LTR activation.

CRISPR/dCas9-mediated activation is specific and does not induce cellular off-target effects

Chemical LRAs such as HDACi and NF-κB agonists are nonspecific, resulting in widespread global gene expression. As such,

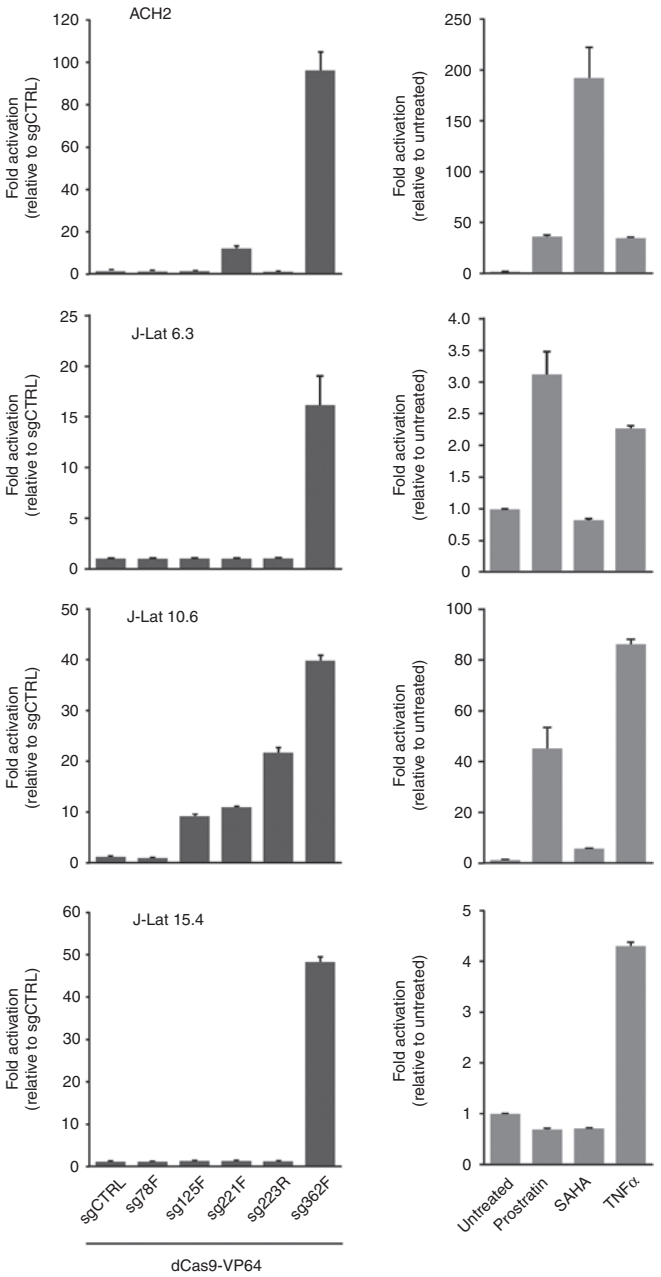

Figure 3 The activation efficacy of dCas9-VP64 and LTR-targeting sgRNAs in multiple in vitro latency models. A panel of LTR-targeting sgRNAs together with dCas9-VP64 was tested for HIV-1 activation efficacy in multiple latency cell models including ACH2 cells and several J-Lat clonal cell lines including J-Lat 6.3, J-Lat 10.6, and J-Lat 15.4. Cells were co-transfected with individual LTR-targeting sgRNAs or a control sgCTRL, dCas9-VP64 and a mCherry-expressing plasmid. In ACH2 cells, total RNA was extracted from cells 72 hours posttransfection and activation of HIV-1 gene expression was measured by qRT-PCR. In J-Lat clonal cell lines, activation was measured by flow cytometry. Transfected cells were gated on mCherry expression and GFP levels were measured as an indication of LTR activity 72 hours posttransfection. Graphed values represent the geometric mean of GFP expression normalized to sgCTRL + SD in cells treated in triplicate. In parallel, each latency cell model was treated with drug compounds known to activate latent HIV-1. Each respective cell line was treated with tumor necrosis factor-α (TNFα), prostratin, and suberoylanilide hydroxamic acid (SAHA). Graphed values represent the geometric mean of mCherry expression normalized to untreated cells + SD in cells treated in triplicate. qRT-PCR, quantitative real-time PCR.

