dCas9 targeted chromatin and histone enrichment for mass spectrometry (Catchet-MS) identifies IKZF1 as a novel drug-able target for HIV-1 latency reversal

Authors:

Enrico Ne[1], Raquel Crespo[1], Ray Izquierdo-Lara[1], Selin Kocer[1], Alicja Gorszka[1], Thomas van Staveren[1], Tsung Wai Kan[1], Dick Dekkers[2], Casper Rokx[3], Panagiotis Moulos[4], Pantelis Hatzis[4], Robert-Jan Palstra[1], Jeroen Demmers[2], Tokameh Mahmoudi[1]*

Affiliations:

1 Department of Biochemistry, Erasmus University Medical Center, Ee622 PO Box 2040

3000CA Rotterdam, The Netherlands

2 Proteomics Center, Erasmus University Medical Center, Ee679a PO Box 2040

3000CA Rotterdam, The Netherlands

3 Department of Internal Medicine, Section of Infectious Diseases, Erasmus University Medical Center, Rg-530, PO BOC 2040 3000CA Rotterdam, The Netherlands

4 Fleming Genomics Facility, Biomedical Sciences Research Center "Alexander Fleming", 16672, Vari, Greece

*Correspondence to Tokameh Mahmoudi, t.mahmoudi@erasmusmc.nl

Summary

A major pharmacological strategy toward HIV cure aims to reverse latency in infected cells as a first step leading to their elimination. While the unbiased identification of molecular targets physically associated with the latent HIV-1 provirus would be highly valuable to unravel the molecular correlates of HIV-1 transcriptional repression and latency reversal, due to technical limitations, this has not been possible. Here we use dCas9 targeted chromatin and histone enrichment strategy coupled to mass spectrometry (Catchet-MS) to isolate the latent and activated HIV-1 5'LTR, followed by MS identification of the differentially locus-bound proteins. Catchet-MS identified known and novel latent 5'LTR-associated host factors. Among these, IKZF1 is a novel HIV-1 transcriptional repressor, required for Polycomb Repressive Complex 2 recruitment to the LTR. We find the drug iberdomide, which targets IKZF1 for degradation, to be a clinically advanced novel LRA that reverses HIV-1 latency in CD4+T-cells isolated from virally suppressed HIV-1 infected participants.

INTRODUCTION

Combination antiretroviral therapy (cART) effectively halts HIV replication and has significantly reduced AIDS-associated mortality. However, cART is not curative, it has side-effects, and apart from the costs of lifelong therapy, the global roll-out of cART, particularly in resource-limited countries, remains an ongoing challenge (UNAIDS fact sheet 2019). HIV persists because subsequent to stable integration into the CD4+ T cell host genome, the provirus can remain in a nonproductive latent state, defined by the absence of HIV-1 gene expression. Because of this reservoir of latently HIV-1 infected cells, interruption of cART leads to a rapid rebound of unrestricted viral replication, necessitating life-long treatment (Siliciano and Siliciano, 2015).

Therapeutic strategies for HIV cure aim to eliminate, inactivate, or reduce the pool of latently infected cells such that the patient's immune system can control viral replication upon cessation of cART. As quiescent memory CD4+ T cells, which constitute the main cellular reservoir of latent HIV infected cells, have a long half-life (Siliciano et al., 2003), pharmacological approaches aim to speed up the decay rate of this infected reservoir. One such strategy is to induce viral expression in latently infected cells using latency reversal agents (LRAs) to increase the viral transcriptional output and viral protein production, rendering the infected cell recognizable to the immune system or susceptible to viral cytopathic effects for elimination.

At the molecular level, the expression of the HIV-1 genome is determined by the activity of the HIV-1 promoter or 5'LTR, which is controlled by the LTR chromatin landscape (Rafati et al., 2011), the engagement of sequence-specific host transcription factors (TFs) and associated cofactors (Ne et al., 2018; Pereira et al., 2000), the recruitment of RNA polymerase II (Pol II) and its efficient transcriptional elongation (Mousseau and Valente, 2017; Ott et al., 2011). Co-transcriptionally recruited host factors then mediate post-transcriptional processing of the nascent HIV-1 RNA template for its efficient splicing, export, and translation (Chun et al., 2003; Karn and Stoltzfus, 2012; Lassen et al., 2006; Rojas-Araya et al., 2015). Host factors mediating these regulatory steps that control HIV-1 gene expression are thus the main targets of ongoing pharmacological strategies to reverse latency (Stoszko et al., 2019). However, due to the complex and heterogeneous nature of latency, the currently available LRAs are incapable of reactivating a significant portion of cells carrying a latent provirus (Battivelli et al., 2018), have limited capacity to induce viral protein expression (Grau-Exposito et al., 2019; Grau-Exposito et al., 2017), and thus have failed to significantly impact the reservoir in patients (Kim et al., 2018). Additionally, LRAs target host molecular complexes with widespread regulatory functions and are therefore prone to non-specific pleiotropic effects. The identification of the full repertoire of regulatory factors and cofactors involved in silencing the latent HIV 5'LTR would be critical for a better understanding of the molecular determinants of proviral latency and for the development of a more specific and effective combination of therapeutics to deplete the HIV reservoir. However, due to technical limitations, unbiased proteomic analysis of host proteins associated with the latent or active HIV promoter, in infected cells, has never been conducted.

Locus-specific strategies to identify in vivo chromatin-bound protein complexes, at a single genomic location, are largely based on a reverse chromatin immunoprecipitation (reverse-ChIP) protocol which relies on the introduction of an affinity bait at the chromatin region of interest, followed by purification of the bait and proteomic analysis of the protein complexes that co-purify

with it (Byrum et al., 2012; Pourfarzad et al., 2013). A related approach, Proteomics of isolated chromatin fragments (PICh), uses biotinylated oligonucleotides that hybridize with the region of interest to specifically purify the target DNA and associated proteins (Dejardin and Kingston, 2009; Ide and Dejardin, 2015). With the introduction of CRISPR-based methods and nuclease-deficient cas9 (dCas9) fusion constructs, reverse-ChIP strategies have become less labor-intensive and more versatile, as introduction of an artificial consensus sequence for targeting of the bait to the region of interest is no longer required (Liu et al., 2017; Waldrip et al., 2014). Additionally, CRISPR-dCas9 strategies can also be combined with, biotin-based, proximity labeling techniques, such as Bio-ID (Roux et al., 2012; Schmidtmann et al., 2016) and APEX enzyme fusion constructs (Myers et al., 2018), that, by relying on the capture of in vivo biotin-labeled proteins, located in the proximity of the dCas9 construct, do not require cell crosslinking.

Locus-specific isolation techniques represent an enormous challenge in protein biochemistry and proteomics, due to the fact that, in each cell, the number of proteins bound at a specific regulatory region is exceedingly lower than the number of proteins bound to the rest of the genome. Consequently, the minimum amount of input required for mass spectrometry identification of proteins that bind to a specific genomic locus of interest, in a single copy, as would be the case for the integrated latent HIV-1 promoter, is estimated to be of at least half-a-billion cells, (Gauchier et al., 2020; Vermeulen and Dejardin, 2020). As the cellular platform for latency used in this study harbors only one copy of the integrated HIV-1 5'LTR per cell, we generated sonicated chromatin from approximately 3 billion cells per condition to improve our purification yield.

While conventional reverse-ChIP approaches for locus-specific proteomics allow the identification of some locus-specific factors, many of these strategies suffer from high signal to noise ratio and low abundance of single loci, rendering the identification of proteins bound to the region of interest challenging (Byrum et al., 2012; Chalkley and Verrijzer, 2004; Hamperl et al., 2014; Pourfarzad et al., 2013; Tsui et al., 2018; Waldrip et al., 2014) (Gauchier et al., 2020). Another caveat in locus-specific proteomics techniques is that there is typically a large fraction of the exogenously expressed bait that is present diffusely in the cells and that is not exclusively bound to the locus of interest. Evidence suggests that this fraction of the bait, which is inevitably captured during the purification pipeline, accumulates in the RNA rich nucleolar compartment (Chen et al., 2013; Ma et al., 2016); thus, significantly contributing to the background signal, affecting sensitivity and representing an additional barrier for reliable identification of locus-specific complexes.

A well-established strategy for background elimination is subtraction from the final dataset using a robust set of negative controls (Gauchier et al., 2020; Liu et al., 2017). This approach effectively eliminates some background and has successfully led to the identification of several locus-specific proteins (Gauchier et al., 2020; Liu et al., 2017). This however does not fully solve the problem as contaminant proteins are not removed from the sample, thus improving sensitivity, but from the final dataset, using a control-guided data filtering procedure. Given the evidence for Cas9 bait localization, this would likely result in the subtraction of nucleolar associated proteins and proteins interacting transiently with the diffusely expressed fraction of the bait, stabilized upon crosslinking. We, therefore, reasoned that a biochemical purification pipeline, specifically designed to eliminate this source of background from the sample, would represent a more effective strategy to reduce the complexity of the sample and of the identified hits and improve the sensitivity of the technique.

Here, we combined a CRISPR-dCas9 based reverse-ChIP approach with a sequential purification step, targeting histones to biochemically remove the unwanted background, prior to semi-quantitative Mass Spectrometry identification; a purification pipeline we refer to as “dCas9 targeted chromatin and histone enrichment for mass spectrometry’’ (Catchet-MS). Addition of histone affinity purification as the last step in Catchet-MS enabled specific enrichment and identification of the chromatin-bound fraction of the dCas9 bait proteins bound to the target region, while non-localized bait complexes in the cytoplasm and nucleoplasm were eliminated. We verified the specificity of our dCas9 targeting targeting by ChIP-sequencing and based our hits selection exclusively on proteins that were distinctly identified to be bound to the latent versus the transcriptionally activated HIV 5’LTR. As potential unwanted background is associated to both experimental conditions, the distinct profile offers a stringent solution for its removal.

Using Catchet-MS we identify proteins preferentially associated with the latent or active HIV-1 promoter that represent potential key regulators of HIV-1 gene expression. Among these regulators, we demonstrate that the transcription factor IKZF1 is a novel transcriptional repressor of the HIV-1 promoter, required for the sequential recruitment of the polycomb repressive complex 2 (PRC2) and critical for maintenance of a repressive LTR chromatin environment which is characterized by the presence of the H3K27me3 chromatin mark. Additionally, we demonstrate that iberdomide (CC-220), a therapeutically advanced drug, which targets the IKZF1 protein for degradation through modulation of CRBN dependent ubiquitination, leads to HIV-1 latency reversal in both ex vivo infected primary CD4+ T cells and cells isolated from HIV-1 infected aviremic individuals on cART. We demonstrate that Catchet-MS provides a powerful approach towards enrichment and mass spec identification of molecular effectors bound to a single DNA locus. Our data suggest that modulation of IKZF1 protein expression with iberdomide, or other thalidomide derived drugs, represents a viable target for therapeutic intervention to reverse latency in context of HIV-1 cure strategies.