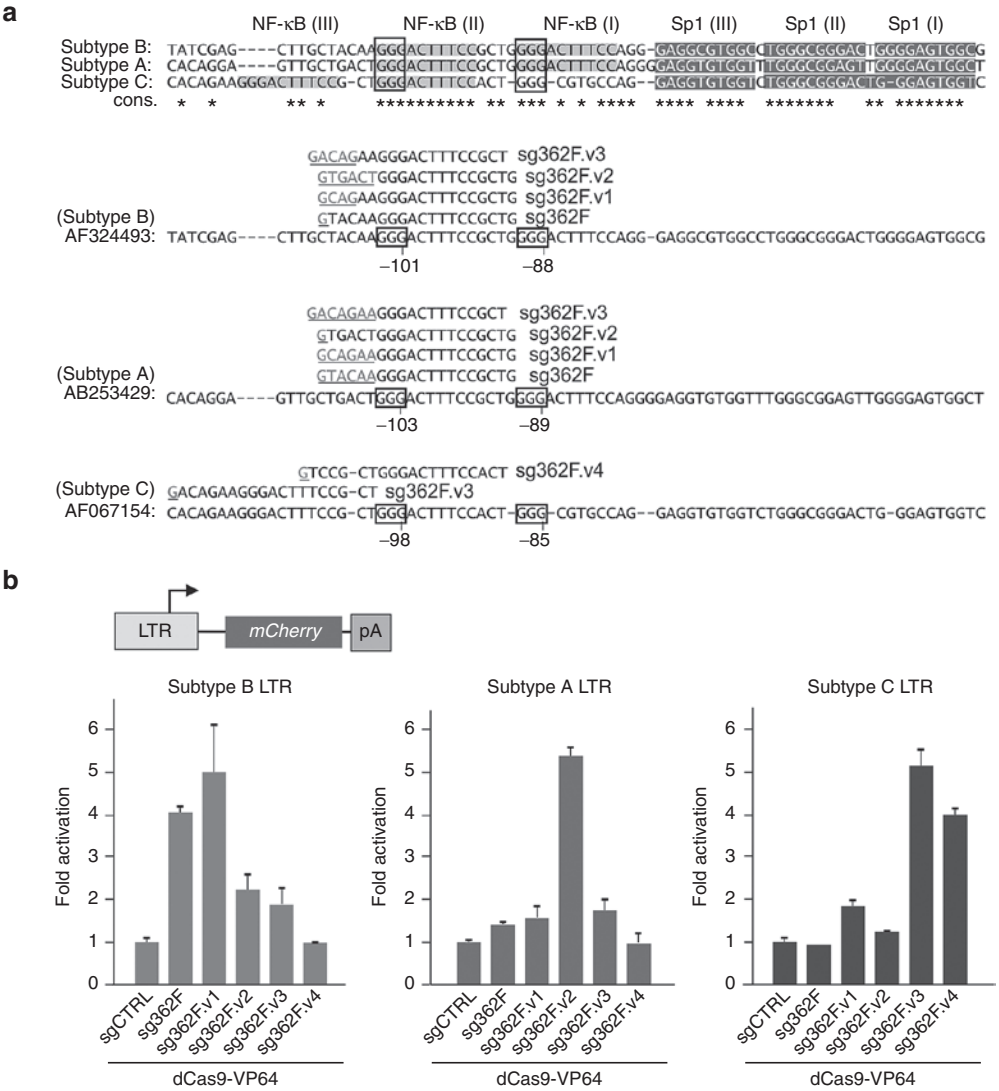

Figure 4 The effects of spatial and sequential modifications on sgRNA efficacy. (**a**) The sequences of the LTR region containing the sg362F target site are depicted for three different viral subtypes: subtypes B, A, and C. Sequences are aligned. NF-κB–binding motifs and Sp-1–binding motifs shaded. The GGG PAM sequences are boxed. Variants of sg362F (sg362F.v1, sg362F.v2, sg362F.v3, or sg362.v4) that retain matched sequence complementarity with representative sequences of the B, A and C HIV subtypes were designed and generated. These variants are shown aligned with the sequence of each respective viral subtype and both the conserved and mismatched regions are depicted (mismatched sequences are shown in blue text). (**b**) HEK293T cells were co-transfected with the sg362F, and variants of the sg362F or a control sgCTRL together with dCas9-VP64, a subtype-specific LTR-driven mCherry reporter, and a GFP-expressing plasmid. Activation was measured 72 hours posttransfection by flow cytometry. Transfected cells were gated on GFP expression and mCherry levels were measured as an indication of LTR activity. Graphed values represent the geometric mean of mCherry expression normalized to sgCTRL + SD in cells treated in triplicate.

a major concern is the risk, especially long-term, of toxic or unwanted effects. To determine if any adverse effects caused by sg362F/dCas9-VP64 could be detected, CEM LChIT 3.2 cells were monitored for cellular toxicity and for perturbations to the cell cycle. Cells were either untransfected, transfected with sg362F and sgCTRL, or treated with different latency reactivation agents. For the cellular toxicity assay, puromycin was used as a positive control (**Figure 7a**); however, no further adverse effects could be detected for any of the treated cells. Cell cycle analysis was determined by propidium iodide staining followed by flow cytometry. Not one of sg362F/dCas9-VP64, sgCTRL/dCas9-VP64, or TNFα affected cell cycle progression, whereas prostratin increased S and decreased G2/M phases (**Figure 7b**). To further analyze the specificity of the sg362F-guided CRISPR activator system, we assessed the changes in host gene expression by RNA deep sequencing analysis. TNFα-treated CEM LChIT 3.2 cells were compared to sg362F/dCas9-VP64 transfected cells and control sgCTRL/dCas9-VP64 (**Figure 7c**,**d**). Cells treated with sg362F potently activated the LChIT-reporter (mCherry signal). No other transcript was dysregulated by more than twofold. As expected, TNFα impacts the transcript levels of many more genes than the sg362F/dCas9-VP64 with genes directly related to NF-κB signaling such as *NFKB2* and *RelB*[40] (together with LChIT/mCherry) being potently upregulated (**Figure 7c**). CD83 is a marker for T-cell activation,[41] and the observed presence of upregulated *CD83* points to TNFα's potential T-cell–activating role not observed

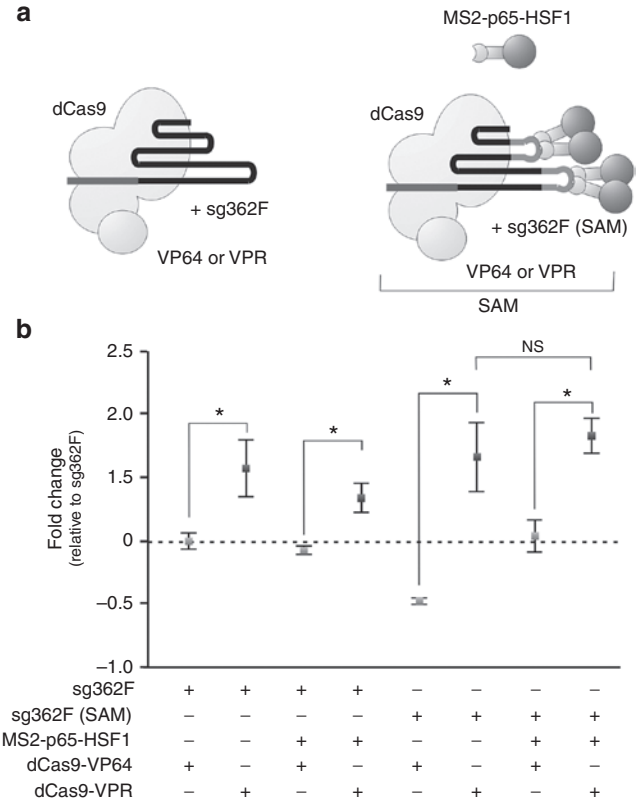

Figure 5 Modified CRISPR-dCas9 activation systems for enhanced LTR activation. (**a**) We made use of dCas9 fused to a C-terminal VP64 or VPR transactivation domain. By utilizing minimal hairpin aptamers that selectively bind dimerized MS2 bacteriophage coat proteins, the synergistic activation mediator (SAM) system was developed for enhanced activation. This system recruits multiple activation domains such as the NF-κB trans-activating subunit, p65, and the activation domain from human heat-shock factor 1 (HSF1) as fusion proteins with MS2 to a single sgRNA-targeted dCas9-VP64 bound locus. (**b**) The activation properties of the different transactivation domains (VP64 or VPR) in the context of the different activation systems were assessed. Different combinations of activation system components were co-transfected into CEM LChIT 3.2 cells as illustrated together with a GFP-expressing plasmid. Activation was measured 72 hours posttransfection by flow cytometry. Transfected cells were gated on GFP expression and mCherry levels were measured as an indication of LTR activity. Graphed values were normalized to cells transfected with sgCTRL and subsequently made relative to activation mediated by dCas9-VP64 and sg362F to set base line activation (set as zero). Values represent the mean + SD of cells treated in triplicate. *$P < 0.05$, one-way ANOVA and Bonferroni's test. ANOVA, analysis of variance; NS, not significant.