RESULTS

Catchet-MS, a method to isolate and identify locus-bound protein complexes at the latent and activated HIV-1 LTR in vivo

We established a CRISPR/dCas9-based proteomics strategy to identify a comprehensive repertoire of protein complexes associated in vivo with the latent and active HIV promoter (**Figure 1A-C**). As a cellular platform for latency we used J-Lat 11.1, a Jurkat derived T cell line, which harbors a single latent copy of the integrated full-length HIV-1 proviral genome, defective in *env*, and in which GFP replaces the *nef* coding sequence as a reporter for viral activation (Jordan et al., 2003) (**Figure 1A**). J-Lat 11.1 cells were first modified to stably express a multiple epitopes (V5, HA, FLAG)-tagged dCas9 bait, and an 18 nucleotide single guide-RNA, targeting the hypersensitive site 2 (HS2) region of the HIV-1 promoter, downstream of the Transcription Start Site (TSS) (**Figure 1C**). We confirmed the expression of the tagged dCas9 bait construct by Western blotting, using antibodies specifically recognizing dCas9 and the construct tags HA, V5, and FLAG (**Figure 1D**). We chose to target our dCas9 bait to the HS2 region downstream of Nuc-1, as recent evidence pointed to transcription elongation and RNA processing as critical steps involved in HIV-1 latency in vivo (Yukl et al., 2018). As a negative control, we generated cells expressing HA-V5-FLAG-dCas9 and a non-targeting control gRNA (nt-gRNA). We validated HIV-1 promoter targeting of the gRNA (HS2-gRNA) sequence in a functional assay, by co-transfection of J-Lat 11.1 cells together with a dCas9-VP64-p65-Rta (dCas9-VPR) fusion construct containing the VPR transcriptional activation domain. Confirming correct targeting of the dCas9-VPR construct to the HIV-1 promoter, co-transfection of the HS2 gRNA, but not the nt-gRNA, resulted in an increase in HIV-1 promoter-driven GFP positive cells (**Supplementary Figure 1A**). We additionally examined the specific binding of the HA-V5-FLAG-dCas9 bait at the HS2 region of the HIV-1 promoter in latent J-Lat 11.1 cells by ChIP-qPCR analysis using a V5 epitope-based IP. As shown in **Figure 1E**, enrichment of the HA-V5-FLAG-dCas9 bait was observed specifically over HIV-1 LTR sequences spanning Nuc-1 and Nuc-2, only in the HS2-gRNA guided cells and not in those expressing the non-targeting control gRNA (nt-gRNA) (**Figure 1E**).

To determine the optimal HA-V5-FLAG-dCas9 bait IP enrichment strategy, we examined specific enrichment of the locus in ChIP-qPCR experiments comparing antibodies directed against the HA, V5, and FLAG epitopes (**Supplementary Figure 1B and C**). While antibodies directed against each tag resulted in efficient and specific enrichment of the targeted HIV-1 DNA region, we chose to use anti V5 affinity beads targeting the V5 epitope as the first step in the large-scale locus enrichment strategy as it demonstrated both optimal purification yield as well as the highest signal to noise ratio (**Supplementary Figure 1C**). To ensure that targeting of the HA-V5-FLAG-dCas9 bait downstream of the HIV-1 promoter would not interfere functionally with viral reactivation, we examined LTR activity in the latent and PMA activated HA-V5-FLAG-dCas9 HS2-gRNA guided J-Lat 11.1 cells. As shown by Flow cytometry and fluorescence microscopy, J-Lat 11.1 cells or those harboring HS2-gRNA guided or non-targeted HA-V5-FLAG-dCas9 bait displayed low basal GFP expression and a similar increase in GFP positive cells in response to PMA treatment (**Figure 1F and Supplementary Figure 1D**). Thus, the targeted dCas9 bait does not interfere with the maintenance of latency or capacity for reactivation. From this, we infer that binding of 5'LTR associated transcription regulatory factors bound in the repressed or activated states is unaltered by dCas9 targeting. We also confirmed that the binding capacity of the HA-V5-

FLAG-dCas9 bait is not affected by the transcriptional state of the 5'LTR by ChIP-qPCR in latent and PMA stimulated cells (**Figure 1G**); HA-V5-FLAG-dCas9 was similarly enriched over the sequences spanning LTR Nuc-1 and Nuc-2 in both the transcriptionally repressed and active states, confirming that the established platform is suitable for purification of the locus.

The large number of non-localized bait molecules (e.g. dCas9) present a caveat in reverse-ChIP approaches for single locus chromatin proteomics, as they are also captured through the purification pipeline, thus preventing reliable identification of the locus associated bait interactome within the background signal from purified unspecific complexes. It has been shown that Cas9 molecules, when non-efficiently guided are diffusely present in the nucleus and tend to accumulate in the RNA-rich nucleolar compartment (Chen et al., 2013; Ma et al., 2016). In our experimental system, cell fractionation followed by western blotting demonstrated a large fraction of the HA-V5-FLAG-dCas9 bait molecules to be present in the cytoplasmic fraction, despite the presence of a nuclear localization sequence (NLS) within the bait construct (**Figure 2A, 2B**). To obtain a purer chromatin fraction, we modified our in-house ChIP protocol to include 10 cycles of low-efficiency sonication after nuclear isolation followed by denaturation in the presence of strong denaturing buffers (adapted from Chromatin enrichment for proteomics protocol (CheP) (Kustatscher et al., 2014). We named the modified protocol Chromatin enrichment for immunoprecipitation (CheIP). Mass spectrometry analysis of individual steps of our purification pipeline demonstrated progressive removal of membrane, cytoplasmic and organelle associated proteins as well as, in part, non-chromatin bound nuclear proteins by CheIP, resulting in the enrichment of chromatin and associated proteins within the insoluble nuclear material (**Supplementary Figure 2A, 2B**). While this biochemical procedure enables a more stringent purification of the chromatin fraction, a substantial amount of non-chromatin associated dCas9 bait complexes remain within the precipitated chromatin fraction, presenting substantial background.

To improve the specificity of the purification, we took advantage of the fact that what distinguishes cross-linked chromatin-associated complexes from non-chromatin associated complexes is the presence of histones within the complex. Therefore, we introduced a histone enrichment step into our purification pipeline; given the high abundance of histones within the purified chromatin fraction, we opted to perform histone affinity purification following the V5 epitope-based purification of the HA-V5-FLAG-dCas9 bait-containing cross-linked complexes (**Figure 2C),** thus eliminating the need for large amounts of histone antibodies. We tested this pipeline in a sequential (V5/histone) ChIP experiment, followed by qPCR and western blotting. As shown in **Figure 2D** (top panel), ChIP, using anti V5 affinity beads efficiently enriched for the HIV-1 promoter region targeted by the HA-V5-FLAG-dCas9. Western blotting of the ChIP decrosslinked eluate confirmed efficient bait recovery by immunoprecipitation **Supplementary Figure 3A**. The second sequential immunoprecipitation, using anti H3- anti H2B antibodies conjugated beads, while retaining the HIV-1 target locus of interest (**Figure 2D)** (bottom panel), results in a concomitant loss of HA-V5-FLAG-dCas9 bait at the protein level (**Supplementary Figure 3A**). Relative abundance of histone proteins is reduced during the V5 enrichment step while, as expected, they are retained after the histone H3/H2B enrichment step (**Supplementary Figure 3B**).

The specificity of the guided dCas9 is a critical point for locus purification. We, therefore, performed ChIP-sequencing experiments to comprehensively investigate the locus specificity of our bait, which is guided and predicted to specifically bind to the integrated HIV-1 5'LTR. ChIP-sequencing was performed using the sequentially purified chromatin, first immunoprecipitated using V5 affinity beads, eluted and re-immunoprecipitated using Histone H3-H2B affinity purification. Both in ChIP-seq (**Figure 2 E**) and its confirmation by ChIP qPCR (**Figure 2F**) we observe a strong enrichment on the HIV-1 5' LTR while enrichment at human chromosomes, aside from centromeric artifacts common to ChIP-seq and a few minor off-target peaks (**Supplementary Table 1**), is low. Next, we specifically monitored the recovery of the top potential dCas9 off-target binding sites as predicted for the HS2 gRNA sequence (5'-GAAGCGCGCACGGCAAG-3') by the online tool http://www.rgenome.net/cas-offinder/. The top predicted off-target, located at chr1:42,957,793 , allows 2 mismatches in the sequence, while all subsequent off-targed locations predicted allow 3 or more mismatches in the sequence. Zooming in on these predicted off-target regions in the ChIP-seq data (**Figure 2E, G**) and by ChIP qPCR (**Figure 2F**) demonstrated the absence of traceable peaks or any substantial PCR signal for these chromosomal loci. Our results demonstrate a highly specific enrichment of the targeted HIV-1 5'LTR at the DNA level, whose protein content was subjected to mass spectrometry, within the Catchet-MS pipeline.

Cachet-MS identifies known and novel host factors associated with the latent and active HIV-1 promoter

We scaled up the Catchet-MS workflow (**Figure 2C)**, to approximately 3 billion cells as starting material in the unstimulated and PMA stimulated conditions to model J-Lat 11.1 HIV-1 latency and activation. **Figure 3A** presents a heatmap of the mass spectrometry analysis performed at different steps of the Catchet-MS purification pipeline. To assess sample complexity prior to bait enrichment, 1% of the input chromatin material, prepared using the optimized CheIP protocol in both latent and activated states, was analyzed by mass spectrometry. As shown in **Figure 3A** the corresponding lanes in the heatmap show high sample complexity, characterized by a large number of high protein intensity values. The samples are enriched for nuclear and chromatin-associated proteins but also contain numerous cellular contaminants, resulting from carryover precipitation of insoluble, non-chromatin associated complexes (deposited Supplementary raw data). As expected, the abundance of the HA-V5-FLAG-dCas9 bait is low, in these input samples (**Figure 3B and C)**. By contrast, the assessment of sample complexity following the anti-V5 affinity purification step shows efficient enrichment for the HA-V5-FLAG-dCas9 bait-containing complexes. This stage of the purification, shown in the third and fourth lanes of **Figure 3A,** results in a marked reduction in sample complexity caused by the elimination of protein complexes that do not contain the HA-V5-FLAG-dCas9 bait concomitant with high enrichment of the HA-V5-FLAG-dCas9 bait (**Figure 3B and C**). Finally, re-immunoprecipitation using Histone H2B and H3 conjugated affinity beads leads to a clear difference in protein abundances between the unstimulated and PMA stimulated samples, as shown in the last two lanes of **Figure 3A**. Concomitantly, at this stage, we observe a loss in HA-V5-FLAG-dCas9 as well as, presumably, the corresponding bait-associated complexes not associated with histones (**Figure 3B**). As expected, given the known accumulation of non-locus bound Cas9 at the RNA-rich nucleolar

compartment (Chen et al., 2013; Ma et al., 2016), V5 affinity purification resulted in an increase in the number of nucleolar proteins detected, followed by their loss following the subsequent histone affinity purification step (**Supplementary Figure 3C**).