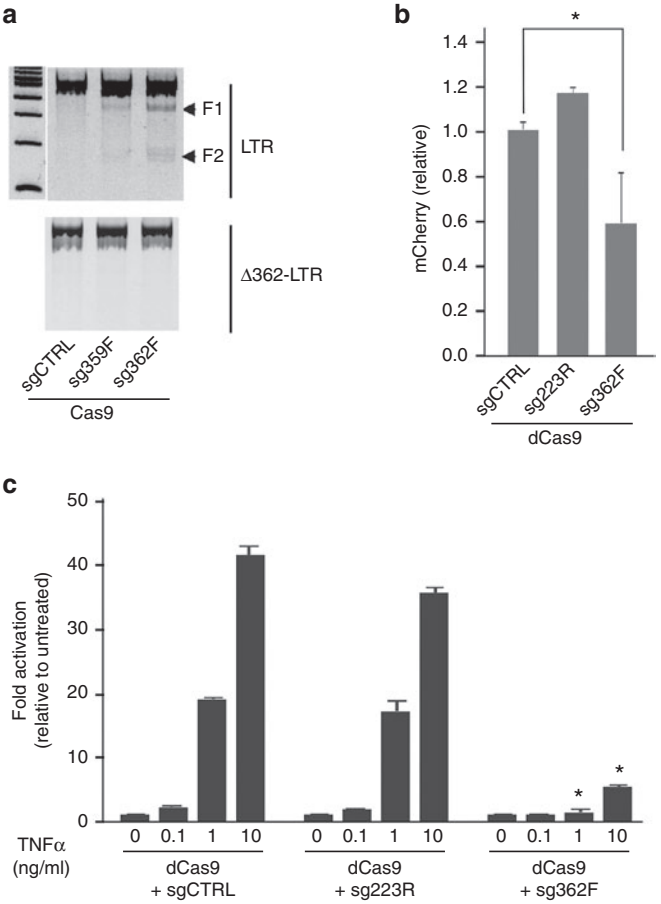

Figure 6 CRISPR/dCas9-mediated activation is specific and functions independently of NF-κB. (**a**) The T7E1 assay was used to determine targeted specificity of sg362F compared with upstream sg359F and a non-targeting control (sgCTRL) in transiently transfected HEK293 cells using a nuclease active Cas9. ΔLTR lacks a 26 bp sequence targeted by sg359F and sg362F. (**b**) CEM LChIT 3.2 cells were transfected with dCas9-GFP and each of sgCTRL, sg362F, and sgLTR223R. Transfected cells were gated on GFP expression and mCherry levels were measured as an indication of LTR activity. Graphed values were normalized to cells transfected with sgCTRL. Values represent the mean + SD of cells treated in triplicate. *$P = 0.031$, one-way ANOVA and *post hoc* test. (**c**) CEM LChIT 3.2 cells transfected (as in **b**) were treated with different TNFα concentrations. Graphed values in each panel were normalized to the relevant transfected cells treated with 0 ng/ml TNFα. Values represent the mean + SD of cells treated in triplicate. *$P < 0.05$, one-way ANOVA and Bonferroni's test. ANOVA, analysis of variance.

in sg362F/dCas9-VP64 transfected cells. Lastly, the density distribution of genes affected by TNFα is significantly different to sg362F/dCas9-VP64 ($P < 10^{-16}$), underscoring the specificity of the CRISPR-based targeted activation approach (**Figure 7d**).

DISCUSSION

The mechanisms underlying the establishment and maintenance of viral latency are multifaceted and poorly understood. One possible theory is that the replicative state of the infected T cell plays a significant role in controlling viral transcription; however, recent findings point to the viral TAT as a master regulator of viral activity.[42] Cell state–dependent establishment of latent reservoirs holds that latency arises as infected, activated T cells return to a resting, memory state because HIV transcription depends largely on host transcription factors.[43] Because of the scarcity of these latently infected cells in patient populations, the mechanisms underlying latency have been most often studied using *in vitro* primary cell models. Generation of a latent state in these models has required culturing infected cells under conditions that return them to a quiescent state. Doing so achieves heterochromatin formation and subsequent epigenetic silencing of the viral promoter.[44] Recently though, models in which a latent state was achieved in actively proliferating T cells have been developed, suggesting that latency may in fact be independent of the replicative state of the cell.[45] For example, the Planelles lab has shown that cells induced