Analysis of the MS data from the final locus enrichment step resulted in a list of 678 proteins, identified with high confidence, Mascot score > 100. We then applied two stringent and unbiased selection criteria, Illustrated in **Figure 3D**, to narrow down the list of candidates by removing potential contaminants. First, we filtered out frequent contaminants using the Crapome database https://www.crapome.org/, a comprehensive database of contaminants commonly present in affinity purification coupled to mass spectrometry (AP-MS) experiments. As a second filter, we applied a GO cellular compartment and function analysis (GO-terms nucleus, chromatin-associated, histones bound, bound to transcription factors, DNA associated, or RNA associated) to include only nuclear proteins in our further analysis. Applying these criteria, we compiled a stringent list (n=38) of proteins predominantly associated with the HIV-1 promoter in the latent state (**Figure 3E and F**), a list (n=122) of common proteins, including the dCas9 bait and potential contaminants, found to associate with the HIV-1 5'LTR under both conditions (**Figure 3E and Supplementary Table 2**), and a list (n=84) of factors predominantly associated with the activated HIV-1 promoter following treatment with PMA (**Figure 3E and Supplementary Table 3).** Previously, it was reported that the latent provirus associates with the nuclear periphery (Dieudonne et al., 2009). Therefore, we re-analyzed our unfiltered data and selected the hits belonging to the "nuclear periphery" GO category (GO:0034399), which includes the subcategories "nuclear matrix" and "nuclear lamina" and indeed find Lamin B1 (LMNB1) among other nuclear body proteins to be more predominantly associated with the latent state (**Supplementary Figure 4**).

Functional validation of proteins identified to associate with the latent promoter

In search of putative molecular targets for therapeutic inhibition and HIV-1 latency reversal, we exclusively selected protein hits that were distinctly associated with the latent HIV-1 LTR, in vivo, comprising of 38 factors presented in the table in **Figure 3F**. As unwanted contaminants are associated to both experimental conditions, this represents an additional stringent solution for filtering out the potential background.

Of these latent HIV-1 LTR-associated candidates, a significant number (n=11) were previously reported to restrict viral expression (**Figure 3F** referenced in the table) and serve as positive controls that validate the quality of the experimental approach, providing confidence towards the veracity of the candidate list. We evaluated the effect of shRNA-mediated knockdown of a selection of the novel factors we identified, on the HIV-1 5'LTR-driven GFP reporter by flow cytometry and on the expression of HIV-1 genes Gag, pol, and Tat by RT-PCR in J-Lat 11.1 cells (**Figure 3F; Supplementary Figure 5**). Depletion of a number of the selected candidates (n=5 out of 10), namely HP1BP3, IKZF1, CDC73, DKC1, and PNN, resulted in significant ($p < 0.05$) upregulation of HIV-1 genes and an increase in the percentage of GFP positive cells (green boxes, **Figure 3F** and **Supplementary figure 5**), while for a small number of factors tested (N=3 out of 10), we did not observe a significant change in HIV-1 expression (yellow boxes, **Figure 3F**) and the remaining factors were found to be essential for cell viability (x mark, **Figure 3F**). Our

data suggested that the selected n=5 candidates are required for repression of HIV-1 expression in J-Lat 11.1 cells.

IKZF1 binds to the latent HIV-1 promoter and is repressive to HIV-1 transcription

Catchet-MS identified the IKAROS Family Zinc Finger 1 protein (IKZF1) as one of the factors to be exclusively associated with the repressed HIV-1 promoter (**Figure 3F**). IKZF1, a critical factor required for normal T cell development that can act both as a transcriptional repressor and activator (Geimer Le Lay et al., 2014; Ng et al., 2009; Zhang et al., 2011) has been reported to regulate gene expression through its association with the nucleosome remodeling and deacetylase complex (NuRD) and the positive transcription elongation factor (P-TEFb) (Bottardi et al., 2015; Davis, 2011; Marke et al., 2018; Schwickert et al., 2014; Zhang et al., 2011). Additionally, IKZF1 mediated gene silencing in T cells has been shown to be facilitated by its interaction with the Polycomb repressive complex 2 (PRC2), which promotes histone H3 lysine 27 trimethylation, a mark of transcriptionally silent chromatin (Oravecz et al., 2015). Consistent with its role as a repressive factor in T cells, depletion of IKZF1 in our functional testing screen using independent shRNA clones led to the strongest reactivation of HIV-1 expression (5 to 10 fold), as measured by the increase of GFP positive cells in flow cytometry and qRT-PCR for HIV-1 genes (**Figure 4A-C**). To ensure that the observed latency reversal is not subject to clonal effects in J-Lat 11.1 cells, we depleted IKZF1 in J-Lat A2 cells, harboring an integrated HIV 5' LTR driven GFP reporter and observed similar reversal of HIV-1 latency (**Supplementary Figure 9A**).

The IKZF family consists of several family members (Marke et al., 2018; Powell et al., 2019), although IKZF1 is the only family member detected in our mass spectrometry data. We assessed the expression of the other IKZF family members in Jurkat cells (RNA sequencing (Palstra et al., 2018)) and proceeded with shRNA mediated depletion of IKZF2 and 5, which are expressed in Jurkat cells (**Supplementary Figure 6A**). Depletion of IKZF2 and IKZF5 also resulted in HIV-1 latency reversal, although modestly compared to IKZF1 (**Supplementary Figure 6B**). These results are consistent with the lower levels of expression of IKZF2 and IKZF5 in T cells, compared to IKZF1, and with the well-established heterodimeric interactions among IKZF family members (Boutboul et al., 2018; Mullighan et al., 2009; Sun et al., 1996). We next sought to confirm and examine the association of IKZF1 with the latent HIV-1 LTR. Downstream of the 5'LTR, at position +818/+864 we identified a putative IKZF binding site, which is composed of a consensus TGGGAA/T sequence and two extra GGGA core sites (Li et al., 2011), (**Figure 4D)**. To assess IKZF1 binding at the predicted sites, we performed ChIP-qPCR experiments in latent and PMA treated J-Lat 11.1 cells using antibodies specific for IKZF1 (**Figure 4E**). In agreement with the Catchet-MS data, ChIP experiments in latent J-Lat 11.1 cells demonstrated IKZF1 enrichment between HIV-1 HS2 and Nuc2, (**Figure 4E**). Importantly, activation of HIV-1 transcription by PMA treatment abrogated IKZF1 binding (**Figure 4E and Supplementary Figure 7A**), consistent with the notion that IKZF1 may be important for maintenance of HIV-1 transcriptional repression. To confirm the specificity of IKZF1 binding, we also performed ChIP in control J-Lat 11.1 cells, infected with a scramble shRNA vector (Sh Control), and in cells infected with an shRNA vector targeting the IKZF1 mRNA (sh IKZF1). Our data show clear enrichment of IKZF1 binding over its predicted binding region in the sh Control cell line together with a dramatic loss of binding upon

shRNA mediated depletion in the IKZF1 depleted cells (**Figure 4F and Supplementary Figure 8A**), thus confirming the specificity of the antibody. To further control for the specificity of the IKZF1 signal we also monitored enrichment of IKZF1 at endogenous IKZF1 targets previously reported to be enriched for IKZF1 binding in B cells (Song et al., 2016) (**Supplementary Figure 8C**).

We then examined the effect of IKZF1 depletion on the chromatin state of the region by histone ChIP-qPCR experiments using antibodies against the active chromatin modification H3 lysine 4 trimethylation (H3K4me3), the repressive chromatin modifications H3 lysine 27 trimethylation (H3K27me3) and H3 lysine 9 trimethylation (H3K9me3), using the total histone H3 signal (**Figure 4G**) for normalization.

Upon IKZF1 depletion the chromatin state is characterized by a drastic loss of enrichment for the H3K27me3 mark (**Figure 4H; Supplementary Figure 8E**), a moderate increase of the H3K4me3 mark (**Figure 4I; Supplementary Figure S8F**) and a concomitant decrease in the repressive chromatin mark H3K9me3 (**Supplementary Figure 8G**). Similar data were obtained upon IKZF1 depletion in J-Lat A2 cells (**Supplementary Figure 9C-G**), thus confirming that the effects on chromatin are not subjected to clonal effects.

We compared the histone modification profile observed, with those obtained in latent J-lat 11.1 cells and upon PMA stimulation (**Supplementary Figure 7B-E**). Upon treatment with PMA, the chromatin state is also characterized by increased H3K4me3 (**Supplementary Figure 7C**) and a concomitant decrease in the repressive chromatin marks H3K27me3 (**Supplementary Figure S7D**) and H3K9me3 (**Supplementary Figure 7E**). IKZF1 depletion, however, led to a more prominent loss of enrichment for the H3K27me3 mark (**Figure 4H**).

In T cells, IKZF1 is necessary for mediating gene silencing through the recruitment of the Polycomb repressive complex 2 (PRC2) and the deposition of the H3K27me3 mark (Oravecz et al., 2015). Interestingly, our Catchet-MS experiment identified SUZ12, a core subunit of the PRC2 complex, and CBX8, a subunit of the Polycomb repressive complex 1 (PRC1) that acts as a reader of the H3K27me3 modification (Bracken et al., 2006; Malik and Hemenway, 2013; Vidal and Starowicz, 2017) to be also associated with the HIV-1 5'LTR in its latent state. We therefore conducted ChIP experiments for probing the IKZF-1 mediated binding of SUZ12 to the region to verify the hypothesis that IKZF1 may be necessary for recruitment of PRC2 complex to the HIV-1 5'LTR. We additionally examined the recruitment of CBX8 to check whether, the reduction in H3K27me3 mark deposition observed in absence of IKZF1, affects CBX8 binding. Consistent with this hypothesis, our data shows that depletion of IKZF1 leads to reduced enrichment of SUZ12 (**Figure 4J and Supplementary Figure 8B**) and CBX8 (**Figure 4K**) over the HS2 5'LTR region. Importantly, as shown by western blotting, depletion of IKZF1 does not lead to decreased levels of expression of SUZ12 and CBX8, suggesting that the effects observed are exclusively consequence of the absence of IKZF1 recruitment to the region (**Figure 4L**).

Targeting IKZF1 by iberdomide treatment reverses HIV-1 latency in ex vivo infected primary CD4+ T cells

Iberdomide (CC-220) and other thalidomide-derived, immunomodulatory drugs (IMIDs), have been shown to cause selective ubiquitination and degradation of IKZF1 and its related family

member IKZF3 (Kronke et al., 2014; Schafer et al., 2018). Thalidomide and its analogs, lenalidomide, and pomalidomide, are FDA approved drugs employed in the treatment for multiple myeloma (MM) leading to improved patient survival rates. The mechanism of action for IMIDs activity lies in their affinity for cereblon (CRBN) which is part of the cullin-ring finger ligase-4 cereblon ($CRL4^{CRBN}$) E3 ubiquitin ligase complex. Other members of the complex include cullin 4A (*CUL4A*), the DNA damage binding protein 1 (*DDB1*) and regulator of cullin 1 (*ROC1*). Iberdomide is a novel compound with higher affinity for CRBN and currently under clinical development for the treatment of systemic lupus erythematosus (SLE) and relapsed/refractory multiple myeloma (RRMM). Binding of iberdomide to CRBN modulates the E3 ubiquitin ligase activity of the complex, increasing its affinity for IKZF1 and IKZF3, and leads to their ubiquitination and proteasome-dependent degradation. We, therefore, set out to establish if iberdomide treatment in latent HIV-1 infected cells would induce degradation of IKZF1 and lead to HIV-1 latency reversal.

As the main reservoir of latent HIV-1 in infected individuals is resting CD4+ T cells, we ex vivo infected, without prior activation, CD4+ T cells obtained from healthy donors, with a defective full-length HIV-1 virus harboring a luciferase reporter to establish latent infections as shown in **Supplementary Figure 10A** (Lassen et al., 2012). Treatment of latent HIV-1 infected primary CD4+ T cells with iberdomide for 48 hours, resulted in moderate but significant ($p < 0.05$) reversal of latency, as observed by an increase in the mean luciferase activity compared to the untreated control (**Figure 5A**) in all 6 donor CD4+ T cells examined. Bromodomain end extra-terminal domain (BET) inhibitors were reported to act synergistically with the activity of IMIDs in treatment of refractory forms of multiple myeloma (Diaz et al., 2017; Moros et al., 2014). We, therefore, examined whether these compounds also synergize in context of HIV-1 latency reversal. Indeed, co-treatment of Iberdomide and JQ1 resulted in a robust and synergistic HIV-1 latency reversal (**Figure 5A**). To confirm that treatment with iberdomide results in degradation of IKZF1at the protein level, we performed western blot analysis of primary cells treated for 48hours with iberdomide. As expected, treatment with iberdomide, alone or in combination with JQ1, resulted in degradation of IKZF1 as shown in **figure 5B** and **Supplementary figure 10B.** Importantly, treatment with iberdomide did not affect IKZF1 expression at the level of transcription (**Figure 5C),** consistent with the known post-transcriptional mechanism by which IMIDs target IKZF1 for CRBN mediated degradation. Conversely, mRNA levels of the endogenous IKZF1 targets, p21 and cMyc (Bjorklund et al., 2020; Fecteau et al., 2014) were significantly affected by the treatment (**Figure 5C)**.

Confounding factors for potential clinical use of candidate LRAs is toxicity and the possibility that they induce unwanted immunomodulatory effects or global immune activation, thus leading to serious adverse effects. Besides, HIV reactivation strategies that aim for a cure will require an intact cytotoxic compartment that promotes viral clearance. Potential toxicity of iberdomide treatment of primary CD4+ T cells was determined by Annexin V staining, followed by flow-cytometry. As shown in **Figure 5D and Supplementary Figure 10C and E**, CD4+T cell viability quantitated as percentage of Annexin V positive cells 24 and 48 hours following iberdomide treatment was not significantly affected.

Given the known immunomodulatory role of Iberdomide and Thalidomide-derived drugs, we investigated the effects of Iberdomide treatment alone or in combination with JQ-1 in T cell

activation, proliferation and functionality. Our results, as shown by extracellular CD69 staining followed by flow-cytometry, demonstrate that Iberdomide treatment alone, or in combination with JQ-1, does not result in global activation of primary CD4+ T cells (**Figure 5E and F**) (**Supplementary Figure 10D and E**).

We then examined proliferation potential upon iberdomide treatment alone or in combination with JQ-1. As previously reported in the literature, and consistent with their anti-lymphoproliferative capacity (Fuchs, 2019), we observed decreased proliferation capacity upon treatment of CD4+ and CD8+ primary T cells (**Figure 5G and H**; **Supplementary Figure 10F and G**). Lastly, we assessed the effect of iberdomide on T cells functionality, by measuring cytokines IFNγ and IL-2 in CD4+ and CD8+ primary T cells. Iberdomide treatment alone or in combination with JQ-1 does not hamper the production of cytokines IFNγ (**Figure 5 I and J; Supplementary Figure 11A and B**) and IL-2 (**Supplementary Figure 11 C-F**) upon stimulation of CD4+ or CD8+ T cells and therefore does not diminish T cell functionality. On the contrary, and consistent with the literature (Gandhi et al., 2014), we observe an increase in the production of IL-2 upon treatment with the LRAs (**Supplementary Figure 11 C-F**).

Targeting IKZF1 by iberdomide treatment reverses latency in cells obtained from cART suppressed HIV-1 infected patients

To validate our findings in a more relevant setting, we tested the efficacy of Iberdomide in CD4+ T cells obtained from virologically suppressed HIV-1 patients. 5 aviremic patients with maintained viremia below 50 copies/mL for at least one year were selected for the study (**Figure 6A**). CD4+ T cells were isolated from PBMCs by negative selection and were left untreated or treated as indicated with Iberdomide, JQ1, or PMA/Ionomycin as a positive control (**Figure 6B**). Treatment with iberdomide alone resulted in a significant increase in cell-associated HIV-1 gag copies in all 5 donors (mean gag copies 42.68) compared to control (mean gag copies 5.262), and was comparable to that observed after treatment of cells with the BET inhibitor JQ1 (mean gag copies 67.72). Due to the low frequency of reservoir cells, we were unable to detect HIV-1 gag copies above the limit of detection for three donors in the untreated control, which we display as 1 copy for representation purposes. PMA/Ionomycin treatment, as expected, resulted in the highest increase in HIV-1 transcription (mean gag copies 758.1). As also observed in primary CD4+ T cells from healthy donors (**Figure 5C**), treatmenent with iberdomide resulted in a modest downregulation of c-Myc, significant when in combination with JQ1, and a significant upregulation of p21 mRNA levels (**Figure 6C**) in cells isolated from aviremic, HIV-1 infected, partecipants.

Importantly, the latency reversal properties of iberdomide treatment in CD4+ T cells obtained from cART suppressed, HIV-1 infected patients, closely match our observations in ex vivo infected primary CD4+ T cells. Remarkably, also in cells derived from HIV-1 infected, aviremic participants we observe a modest but significant increase upon Iberdomide treatment and a robust, synergistic increase in cell-associated HIV-1 RNA upon co-treatment with JQ-1 (mean gag copies 254.72) (**Figure 6B**).

DISCUSSION

Here we identify regulatory protein complexes distinctly bound to the HIV-1 promoter in the latent and active states in vivo. For this we set-up Catchet-MS, a reverse-ChIP strategy which couples CRISPR/dCas9 targeting and purification of a genomic locus to an additional biochemical histone affinity step to enrich for chromatin-associated dCas9 locus-bound complexes. Using targeted dCas9 as bait makes the system versatile and renders Catchet-MS applicable to study other single genomic loci in live cells, without the need to modify the site of interest. Additionally, the introduction of a second, histone based, affinity purification step ensures that Catchet-MS effectively removes unwanted background originating from cellular contaminants and non-localized bait molecules from the mass spectrometry analysis. Catchet-MS can thus be considered a discovery tool to identify proteins distinctively associated with a genomic region of interest in two functionally disparate states, including transcriptional on-off state, as with the latent and active HIV-5'LTR here, states of differentiation, or distinct genomic SNPs.

As Locus-specific proteomics is a developing field that still represents tremendous biochemical challenges, we acknowledge that our strategy also presents limitations common to locus-specific strategies. The high signal to noise ratio together with the low abundance of single-loci, represent a significant hurdle to overcome in locus-specific strategies to identify the complete repertoire of factors bound to a single locus in vivo (Gauchier et al., 2020). Aware of these limitations, we verified the locus specificity of our dCas9 bait targeting by ChIP sequencing, made use of a large input (~3 billion cells) and introduced a two-step (V5/histones) affinity purification, increasing sensitivity and specificity over a single-step affinity purification, and used stringent filtering and exclusive selection of MS hits detected with high confidence. Importantly, we describe the distinct proteome of latent versus activated HIV-1 5'LTR bound factors and not that of the total HIV 5'LTR bound proteome. In our analysis, the chromatin obtained from cells containing an activated HIV-1 5'LTR serves as a control for the chromatin obtained from cells harboring latent HIV-1 5'LTR, and vice versa. Using this approach, the majority of the proteins common to both conditions, including the dCas9 bait, potential non-specific contaminants, but also well-established functionally relevant factors such as EED (Nguyen et al., 2017; Turner et al., 2020), MTA2 (Cismasiu et al., 2008) and UCHL5 (Rathore et al., 2020), are excluded from further analysis, thus confirming the stringency of the approach. A point of concern with the use of dCas9 based bait guided to the HIV-1 promoter was the potential displacement of or interference with binding of putatively important interactors biasing our candidate list. Although we cannot formally exclude this possibility, dCas9 targeting to the 5'LTR did not appear to interfere, functionally, with viral latency or reactivation (Figure 1F and Supplementary Figure 1D), suggesting that any such interference would have been minor. Thus, despite some limitations, we are confident that Catchet-MS has identified bona fide interactors of the HIV-1 5' LTR and thus represents an effective discovery tool. The validity of our approach is justified by the identification of previously identified and characterized interactors serving as positive controls as well as our extensive functional validation of several candidate proteins with a special focus on IKZF1.

Among the list of interactors bound to the HIV-1 5'LTR locus, in addition to the well-established candidates, we find a number of novel interactors of the HIV-1 promoter (5'LTR) in its latent and active states as shown in Figure 3F and Supplementary Table 3. Among the latent 5'LTR associated factors, we find proteins belonging to chromatin remodeling complexes with well-established roles in HIV-1 latency including the Polycomb repressive complex1 (PRC1),

Polycomb repressive complex 2 (PRC2), the SWItch/Sucrose Non-Fermentable complex (SWI/SNF) and the Nucleosome Remodeling Deacetylase (NuRD) complexes (Khan et al., 2018; Ne et al., 2018). Sequence-specific transcription factors are another class of proteins we identify enriched on the latent HIV-1 LTR, with YY1 (Coull et al., 2000) and POU2F1 (Goffin et al., 2005) previously reported to repress HIV-1 transcription. We additionally identify proteins structurally bound to chromatin such as the DNA topoisomerase 2-alpha enzyme (TOP2A) and the Sister chromatid cohesion protein PDS5 homolog A (PDS5A) and B (PDS5B). As expected, the majority of chromatin structural components were enriched in both the activated and repressed states (Supplementary Table 2). Associated with the repressed locus, we also find the facilitates chromatin transcription (FACT) complex component, SUPT16H, previously described to promote viral latency by interfering with Tat-mediated recruitment of P-TEFb (Huang et al., 2015).Dyskerin pseudouridine synthase 1 (DKC1) is another interesting latent LTR-bound candidate, previously described to promote latency by catalyzing the pseudouridylation and stabilization of the 7SK snRNP complex, inhibiting the release of P-TEFb (Zhao et al., 2016). Moreover, given that HIV-1 preferentially integrates in the nuclear periphery and is essentially regulated by proximity to the PML nuclear bodies (Lusic and Giacca, 2015), we re-analysed our unfiltered data for hits belonging to the “nuclear periphery” GO category (GO:0034399). Interestingly, we find that many factors in this category are associated with the HIV-1 5’LTR (Supplementary Figure S4) and that Lamin B1 (LMNB1) is more predomaninantly associated with the latent state consistent with a previous report that the latent provirus associates with the nuclear periphery (Dieudonne et al., 2009). Interestingly in this context, the HIV-1 LTR binding repressor YY-1, has also been reported to interact with and mediate H3K27me3 -dependent formation of lamina-associated domains (LADs), and is more strongly associated with the latent state (Harr et al., 2015). As expected, upon PMA stimulation and HIV-1 transcriptional activation, a much larger number of ribonucleoproteins, splicing factors as well as proteins involved in mRNA processing and nuclear-cytoplasmic transport are identified by Catchet-MS, most likely recruited to the locus following transcriptional activation (Supplementary Table 3).