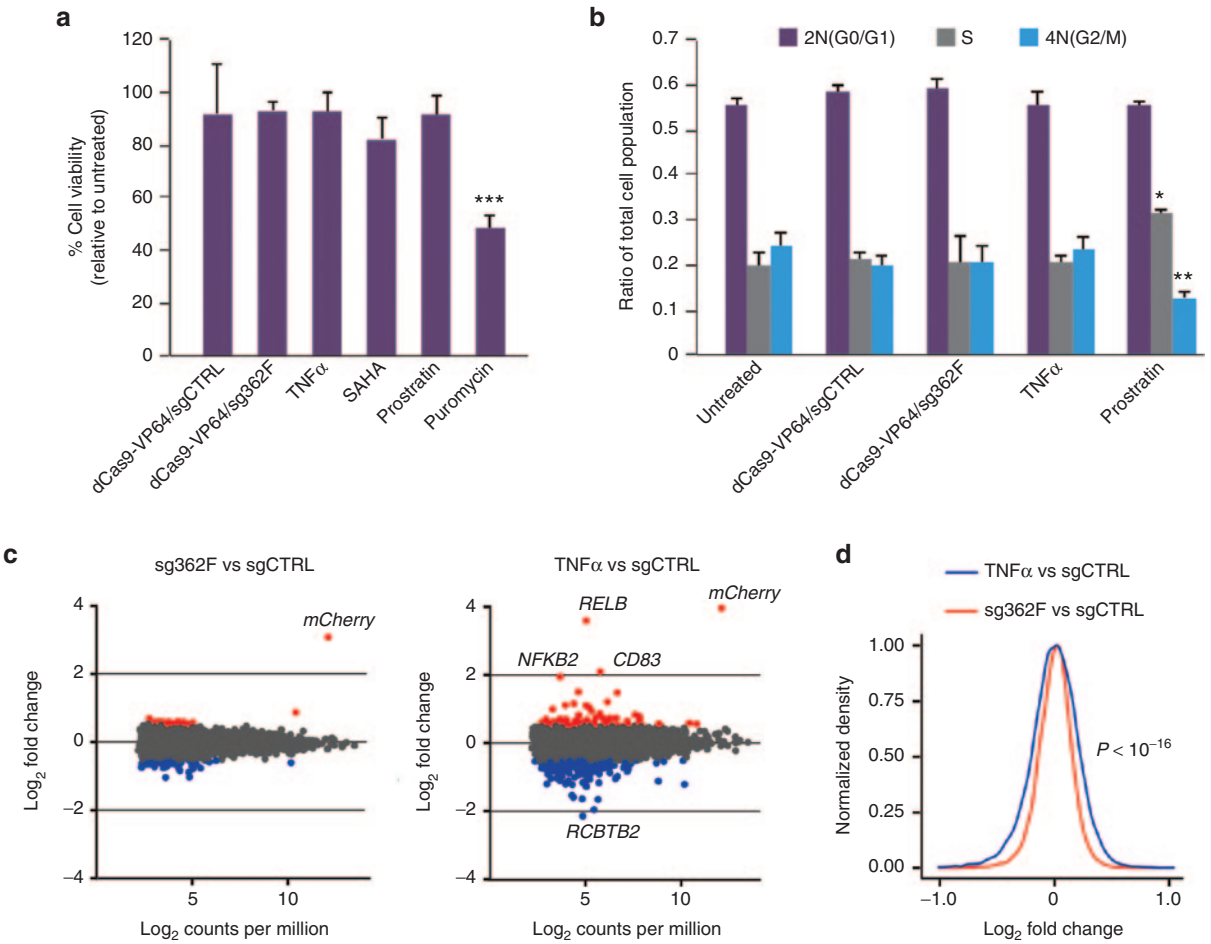

Figure 7 CRISPR/dCas9-mediated activation does not result in off-target effects. (**a**) HEK293T cells were co-transfected with dCas9-VP64 together with sg362F or control sgCTRL, or treated with the latency reversal agents TNFα, SAHA or prostratin, or treated with puromycin. Cell viability was analyzed 72 hours posttreatment. (**b**) CEM LChIT 3.2 bimodal cells were co-transfected with dCas9-VP64 together with sg362F or control sgCTRL, or treated with the latency reversal agents TNFα or prostratin. 2N, S, and 4N phases were identified in treated cells 72 hours posttreatment. For **b** and **c**, values represent the mean + SD of cells treated in triplicate. *$P = 0.0217$; **$P = 0.0067$; ***$P < 0.001$ by one-way ANOVA and Dunnett's test. (**c**) RNA sequencing analysis of cells activated by dCas9-VP64 together with the sg362F vs. TNFα activation. MA plots of sg362F (left) and TNFα (right) activation demonstrating the specific activation of mCherry by sg362F as compared with TNFα activation of NF-κB in addition to mCherry. CEM LChIT 3.2 cells were co-transfected with dCas9-VP64 together with the sg362F or a control sgCTRL and a GFP-expressing plasmid. Cells transfected with the sgCTRL were cultured in the presence of TNFα or left untreated. Transfections were performed in triplicate for each condition. Transfected cells were FACS sorted for GFP expression 72 hours posttransfection and total RNA was isolated from the sorted cells for sequencing library preparation and RNA sequencing. Transcripts above +0.5 log2FC and below −0.5 log2FC were shaded in red and blue, respectively. (**d**) Histogram of observed fold changes observed for dCas9-VP64 together with the sg362F and TNFα activation showing decreased variation for sg362F as compared with TNFα. For the modified robust Brown-Forsythe Levene-type test was applied to the distributions of log2FC for the two samples and the variances are significantly different ($P < 10^{-16}$; ref. 58). ANOVA, analysis of variance; FACS, fluorescence activated cell sorting.

to proliferate via IL-7 can propagate latent viruses by cell division, in the absence of viral reactivation from latency.[46] An analysis of more than 200 infected patient samples discovered a large subset of the latent reservoir for which the transcriptional activity of the virus is not dependent on T-cell activity.[47] In addition, the viral LTR of these integrated proviruses showed no CpG methylation, suggesting that latency may in fact be under the control of a stochastic TAT feedback loop.[47] Recent modeling of HIV latency, using both mathematical and TAT-inducible systems, strongly supports this notion of the TAT-feedback loop acting as a master regulator of viral state.[48] Instead of latency being primarily dependent on cellular host factors, the findings argue that the TAT feedback loop plays the predominate role in toggling the virus between active and latent states. The latter holds true for cells comprising functional P-TEFb complex. That an intrinsic viral program autonomously regulates viral latency independently of the cellular state may prove problematic to current proposed treatment strategies that depend largely on the use of cellular transcription factors and epigenetic regulators to activate latent virus. Should a large subset of the latent reservoir be found to exist independent of cell state, such treatments may prove largely ineffective. Because of this, targeting the viral promoter using a CRSIPR activation approach may prove to be a more efficient long-term treatment strategy as activation is decoupled from cellular activity.

A major challenge in designing treatment strategies to efficiently target latent reservoirs has remained the lack of models

that accurately represent the cellular environment of latent reservoirs.[36] The primary challenge in developing such models stems from the close association between HIV transcriptional activity and T-cell state.[44] Most *in vitro* models have aimed, and often failed, to accurately represent the transition from active to resting state due to the limitation of cell culture conditions. The primary challenge stems from the fact that once activated, the preponderance of cultured T cells undergoes apoptosis producing few cells that achieve quiescence. Groups have attempted to inhibit apoptosis through a variety of mechanisms, yet in some of these latency models, infected cells are shown to express the activation markers such as CD69 and CD25, suggesting that these cells may not have achieved a quiescent state.[44] This variability in the cellular activation state of various models provides a possible reason behind the ineffectiveness of a single chemical activator to achieve complete activation of a latent population across multiple models (**Figure 3** and **Supplementary Figure S1**). Furthermore, it reiterates the effectiveness of the CRISPR activation system described which directly and specifically activates HIV proviral gene expression in a wide range of latency models tested.

Ongoing studies continue to define parameters that enable the optimal design of systems with increased potency, yet we have identified a sgRNA, sg362F, that is not only potent across a plethora of latency cell models but also highly specific in its actions by not involving cellular activation pathways. By uniquely targeting sequences that incorporate the viral NF-κB–binding sites doublet, sg362F blocks NF-κB transactivation and likely functions independently of NF-κB activity. Furthermore, this system does not induce any cellular toxicity nor does it perturb cell cycle progression. These elements all represent highly desirable features in the quest for a novel latency-purging agent. In conclusion, we have identified and characterized a highly potent latency activator using the CRISPR activation machinery and unique sgRNA that target the viral enhancer sequence. While more is needed to develop a targeted delivery approach and functional analysis of latency reactivation in a clinical setting, CRISPR activation represents a promising tool for the specific and targetable reactivation of latent HIV reservoirs.