In search of putative novel molecular targets for HIV-1 latency reversal, we focused on factors identified to be significantly enriched or uniquely bound to the latent and not activated HIV-1 LTR. Using shRNA-mediated depletion, we functionally validated the activity of a number of novel putative candidates in J-Lat 11.1 cells. Of the 10 putative candidates tested, shRNA depletion of 5 factors (HP1BP3, IKZF1, CDC73, DKC1, PNN) resulted in HIV-1 latency reversal as observed by a significant increase in the percentage of GFP positive cells and expression of HIV-1 genes. One of the strongest observed latency reversals upon depletion was due to the sequence-specific transcription factor IKZF1, an attractive candidate for which pharmacological drug targeting is also available (Kronke et al., 2014; Schafer et al., 2018). IKZF1, a critical factor for lymphoid lineage specification in hematopoietic stem cells, suppresses the stem cell and myeloid programs and primes for the expression of lymphoid specific genes (Ng et al., 2009). IKZF1 is also an important regulator of T cell function (Georgopoulos, 2017; Powell et al., 2019). Mechanistically, IKZF1 has a dual role in transcription, acting, depending on the context, both as a repressor and an activator (Geimer Le Lay et al., 2014; Ng et al., 2009; Zhang et al., 2011). In an activating capacity, IKZF1 can restrict the activity of the NuRD complex, promoting chromatin accessibility (Davis, 2011; Marke et al., 2018; Schwickert et al., 2014; Zhang et al., 2011). IKZF1 has also been proposed to act as an adaptor protein for the local recruitment of p-TEFb and the protein phosphatase 1α at

IKZF1 target genes, facilitating transcription elongation (Bottardi et al., 2015). In T cells, however, IKZF1 has been shown to associate with PRC2 and is required for repression of Notch target genes and the hematopoietic stem cell program (Oravecz et al., 2015). IKZF1-mediated gene silencing may also depend on its interaction with the co-repressors CtBP, CtIP and SWI/SNF-related complex and HDAC-containing Sin3 complexes (Marke et al., 2018).

Consistent with its described repressive mechanisms and functions, depletion of IKZF1 led to a strong reactivation of HIV-1 expression in J-Lat 11.1 cells and J-Lat A2. Upon PMA-induced transcription, IKZF1 binding downstream of the HIV-1 promoter is abrogated, indicating that IKZF1 binding is required for transcriptional repression. Remarkably, depletion of IKZF1 in J-Lat 11.1 cells and J-Lat A2, reveals that IKZF1 is required for the establishment of a repressive chromatin environment, characterized by the presence of the repressive chromatin marks H3K9me3 and H3K27me3. Upon IKZF1 depletion the region becomes de-repressed, displaying a drastic loss of enrichment for the H3K27me3 mark which is even more robust than what is observed following PMA induced activation. Evidence points to a model in which, in T-cells, IKZF1 regulates the epigenetic silencing of IKZF1 target genes via recruitment of the PRC2 complex (Oravecz et al., 2015). The role of PRC2 as a regulatory complex, enriched on the latent HIV-1 promoter with a critical role in compaction of chromatin, recruitment of PRC1 and repression of HIV-1 transcription is well established (Bracken et al., 2006; Khan et al., 2018; Marke et al., 2018). Consistent with these reports our Catchet-MS data identified SUZ12, a core subunit of the PRC2 complex, and CBX8, a subunit of the PRC1 complex, to be also associated with the HIV-1 5'LTR in its latent state, pointing to a possible role for these complexes in regulating the chromatin landscape of this region (Figure 3F). We investigated the potential IKZF1-dependent recruitment of the PRC2 subunit SUZ12 to the HIV-1 5'LTR by ChiP-qPCR; depletion of IKZF1 led to reduced SUZ12 enrichment over the HS2 region. Interestingly, we find that upon IKZF1 depletion, enrichment of the PRC1 component CBX8 at these regions was also reduced. Our results are consistent with a model in which IKZF1 directly recruits PRC2 to the HIV-1 5'LTR, as previously described at endogenous target genes in T cells (Oravecz et al., 2015). The PRC2-modified, H3K27 trimethylated region subsequently serves as a docking site for and results in recruitment of PRC1 (Figure 6 D).

Expression of IKZF1 and related family member IKZF3 at the protein level can be controlled by modulating ubiquitination levels. IMIDs such as thalidomide, lenalidomide, pomalidomide and iberdomide promote ubiquitin-dependent proteasomal degradation of IKZF1 and IKZF3 by redirecting the substrate specificity of the CRL4CRBN ubiquitin ligase complex (Kronke et al., 2014; Lu et al., 2014). Among these drugs, iberdomide (CC-220), is a novel compound with the highest reported specificity for CRBN and the lowest IC50, currently in phase 2 Clinical trial for Systhemic lupus erythematosus (SLE) and Multiple Myeloma (MM). Due to its high potency we selected iberdomide for treatment of primary ex vivo HIV-1 infected CD4+ T cells and those obtained from cART suppressed aviremic participants. Iberdomide mediated depletion of IKZF1 was accompanied by a significant reversal of HIV-1 latency, with no significant associated toxicity. A class of compounds previously reported to synergize with the activity of IMIDs in inhibiting the growth of refractory forms of MM are BET inhibitors such as JQ1 (Diaz et al., 2017; Moros et al., 2014). As BET inhibitors are also a well-established class of HIV-1 LRAs, we examined the

latency reversal effect of combination treatment with iberdomide and JQ1. Remarkably, co-treatment of latently infected primary CD4+ T cells results in strong and synergistic induction of HIV-1 transcription in cells obtained from HIV infected cART suppressed aviremic patients, quantitated as increase in cell-associated HIV-1 Gag RNA. Thus, similar to what was observed in the treatment of resistant MM, and of therapeutic interest, combination of these two classes of compounds also leads to synergism in context of HIV-1 latency reversal.

Crucial to the potential applicability of Iberdomide treatment alone or in combination with other LRAs in future clinical studies is the absence of detrimental effects on the overall T cell biology and the preservation of a functional cytotoxic compartment. Interestingly, and consistent with its role as an immunomodulatory treatment for lymphoid and myeloid malignancies (Fuchs, 2019), we observed a reduction of T cell proliferative capacity after treatment with Iberdomide alone or in combination with JQ-1. Importantly, we demonstrate that the use of Iberdomide does not affect T cell functionality, as measured by cytokine expression in CD4+ and CD8+ T cells, and does not cause global immune activation. Together, our results support the potential use of Iberdomide as a novel LRA as it reactivates HIV-1 latency without significant toxicity, global T cell activation or impaired T cell functionality. Still, future studies including Iberdomide or other thalidomide class compounds as LRA alone or in combination with other LRAs should take into consideration the anti-proliferative activity of these compounds and their direct effects on T cell biology.

Our data is consistent with a model in which iberdomide treatment of latent HIV-1 infected cells leads to IKZF1 degradation through CRBN and as consequence inhibits IKZF1 mediated recruitment of the PRC2 and PRC1 complexes downstream of the HIV-1 promoter (Figure 6D). Thus, iberdomide and other FDA/EMA approved IMID family compounds that target IKZF1 for degradation represent attractive candidates for inclusion in proof of concept clinical studies aiming to reduce the latent HIV-1 reservoir. The specificity of the IMID/IKZF1-targeting pharmacological strategy for HIV-1 latency reversal may also be enhanced by combination treatment with other LRA class compounds such as BET inhibitors, which as our data demonstrates leads to synergistic latency reversal.

Figure 1

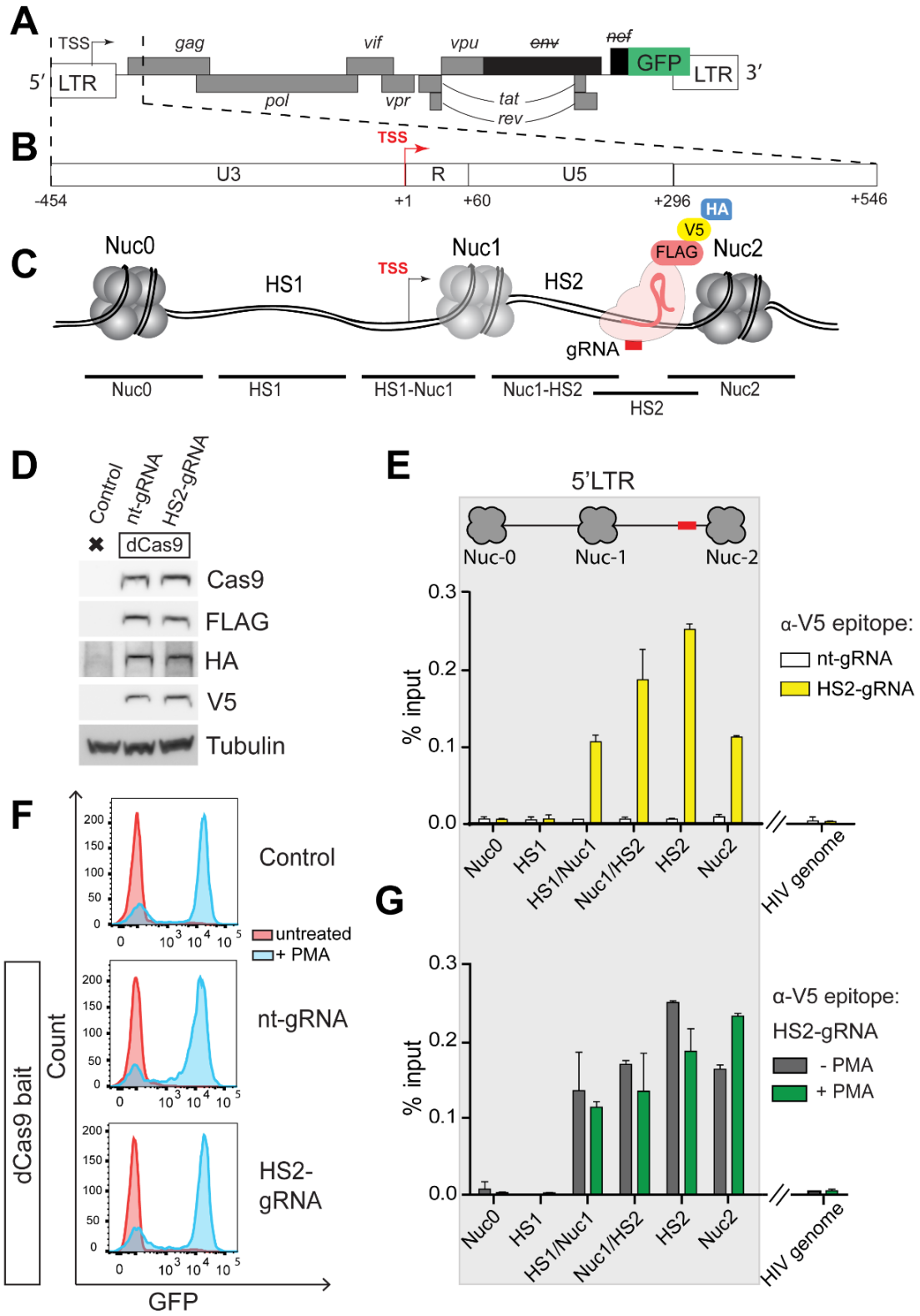

A
TSS
5′ LTR
gag
pol
vif
vpr
vpu
tat
rev
env
nef
GFP
LTR 3′
B
TSS
U3
R
U5
-454
+1
+60
+296
+546
HA
C
Nuc0
HS1
TSS
Nuc1
HS2
V5
FLAG
Nuc2
gRNA
Nuc0
HS1
HS1-Nuc1
Nuc1-HS2
HS2
Nuc2
D
Control
nt-gRNA
HS2-gRNA
dCas9
Cas9
FLAG
HA
V5
Tubulin
E
5'LTR
Nuc-0
Nuc-1
Nuc-2
% input
α-V5 epitope:
nt-gRNA
HS2-gRNA
Nuc0
HS1
HS1/Nuc1
Nuc1/HS2
HS2
Nuc2
HIV genome
F
Control
untreated
+ PMA
nt-gRNA
HS2-gRNA
dCas9 bait
Count
GFP
G
% input
α-V5 epitope:
HS2-gRNA
- PMA
+ PMA
Nuc0
HS1
HS1/Nuc1
Nuc1/HS2
HS2
Nuc2
HIV genome

Figure 1. Generation and characterization of J-Lat 11.1 polyclonal cell line expressing a multiple epitope-tagged HA-V5-FLAG-dCas9 bait, guided to the 5'HIV-1 LTR HS2 region.

(A) Schematic representation of the genomic organization of the integrated HIV-1 provirus in J-Lat 11.1 cells, encoding GFP and containing a frameshift mutation in *env* and a partial deletion of *nef*.
(B) The 5' LTR region is further segmented into the U3, R, and U5 regions.
(C) Three strictly positioned nucleosomes, Nuc-0, Nuc-1 and Nuc-2, delimit the nucleosome-free regions HS1 and HS2, hypersensitive to nuclease digestion, as indicated. The HS2 region, to which the multiple epitope-tagged HA-V5-FLAG-dCas9 bait is guided, is indicated. The amplicons used to scan the chromatin region in ChIP-qPCR are shown.
(D) Western blot analysis indicates expression of multiple epitope-tagged HA-V5-FLAG-dCas9 bait in modified J-Lat 11.1 cells using antibodies specific for Cas9, V5, FLAG and HA as indicated. Parental J-Lat 11.1 cell lysate is used as a negative control and α-Tubulin is used as a loading control.
(E) Flow cytometry histograms show the distribution of GFP positive cells in unstimulated and PMA stimulated control J-Lat 11.1 cells, cells expressing the dCas9 bait and a non-targeting gRNA (nt-gRNA) and cells expressing the bait and the HS2 targeting gRNA (HS2-gRNA).
(F) ChIP-qPCR analysis with anti V5 epitope affinity beads indicate specific enrichment of the HA-V5-FLAG-dCas9 bait over the guided HIV-1 LTR region. White bars represent data generated in cells expressing the dCas9 bait together with a non-targeting gRNA (nt-gRNA), yellow bars represent data generated in cells expressing the bait and the single guide RNA targeting the HS2 region of the HIV-1 5'LTR (HS2-gRNA). Data show a representative experiment, error bars represent the standard deviation (±SD) of two separate real-time PCR measurements. HIV-1 5'LTR sequences recovery is calculated as a percentage of the input.
(G) ChIP-qPCR analysis with anti V5 epitope affinity beads in latent (-PMA; grey bars) and PMA treated cells (+PMA; green bars) expressing the dCas9 bait and the HS2-gRNA. Data are the mean of 2 independent experiments (±SD). HIV-1 5'LTR sequences recovery is calculated as a percentage of the input.

Figure 2

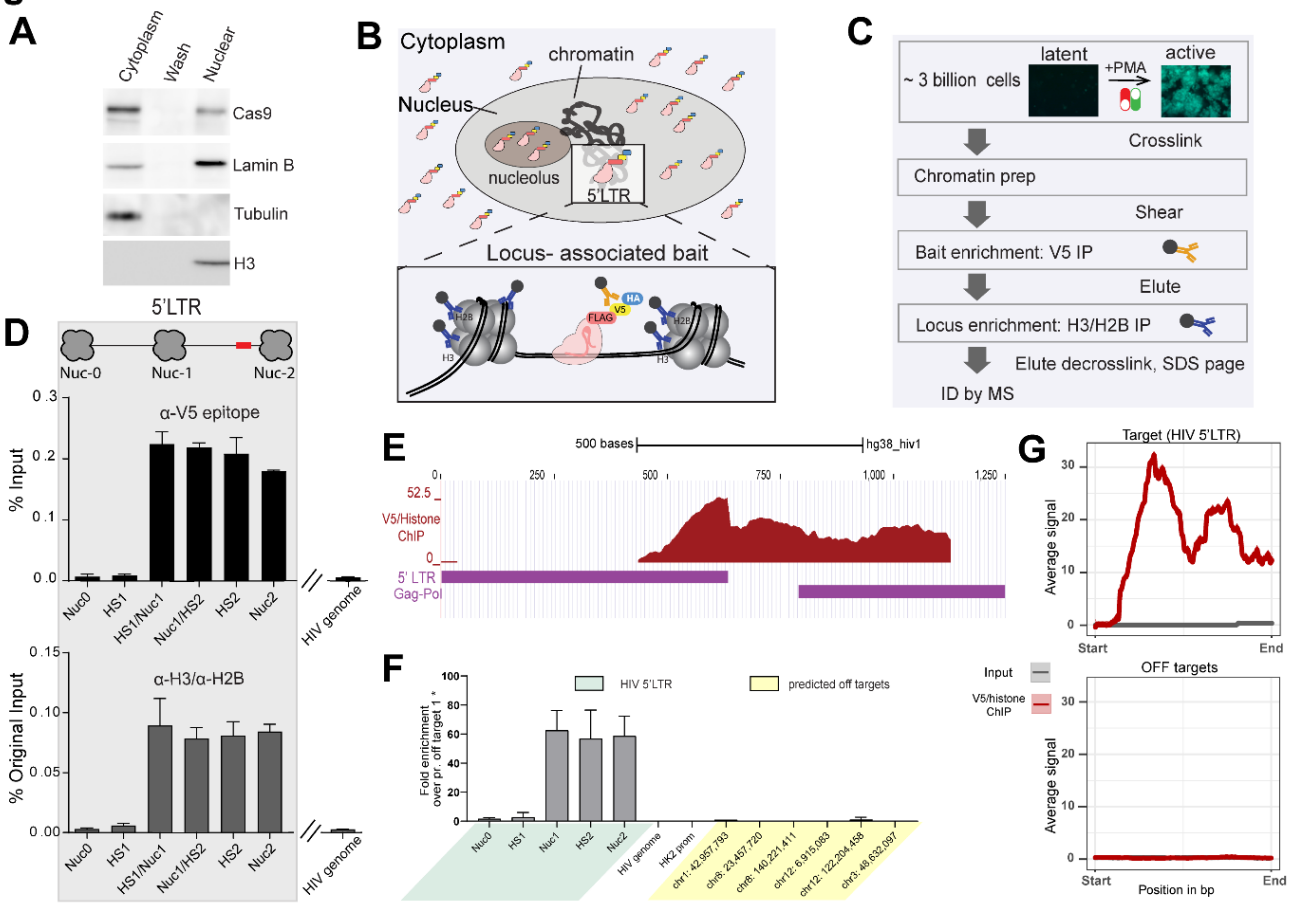


Figure 2. dCas9 targeted chromatin and histone enrichment for mass spectrometry (Catchet-MS), a method to isolate and identify locus-bound protein complexes in vivo.

A) Western blot analysis using antibody specific for dCas9 indicates localization of HA-V5-FLAG-dCas9 bait following a nuclear and nucleolar fractionation protocol. α-Tubulin is used as a cytoplasmic marker, Histone H3 is used as a chromatin marker while Lamin Bis used as a nuclear marker.

(B) Schematic representation of the dCas9 bait cellular localization.

(C) Schematic representation of the Catchet-MS workflow. Approximately 3 billion cells per condition are cross-linked with formaldehyde to stabilize the protein-protein and protein DNA interaction. Following a stringent chromatin enrichment protocol, the cross-linked chromatin is isolated and fragmented by ultrasound sonication. The dCas9 containing complexes are immunoprecipitated using anti V5 antibody conjugated affinity beads, eluted from the beads, and used as input material for a second round of purification with anti-histones (H3, H2B) antibody-conjugated beads, in order to remove the non-chromatin bound fraction of the HA-V5-FLAG-dCas9 bait complexes and to enrich for the locus associated bait complexes. Immunoprecipitated material is finally decrosslinked, resolved on an SDS-page and prepared for mass spectrometry analysis.

(D) (Top-panel) ChIP-qPCR analysis with anti V5 epitope affinity beads in chromatin fraction prepared from unstimulated cells indicates specific enrichment of HA-V5-FLAG-dCas9 over the HIV-1 5'LTR. Data show a representative experiment, error bars represent the standard deviation (SD) of two separate real-time PCR measurements, HIV-1 5'LTR levels are calculated as percentages of the input. The V5 affinity-purified chromatin (represented in top panel) was eluted and used as input for a sequential immunoprecipitation

with a mix of histone H2B and H3 conjugated affinity beads and the isolated material was analyzed by qPCR (Bottom panel). ChIP-qPCR with anti H3/H2B conjugated affinity beads indicates 5'LTR enrichment of HA-V5-FLAG-dCas9 bait after sequential V5/histone affinity purification.Data show a representative experiment, error bars represent the standard deviation (SD) of two separate real-time PCR measurements, HIV-1 5'LTR levels in both the V5 and histone affinity purification are calculated as percentages of the same original input.

(E) ChIP-sequencing tracks of the V5/histone (H3/H2B) sequentially purified chromatin over the HIV-1 5' LTR.

(F) Average coverage profiles using ChIP sequencing reads mapped 500bp upstream and downstream of the peak center at the 5'LTR region of the HIV genome ('Targets') to the respective coverage around the predicted off targets ('OffTargets'). 'Start' denotes the starting base pair of the aforementioned 1kb region around the peak centers and 'End' the ending base pair respectively.