MATERIALS AND METHODS

Plasmid construction. Generic sgRNA-expressing plasmids, pcDNA.H1sgRNA and pcDNA.H1sgRNA (SAM), were constructed using full-length gBlocks (IDT, Coralville, IA) comprising an H1 promoter, chimeric guide RNA backbone, and sgRNA cloning site (**Supplementary Data**). The gBlocks were cloned by Gibson Assembly into the *Bgl*II/*Bbs*I site of pcDNA3.1(+) (Thermo Fisher Scientific, MA) and contain *Bsm*BI sites allowing for facile cloning of sgRNAs by oligo annealing as described previously (**Supplementary Data**).[19] To generate subtype-specific LTR reporter vectors, full-length 3′ LTR promoters from HIV subtype B (Genbank accession: AF324493), subtype A (AB253429), and subtype C (AF067154) were amplified from LGIT plasmid templates[49] and cloned by Gibson Assembly into the *Pci*I/*Age*I digested pmCherry-C1 (Clontech, CA) using the following primers: LTR_F 5′-TGGCCTTTTGCTGGCCTTTTGCTCAACTAGTGACTTACAAGGCAGCTGTAGATC-3′ and LTR_R 5′- CTCACCATGGTGGCGACCGGTAGCGACCACAGGAAACAGCTATGACCATGATTA-3′. The dCas9-activator vectors include: pAC154-dual-dCas9VP160-sgExpression (Addgene plasmid #48240),[29] pHAGE EF1α dCas9-VP64 (Addgene plasmid #50918),[35] SP-dCas9-VPR (Addgene plasmid #63798),[30] and lenti-MS2-P65-HSF1_Hygro (Addgene plasmid #61426).[33] The following reagent was obtained through the NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH: pNL4-3.Luc.R$^-$.E$^-$ from Dr Nathaniel Landau.[50,51]

LChIT 3.2 latency model in CEM cells. The pCLChIT plasmid was derived from pCLGIT (CMV-LTR-GFP-IRES-Tat).[34] The GFP gene in pCLGIT was excised with *BamH*I-*EcoR*I, and the mCherry PCR fragment was inserted using the same restriction sites. The lentiviral plasmid for LChIT (pCLChIT) was packaged and harvested in HEK293T cells using 10 µg of vector, 5 µg pMDLg/pRRE, 3.5 µg pVSV-G, and 1.5 µg pRSV-Rev, as previously detailed.[49] Viral supernatant was collected 48 hours after transfection and passed through a 0.45-µm filter to remove cell debris. The virus was then loaded onto a 20% (wt/wt) sucrose cushion and concentrated by ultracentrifugation in an SW28 rotor on Optima XL-100K Ultracentrifuge (Beckman Coulter) for 1.5 hours at 25,000 rpm at 4 °C. The viral pellet was resuspended in 100 µl of phosphate-buffered saline (pH 7.0). To ensure that cells were not transduced with multiple copies of the LChIT vector, cells were transduced at low multiplicity of infection (>0.1), which corresponded to less than 10% of the total cells having a single LChIT integration event. One week after transduction, cells were treated with TNFα to stimulate TAT (and mCherry) expression; 1 day later, the "On" (mCherry+) population was sorted with fluorescence activated cell sorting. The polyclonal LChIT "On" cells were cultured for 2 weeks, in which 16.6% of the "On" cells relaxed into the "Off" (mCherry-) state. Individual clones were isolated from the polyclonal "Off" subpopulation by fluorescence activated cell sorting single cells into 96-well plates. After 4 weeks of culturing, clones were analyzed by flow cytometry, and the LChIT 3.2 clone was identified.

Cell culture. HEK293T cells were maintained in Dulbecco's Modified Eagle Medium (Mediatech, VA) supplemented with 10% fetal bovine serum (Life Technologies, CA) and 50 µg/ml Pen/Strep (Mediatech) at 37 °C and 5% CO_2. ACH2 and J1.1 cells are chronically infected cell lines harboring the HIV-1 LAI strain. ACH2 and J1.1 cell lines were maintained in RPMI 1640 (Mediatech) supplemented with 10% fetal bovine serum and 50 µg/ml Pen/Strep at 37 °C and 5% CO_2. The following reagents were obtained through the NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH: J-Lat Full Length Clones 6.3, 8.4, 9.2, 10.6, and 15.4, henceforth referred to as J-Lat cells, from Dr Eric Verdin.[52] The LChIT cell line is a CEM T-cell–based reporter system, comprising a single copy of HIV that encodes mCherry-IRES-Tat from the full-length HIV LTR (LChIT). J-Lat and LChIT cell lines were maintained in RPMI 1640 supplemented with 10% fetal bovine serum, 2 mmol/l L-glutamine, and 50 µg/ml Pen/Strep at 37 °C and 5% CO_2.

Transient transfection assays and treatment with latency reversal agents. To screen sgRNAs targeted to the HIV-1 5′LTR, HEK293T cells were seeded at 40,000 cells per well in a 48-well plate, 24 hours prior to transfection. Cells were transfected in triplicate using Lipofectamine 2000 (Thermo Fisher Scientific) with 50 ng of pNL4-3.Luc.R$^-$E$^-$,[50,51] 200 ng of a Pol-III expressed sgRNA vector, and 200 ng of pAC154-dual-dCas9VP160-sgExpression or pHAGE EF1α dCas9-VP64 (Addgene plasmid #48240 and #50918, respectively). Fifty nanograms of a plasmid expressing *Renilla* luciferase (pRL-CMV) was included as a background control (Promega, WI). At 48 hours posttransfection, a luciferase assay was performed using the Dual-Luciferase Reporter Assay System (Promega), and the levels of Firefly luciferase were normalized to those of *Renilla* luciferase.

To assess the efficacy of LTR-sgRNAs to activate HIV-1 gene expression across various latency cell models, the following cell lines were transfected: CEM LChIT 3.2 cells, ACH2 cells, J1.1 cells, and J-Lat cells. For RNA sequencing experiments, CEM LChIT 3.2 cells were transfected. Two million cells were co-transfected in triplicate with 5 µg of pHAGE EF1α dCas9-VP64 (Addgene plasmid #50918) and 5 µg of the sgRNA expressing plasmid using the Neon Transfection System (Thermo Fisher Scientific) with the following electroporation parameters: three pulses of

1,350 V and 10 ms at a cell density of 2×10^7/ml. To evaluate transfection efficiency in the fluorescent reporter cell lines, 500 ng of either a GFP or a mCherry expressing plasmid was co-transfected into the LChIT or J-Lat cells, respectively.

The effects of spatial and sequential modifications on sgRNA efficacy were determined in HEK293T cells, which were seeded at 120,000 cells per well in a 24-well plate, 24 hours prior to transfection. Cells were transfected in triplicate using Lipofectamine 2000 with 400 ng of pHAGE EF1α dCas9-VP64 (Addgene plasmid #50918), 400 ng of the sgRNA-expressing plasmid, 100 ng pmCherry-C1-LTR reporter, and 100 ng of pcDNA-eGFP.