(G) ChIP-qPCR confirmation analysis of the V5/histone (H3/H2B) sequentially purified chromatin at the HIV-1 5'LTR and at dCas9 bait predicted off target regions. Predicted off-target regions were calculated allowing 2 to 3 mismatches in the gRNA sequence. Data are presented as fold enrichment over the predicted off-target region on genomic location chr1: 42,957,793 and indicated with *.

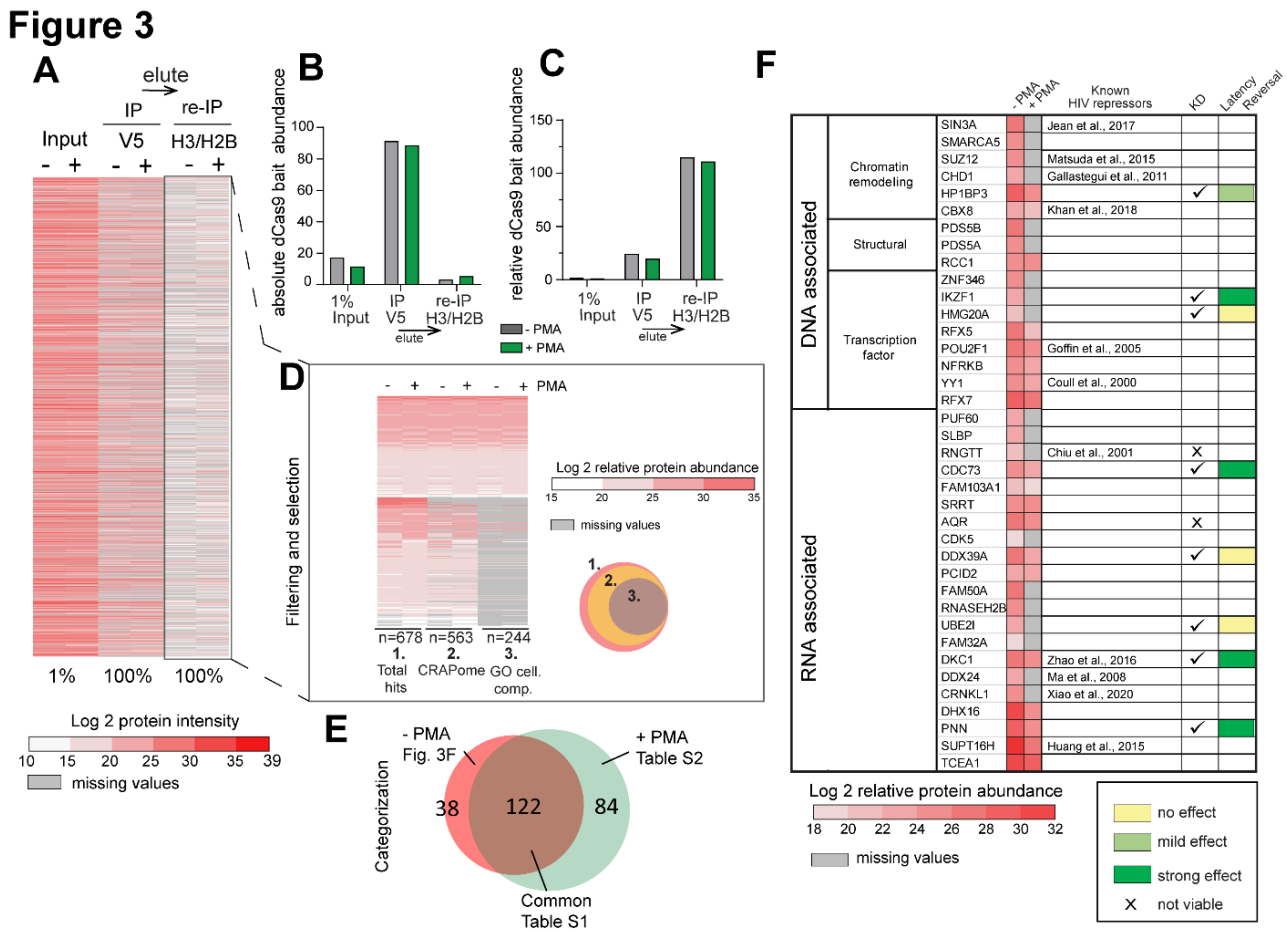


Figure 3. Cachet-MS identifies known and novel host factors associated with the latent and active HIV-1 promoter

(A) Heatmap displays the protein content at each step of the Catchet-MS purification pipeline. The colors represents the Log2 transformation of the protein intensities scores. Values corresponding to the V5 based immunopurification are adjusted to 100% as only a fraction (1:40) of the material that went into the second, histone based (H2B/H3) immunopurification, was analyzed by mass spectrometry. For the V5/histone experiment, 100% of the material was subjected to mass spectrometry. Missing values are represented by grey lines. The heatmap represents data from one Catchet-MS experiment, input material for the experiment corresponds to chromatin generated from a starting material of 3 billion cells per condition.

(B) Bar plots showing the absolute HA-V5-FLAG-dCas9 bait abundance at each step of the purification pipeline. The absolute abundance was calculated based on protein/peptide spectral intensity values and adjusted to the fraction of material analyzed by mass spectrometry. The values are displayed in arbitrary units.

(C) Bar plots showing the relative HA-V5-FLAG-dCas9 bait abundance at each step of the purification pipeline. The relative abundance was calculated based on protein/peptide spectral intensity values and normalized to the total protein content. The values are displayed in arbitrary units.

(D) Filtering criteria applied to Catchet-MS data. Hits are filtered by comparison with the Crapome repository for contaminants of affinity purification experiments . as well as through GO categorization to select hits present in the cell nucleus and classified to be RNA bound, DNA bound or alternatively bound to histones, transcription factors or associated to chromatin. Colors represent the Log2 transformation of the protein

relative abundance. The protein relative abundance was calculated based on protein/peptide spectral intensity values and normalized to the total protein content. Missing values are represented by grey lines.
(E) Venn diagram graphically summarizing the overlap between the hits identified to be exclusively, or more abundantly, associated with the unstimulated state (-PMA) and the hits associated with the activated state (+PMA).
(F) Selection and functional classification of hits associated with the unstimulated state and considered for a potential role in the maintenance of HIV-1 latency. The first columns of the table summarizes in a heatmap the Log2 trasfomation of the protein relative abundance values in the unstimulated (–PMA) sample and the activated (+PMA) sample. Missing values are represented by grey lines. Candidates previously reported to restrict HIV-1 expression are referenced in the third column. Checkmarks indicate the hits functionally validated by shRNA mediated depletion experiments (Supplementary Figure 3C). Colored boxes (yellow to dark green) summarize the effect of the target depletion on HIV- 1 expression and refers to the effects summarized in S5A.

Figure 4

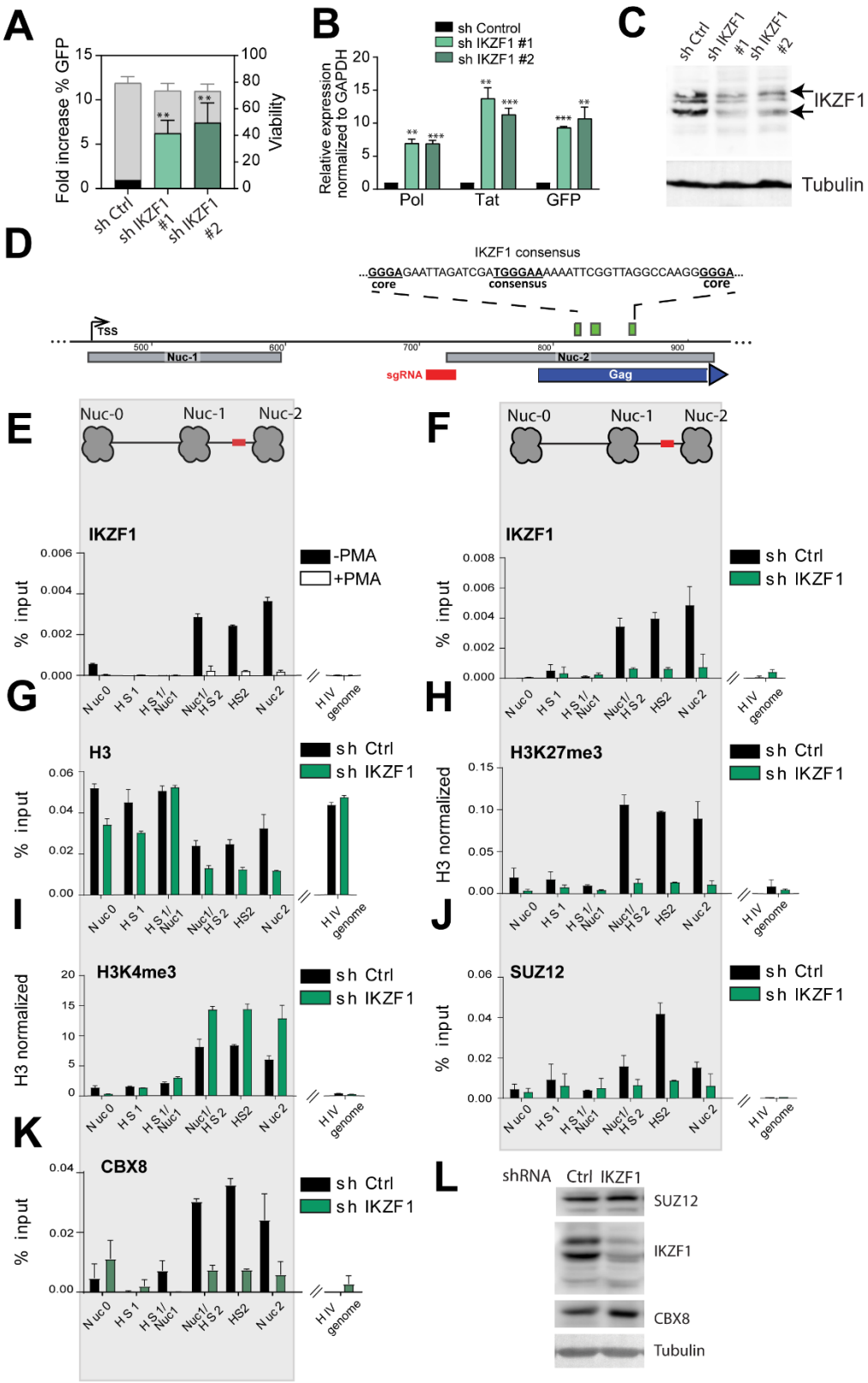

A
Fold increase % GFP
Viability
sh Ctrl
sh IKZF1 #1
sh IKZF1 #2
B
sh Control
sh IKZF1 #1
sh IKZF1 #2
Relative expression normalized to GAPDH
Pol
Tat
GFP
C
sh Ctrl
sh IKZF1 #1
sh IKZF1 #2
IKZF1
Tubulin
D
IKZF1 consensus
...GGGAGAATTAGATCGATGGGAAAAAATTCGGTTAGGCCAAGGGGGGA...
core
consensus
core
TSS
Nuc-1
Nuc-2
sgRNA
Gag
E
Nuc-0
Nuc-1
Nuc-2
IKZF1
% input
-PMA
+PMA
F
IKZF1
sh Ctrl
sh IKZF1
G
H3
H
H3K27me3
H3 normalized
I
H3K4me3
J
SUZ12
K
CBX8
L
shRNA
Ctrl
IKZF1
SUZ12
IKZF1
CBX8
Tubulin

Figure 4. IKZF1, required for maintenance of HIV-1 latency, binds downstream of the latent HIV-1 5'LTR to promote PRC2/PRC1 recruitment and establish a repressive chromatin environment.