Modified activation systems were tested in CEM LChIT 3.2 cells. Two million CEM LChIT 3.2 cells were co-transfected in triplicate with 5 μg of pHAGE EF1α dCas9-VP64 (Addgene plasmid #50918) or 5 μg SP-dCas9-VPR (Addgene plasmid #63798), 5 μg of the sgRNA-expressing plasmid, 5 μg lenti MS2-P65-HSF1_Hygro (Addgene plasmid #61426) or pUC19, and 1 μg of a GFP-expressing plasmid pcDNA-eGFP, using the Neon Transfection System with the following electroporation parameters: three pulses of 1,350 V and 10 ms at a cell density of 2×10^7/ml. For experiments involving nonactivating dCas9, CEM LChIT 3.2 cells were transfected with 5 μg pX461B.GFP (dCas9) and 5 μg of sg362F or sg223R or sgCTRL. Similar transfection conditions were used. Forty-eight hours posttransfection, cells were treated with 0 (untreated), 0.1 ng/ml, 1 ng/ml, and 10 ng/ml of TNFα.

To evaluate the efficacy of chemical stimuli to activate latent cell line models, cells were treated for 72 hours with the following latency reversal agents: 10 ng/ml of TNFα (Thermo Fisher), 1 μmol/l SAHA (Sigma-Aldrich, MO), or 3 μmol/l 12-deoxyphorbol-13-acetate (Prostratin; Sigma-Aldrich).

Gene expression analysis by flow cytometry and quantitative real-time PCR. Fluorescent data for the CEM LChIT 3.2 cell line, J-Lat clones, and HEK293T cells were acquired by flow cytometry on the BD LSR II Flow Cytometer (BD Biosciences, Franklin Lakes, NJ) 72 hours posttransfection. A minimum of 10,000 events were collected per sample, and the data was analyzed using FlowJo vX.0.7 software (Tree Star, Ashland, OR). To account for transfection efficiency, CEM LChIT 3.2 and HEK293T cells were gated on GFP expression and J-Lat cells were gated on mCherry expression.

To determine HIV expression in ACH2 and J1.1 cell lines, total cellular RNA was isolated using the RNeasy Mini kit (QIAGEN, CA) 72 hours posttransfection. RNA samples were DNase treated using the Turbo DNA-free Kit (Thermo Fisher Scientific). DNase-treated RNA samples were then standardized and reverse transcribed with Mu-MLV (Life Technologies) using an oligo-dT/random nonamer primer mix. Quantitative real-time PCR was carried out using Kapa Sybr Fast universal qPCR mix (Kapa Biosystems, MA) on an Eppendorf Mastercycler realplex. Thermal cycling parameters started with 3 minutes at 95 °C, followed by 40 cycles of 95 °C for 3 seconds and 60 °C for 30 seconds. Specificity of the PCR products was verified by melting curve analysis. The following primers were used for qPCR: HIV F: 5′- AGGGATGGAAAGGATCACCAGCAA-3′; HIV R: 5′- CCCACCTCAACAGATGTTGTCTCA-3′; β-actin F: 5′- AGGTCA TCACCATTGGCAATGAG-3′; β-actin R: 5′- TCTTTGCGGATGTCC ACGTCA-3′.

T7 endonuclease I assay. HEK293T cells were transfected with plasmids expressing sg362F, sg369F, and sgCTRL together with target plasmids pLTR-mCherry or pΔLTR-mCherry. At 48 hours posttransfection, cells were lysed using a KAPA express extract kit (KAPA Biosystems, MA). PCR was performed on the lysate with LTR-specific primers LTR-F2 5′-TTGACAGCCGCCTAGCATTTCA-3′ and LTR-R2 5′-CACACAGGAAACAGCTATGACCATGATTA-3′. The assay was performed using T7 endonuclease I (New England Biolabs, MA) as described previously.[19]

RNA sequencing. CEM LChIT 3.2 bimodal cells were co-transfected with the plasmid pHAGE EF1α dCas9-VP64 together with sg362F or control sgCTRL and a GFP-expressing plasmid as described above. Cells transfected with the sgCTRL were cultured in the presence of TNFα (10 ng/ml) or left untreated to serve as a control. Transfections were performed in triplicate for each condition. To control for transfection efficiency, cells were sorted by fluorescence activated cell sorting based on GFP expression using the BD FACSAria II 72 hours posttransfection. Total RNA was isolated from ~25,000–30,000 sorted cells using the Promega Maxwell 16 LEV Simply RNA Cell Kit (Promega). The sequencing library was prepared using the NEBNext Ultra RNA Library Prep Kit reagent. Sequencing was performed on the Illumina HiSeq2500 platform, in a 1 × 50 bp single-read configuration in Rapid Run mode, with a total of at least 120 million reads per lane. Raw, as well as processed, data are available online (Gene Expression Omnibus accession number GSE72259). Raw reads were mapped to GRCh38.P2 with the LChIT sequence appended using the STAR aligner.[53] Aligned reads were counted using htseq[54] utilizing the Gencode version 22 gene annotations[55,56] with an annotation of the LChIT sequence appended. Analysis of differential expression was performed using edgeR[57] after filtering the data such that all samples had a minimum of 20 reads per gene.

Cell cycle analysis. CEM LChIT 3.2 bimodal cells were co-transfected with the plasmid pHAGE EF1α dCas9-VP64 together with sg362F or control sgCTRL and a GFP-expressing plasmid as described above, or treated with the latency reversal agents as described before. Seventy-two hours following treatment, cells were fixed in 70% ethanol at 4 °C overnight, washed in phosphate-buffered saline, and then stained at room temperature for 1 hour with propidium iodide staining solution (1% Triton X-100, 20 μg/ml propidium iodide). A minimum of 75,000 cells were collected using the BD LSR II flow cytometer (BD Biosciences). Data were analyzed using FlowJo vX.0.7 software. To remove doublets and debris, cells were first gated using forward scatter and side scatter. Cell cycle kinetic data were generated using the FloJo univariate cell cycle algorithm to identify the 2N, S, and 4N phases.

Toxicity assay. Ten thousand HEK293T cells were seeded into a 96-well plate and 24 hours later co-transfected with the plasmid pHAGE EF1α dCas9-VP64 together with sg362F or control sgCTRL, or treated with the latency reversal agents as described before, or treated with puromycin (500 ng/ml). Seventy-two hours posttreatment, cell viability was analyzed using the ab112119 Cell Cytotoxicity Assay Kit (Fluorometric) (Abcam, Cambridge, MA). Reactions were performed in quadruplicate.

Statistical analysis. All experiments were performed in triplicate. Results are presented as the mean plus SD. Statistical analysis was performed using Prism 5 software (Graph Pad, San Diego, CA). One-way analysis of variance followed by Bonferroni's or Dunnett's multiple comparison test was performed to evaluate the statistical significance of the treatments. A result of $P < 0.05$ was considered to be statistically significant.

SUPPLEMENTARY MATERIAL

Figure S1. The activation efficacy of dCas9-VP64 and LTR-targeting sgRNAs in multiple *in vitro* latency models J1.1, J-Lat 8.4 and J-Lat 9.2.
Supplementary Data

ACKNOWLEDGMENTS

M.S.W. acknowledges support from the Strategic Health Innovation Partnerships (SHIP) Unit of South African Medical Research Council (SAMRC), with funds received from the South African Department of Science and Technology (DST). Support for K.V.M. is acknowledged from NIAID PO1 AI099783-01, R01 AI111139-01, R01 DK104681-01, and Australian Research Council FT1300100572. The authors declare no conflict of interest.

REFERENCES

1. Palella, FJ Jr, Delaney, KM, Moorman, AC, Loveless, MO, Fuhrer, J, Satten, GA *et al.* (1998). Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV Outpatient Study Investigators. *N Engl J Med* **338**: 853–860.
2. Siliciano, JD, Kajdas, J, Finzi, D, Quinn, TC, Chadwick, K, Margolick, JB *et al.* (2003). Long-term follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting CD4+ T cells. *Nat Med* **9**: 727–728.
3. Crooks, AM, Bateson, R, Cope, AB, Dahl, NP, Griggs, MK, Kuruc, JD *et al.* (2015). Precise quantitation of the latent HIV-1 reservoir: implications for eradication strategies. *J Infect Dis* **212**: 1361–1365.
4. Archin, NM and Margolis, DM (2014). Emerging strategies to deplete the HIV reservoir. *Curr Opin Infect Dis* **27**: 29–35.
5. Siliciano, JD, Lai, J, Callender, M, Pitt, E, Zhang, H, Margolick, JB *et al.* (2007). Stability of the latent reservoir for HIV-1 in patients receiving valproic acid. *J Infect Dis* **195**: 833–836.
6. Archin, NM, Eron, JJ, Palmer, S, Hartmann-Duff, A, Martinson, JA, Wiegand, A *et al.* (2008). Valproic acid without intensified antiviral therapy has limited impact on persistent HIV infection of resting CD4+ T cells. *AIDS* **22**: 1131–1135.
7. Sagot-Lerolle, N, Lamine, A, Chaix, ML, Boufassa, F, Aboulker, JP, Costagliola, D *et al.*; ANRS EP39 study. (2008). Prolonged valproic acid treatment does not reduce the size of latent HIV reservoir. *AIDS* **22**: 1125–1129.
8. Archin, NM, Espeseth, A, Parker, D, Cheema, M, Hazuda, D and Margolis, DM (2009). Expression of latent HIV induced by the potent HDAC inhibitor suberoylanilide hydroxamic acid. *AIDS Res Hum Retroviruses* **25**: 207–212.
9. Archin, NM, Keedy, KS, Espeseth, A, Dang, H, Hazuda, DJ and Margolis, DM (2009). Expression of latent human immunodeficiency type 1 is induced by novel and selective histone deacetylase inhibitors. *AIDS* **23**: 1799–1806.
10. Archin, NM, Liberty, AL, Kashuba, AD, Choudhary, SK, Kuruc, JD, Crooks, AM *et al.* (2012). Administration of vorinostat disrupts HIV-1 latency in patients on antiretroviral therapy. *Nature* **487**: 482–485.
11. Contreras, X, Schweneker, M, Chen, CS, McCune, JM, Deeks, SG, Martin, J *et al.* (2009). Suberoylanilide hydroxamic acid reactivates HIV from latently infected cells. *J Biol Chem* **284**: 6782–6789.
12. Edelstein, LC, Micheva-Viteva, S, Phelan, BD and Dougherty, JP (2009). Short communication: activation of latent HIV type 1 gene expression by suberoylanilide hydroxamic acid (SAHA), an HDAC inhibitor approved for use to treat cutaneous T cell lymphoma. *AIDS Res Hum Retroviruses* **25**: 883–887.
13. Rasmussen, TA, Tolstrup, M, Brinkmann, CR, Olesen, R, Erikstrup, C, Solomon, A *et al.* (2014). Panobinostat, a histone deacetylase inhibitor, for latent-virus reactivation in HIV-infected patients on suppressive antiretroviral therapy: a phase ½, single group, clinical trial. *Lancet HIV* **1**: e13–e21.
14. Bullen, CK, Laird, GM, Durand, CM, Siliciano, JD and Siliciano, RF (2014). New *ex vivo* approaches distinguish effective and ineffective single agents for reversing HIV-1 latency *in vivo*. *Nat Med* **20**: 425–429.
15. Laird, GM, Bullen, CK, Rosenbloom, DI, Martin, AR, Hill, AL, Durand, CM *et al.* (2015). *Ex vivo* analysis identifies effective HIV-1 latency-reversing drug combinations. *J Clin Invest* **125**: 1901–1912.
16. Reardon, B, Beliakova-Bethell, N, Spina, CA, Singhania, A, Margolis, DM, Richman, DR *et al.* (2015). Dose-responsive gene expression in suberoylanilide hydroxamic acid-treated resting CD4+ T cells. *AIDS* **29**: 2235–2244.
17. Mojica, FJ, Díez-Villaseñor, C, García-Martínez, J and Almendros, C (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* **155**: 733–740.
18. Mojica, FJ, Díez-Villaseñor, C, García-Martínez, J and Soria, E (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J Mol Evol* **60**: 174–182.
19. Cong, L, Ran, FA, Cox, D, Lin, S, Barretto, R, Habib, N *et al.* (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**: 819–823.
20. Jinek, M, East, A, Cheng, A, Lin, S, Ma, E and Doudna, J (2013). RNA-programmed genome editing in human cells. *Elife* **2**: e00471.
21. Mali, P, Yang, L, Esvelt, KM, Aach, J, Guell, M, DiCarlo, JE *et al.* (2013). RNA-guided human genome engineering via Cas9. *Science* **339**: 823–826.
22. Doudna, JA and Charpentier, E (2014). Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* **346**: 1258096.
23. Weinberg, MS and Morris, KV (2014). A new world order: tailored gene targeting and regulation using CRISPR. *Mol Ther* **22**: 893.
24. Gilbert, LA, Larson, MH, Morsut, L, Liu, Z, Brar, GA, Torres, SE *et al.* (2013). CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154**: 442–451.
25. Qi, LS, Larson, MH, Gilbert, LA, Doudna, JA, Weissman, JS, Arkin, AP *et al.* (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**: 1173–1183.
26. Mali, P, Aach, J, Stranges, PB, Esvelt, KM, Moosburner, M, Kosuri, S *et al.* (2013). CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat Biotechnol* **31**: 833–838.
27. Maeder, ML, Linder, SJ, Cascio, VM, Fu, Y, Ho, QH and Joung, JK (2013). CRISPR RNA-guided activation of endogenous human genes. *Nat Methods* **10**: 977–979.
28. Perez-Pinera, P, Kocak, DD, Vockley, CM, Adler, AF, Kabadi, AM, Polstein, LR *et al.* (2013). RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat Methods* **10**: 973–976.
29. Cheng, AW, Wang, H, Yang, H, Shi, L, Katz, Y, Theunissen, TW *et al.* (2013). Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. *Cell Res* **23**: 1163–1171.
30. Chavez, A, Scheiman, J, Vora, S, Pruitt, BW, Tuttle, M, P R Iyer, E *et al.* (2015). Highly efficient Cas9-mediated transcriptional programming. *Nat Methods* **12**: 326–328.
31. Hilton, IB, D'Ippolito, AM, Vockley, CM, Thakore, PI, Crawford, GE, Reddy, TE *et al.* (2015). Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat Biotechnol* **33**: 510–517.
32. Tanenbaum, ME, Gilbert, LA, Qi, LS, Weissman, JS and Vale, RD (2014). A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell* **159**: 635–646.
33. Konermann, S, Brigham, MD, Trevino, AE, Joung, J, Abudayyeh, OO, Barcena, C *et al.* (2015). Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* **517**: 583–588.
34. Weinberger, LS, Burnett, JC, Toettcher, JE, Arkin, AP and Schaffer, DV (2005). Stochastic gene expression in a lentiviral positive-feedback loop: HIV-1 Tat fluctuations drive phenotypic diversity. *Cell* **122**: 169–182.
35. Kearns, NA, Genga, RM, Enuameh, MS, Garber, M, Wolfe, SA and Maehr, R (2014). Cas9 effector-mediated regulation of transcription and differentiation in human pluripotent stem cells. *Development* **141**: 219–223.
36. Spina, CA, Anderson, J, Archin, NM, Bosque, A, Chan, J, Famiglietti, M *et al.* (2013). An in-depth comparison of latent HIV-1 reactivation in multiple cell model systems and resting CD4+ T cells from aviremic patients. *PLoS Pathog* **9**: e1003834.
37. Roach, HI, Bronner, F, and Oreffo, ROC (2011). *Epigenetic Aspects of Chronic Diseases.* Springer: London.
38. Emiliani, S, Van Lint, C, Fischle, W, Paras, P Jr, Ott, M, Brady, J *et al.* (1996). A point mutation in the HIV-1 Tat responsive element is associated with postintegration latency. *Proc Natl Acad Sci USA* **93**: 6377–6381.
39. Naghavi, MH, Schwartz, S, Sönnerborg, A and Vahlne, A (1999). Long terminal repeat promoter/enhancer activity of different subtypes of HIV type 1. *AIDS Res Hum Retroviruses* **15**: 1293–1303.
40. Bren, GD, Solan, NJ, Miyoshi, H, Pennington, KN, Pobst, LJ and Paya, CV (2001). Transcription of the RelB gene is regulated by NF-kappaB. *Oncogene* **20**: 7722–7733.
41. Prazma, CM, Yazawa, N, Fujimoto, Y, Fujimoto, M and Tedder, TF (2007). CD83 expression is a sensitive marker of activation required for B cell and CD4+ T cell longevity *in vivo*. *J Immunol* **179**: 4550–4562.
42. Zack, JA, Arrigo, SJ, Weitsman, SR, Go, AS, Haislip, A and Chen, IS (1990). HIV-1 entry into quiescent primary lymphocytes: molecular analysis reveals a labile, latent viral structure. *Cell* **61**: 213–222.
43. Siliciano, RF and Greene, WC (2011). HIV latency. *Cold Spring Harb Perspect Med* **1**: a007096.
44. Hakre, S, Chavez, L, Shirakawa, K and Verdin, E (2012). HIV latency: experimental systems and molecular models. *FEMS Microbiol Rev* **36**: 706–716.
45. Jeeninga, RE, Westerhout, EM, van Gerven, ML and Berkhout, B (2008). HIV-1 latency in actively dividing human T cell lines. *Retrovirology* **5**: 37.
46. Bosque, A, Famiglietti, M, Weyrich, AS, Goulston, C and Planelles, V (2011). Homeostatic proliferation fails to efficiently reactivate HIV-1 latently infected central memory CD4+ T cells. *PLoS Pathog* **7**: e1002288.
47. Ho, YC, Shan, L, Hosmane, NN, Wang, J, Laskey, SB, Rosenbloom, DI *et al.* (2013). Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure. *Cell* **155**: 540–551.
48. Razooky, BS, Pai, A, Aull, K, Rouzine, IM and Weinberger, LS (2015). A hardwired HIV latency program. *Cell* **160**: 990–1001.
49. Burnett, JC, Lim, KI, Calafi, A, Rossi, JJ, Schaffer, DV and Arkin, AP (2010). Combinatorial latency reactivation for HIV-1 subtypes and variants. *J Virol* **84**: 5958–5974.
50. Connor, RI, Chen, BK, Choe, S and Landau, NR (1995). Vpr is required for efficient replication of human immunodeficiency virus type-1 in mononuclear phagocytes. *Virology* **206**: 935–944.
51. He, J, Choe, S, Walker, R, Di Marzio, P, Morgan, DO and Landau, NR (1995). Human immunodeficiency virus type 1 viral protein R (Vpr) arrests cells in the G2 phase of the cell cycle by inhibiting p34cdc2 activity. *J Virol* **69**: 6705–6711.
52. Jordan, A, Bisgrove, D and Verdin, E (2003). HIV reproducibly establishes a latent infection after acute infection of T cells *in vitro*. *EMBO J* **22**: 1868–1877.
53. Dobin, A, Davis, CA, Schlesinger, F, Drenkow, J, Zaleski, C, Jha, S *et al.* (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**: 15–21.
54. Anders, S, Pyl, PT and Huber, W (2015). HTSeq–a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**: 166–169.
55. Engström, PG, Steijger, T, Sipos, B, Grant, GR, Kahles, A, Rätsch, G *et al.*; RGASP Consortium. (2013). Systematic evaluation of spliced alignment programs for RNA-seq data. *Nat Methods* **10**: 1185–1191.
56. Steijger, T, Abril, JF, Engström, PG, Kokocinski, F, Hubbard, TJ, Guigó, R *et al.*; RGASP Consortium. (2013). Assessment of transcript reconstruction methods for RNA-seq. *Nat Methods* **10**: 1177–1184.
57. Robinson, MD, McCarthy, DJ and Smyth, GK (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**: 139–140.
58. Brown, MB, and Forsythe, AB (1974). Robust tests for equality of variances. *J Am Stat Assoc* **69**: 364–367.