(A) Bar plot showing the fold increase in % GFP positive cells (left y axes) measured by FACS analysis, following IKZF1 depletion in J-Lat 11.1 with two different shRNA constructs (#1 and #2). Data are the mean of three independent experiments (±SD). The right y axis represents the percentage of live cells.
(B) qRT-PCR analysis measuring expression of HIV genes (pol, GFP, tat) in J-lat 11 transduced with scramble shRNA (sh Control) and sh IKZF1 #1 and #2. Data, normalized to GAPDH are represented as fold enrichment over sh Control and are the mean of three independent experiments (±SEM). Statistical significance was calculated using a ratio paired t test, * – p<0,05; **– p<0,01– ***p<0,001.
(C) Western blotting for IKZF1 in J-lat 11 transduced with scramble shRNA (sh Control) and shIKZF1 #1 and #2 iondicates depletion of IKZF1. α-Tubulin is used as a loading control.
(D) Putative IKZF1 binding site (835-840) downstream of the HIV-1 5'-LTR in the proximity of the sequence targeted by the gRNA. The IKZF1 binding site is composed of a consensus sequence (TGGGAA/T) and at least one more extra core site (GGGA) in a 40 bp range (Li et al., 2011).
(E) ChIP-qPCR analysis with IKZF1 antibody in latent cells and PMA stimulated J-Lat 11.1 cells as indicated.
(F) ChIP qPCR analysis with IKZF1 antibody in J-Lat 11. 1 cells transduced with scramble shRNA (shControl) and shIKZF1. Data in (E) and (F) are presented as % input, error bars represent the standard deviation (SD) of two separate real-time PCR measurements.
(G, H, I) ChIP-qPCR using antibodies specific for distinct histone marks in J-Lat 11. 1 cells transduced with scramble shRNA (shControl) and shIKZF1. Total histone H3 (G), H3K27me3 (H), H3K4me3 (I). Total histone H3 data (G) are represented as % input mean (±SD), histone marks data (H)(I) are expressed as fold change over H3 signal (±SD).
(J) ChIP-qPCR analysis with SUZ12 in J-Lat 11. 1 cells transduced with scramble shRNA (shControl) and shIKZF1.
(K) ChIP-qPCR analysis with CBX8 in J-Lat 11. 1 cells transduced with scramble shRNA (shControl) and shIKZF1. Data in (J) and (K) are presented as % input, error bars represent the standard deviation (SD) of two separate real-time PCR measurements. The ChIP analysis presented (E-K) are a representative experiment, biological replicate experiments are shown in S8.
(L) Western blotting for SUZ12, IKZF1, CBX in J-lat 11 transduced with scramble shRNA (sh Control) and shIKZF1. α-Tubulin is used as a loading control.

Figure 5

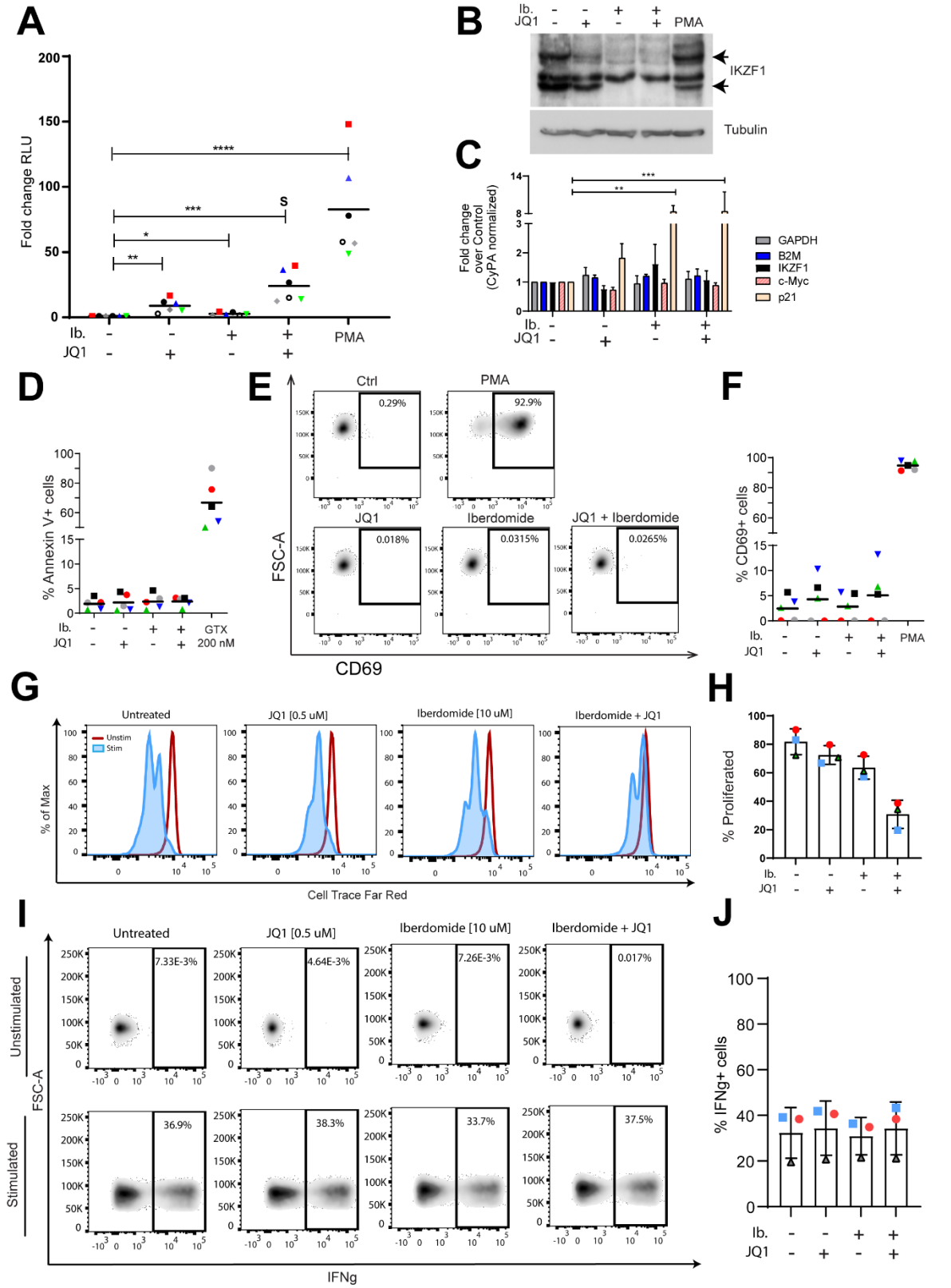

A
Fold change RLU
Ib.
JQ1
PMA
B
IKZF1
Tubulin
C
Fold change over Control (CyPA normalized)
GAPDH
B2M
IKZF1
c-Myc
p21
D
% Annexin V+ cells
GTX 200 nM
E
Ctrl
PMA
JQ1
Iberdomide
JQ1 + Iberdomide
0.29%
92.9%
0.018%
0.0315%
0.0265%
FSC-A
CD69
F
% CD69+ cells
G
Untreated
JQ1 [0.5 uM]
Iberdomide [10 uM]
Iberdomide + JQ1
Unstim
Stim
% of Max
Cell Trace Far Red
H
% Proliferated
I
Unstimulated
Stimulated
7.33E-3%
4.64E-3%
7.26E-3%
0.017%
36.9%
38.3%
33.7%
37.5%
IFNg
J
% IFNg+ cells

Figure 5. Targeting IKZF1 by iberdomide treatment reverses latency in ex vivo infected primary CD4+ T cells with minimal effect on toxicity and effector function

(A) The latency reversal activity of iberdomide alone (10uM) and in combination with JQ1 (500nM) was tested, in primary human CD4+ T cells ex vivo infected to harbour latent HIV-1. The dot plot in panel A shows the fold increase in luciferase activity after treatment as indicated. Each dot represents a single measurement while the black horizontal lines represent the average fold increase for each treatment in the pool of donors. Experiments were performed in duplicate using cells obtained from 6 healthy donors. Statistical significance was determined by repeated measures one-way ANOVA on the log-transformed fold changes followed by Tukey's multiple comparison test * – $p < 0,05$, ** – $p < 0,01$, *** – $p < 0,001$, ****– $p < 0,0001$.

(B) Western blotting shows protein levels of IKZF1 after 24 hours following treatment with Iberdomide, JQ1, and their combination in primary CD4+ T cells. α-Tubulin is used as a loading control.

(C) qRT-PCR analysis of IKZF1 and IKZF1 targets p21 and c-myc upon treatment with JQ1, iberdomide, and their combination. RT-PCR was performed in primary CD4+ T cells isolated from three healthy donors. Data are represented as fold change (±SEM) over untreated and are normalized to Cyclophilin A (CyPA). B2M and GAPDH are used as housekeeping genes. Statistical significance was calculated using ratio-paired t-test * – $p < 0,05$, ** – $p < 0,01$, *** – $p < 0,001$,

(D) Iberdomide treatment is not cytotoxic in primary CD4+ T cells. Percentage of cells expressing the Annexin V marker of apoptosis in primary CD4+ T cells treated with iberdomide, JQ1 and the combination of both compounds for 48 hours. Treatment with a toxic concentration of Gliotoxin (GTX) 200nM was used as a positive control. Experiments were performed in uninfected cells obtained from five healthy donors, represented by the dots.

(E) Iberdomide treatment does not activate CD4+ T cells. Representative flow cytometry plot of extracellular CD69 marker staining analysis in primary CD4+ T cells, upon treatment with JQ1, iberdomide, or a combination of both compunds for 48 hours. CD69 expression was assessed by extracellular staining and analyzed by flow cytometry

(F) Percentage of cells expressing the CD69 marker of cell activation in primary CD4+ T cells from 5 healthy donors as described in E. Treatment with PMA was used as a positive control.

(G) Representantive histogram of proliferative capacity of unstimulated or aCD3/CD28 stimulated CD4+ T cells in the presence or absence of LRAs. Cells were stained with a proliferation dye and analyzed 72 hours later by flow cytometry. Dividing cells show decreased intensity of prolifertion dye as it becomes diluted upon cell division.

(H) Percentage of proliferated CD4+ T cells from 3 healthy donors as described in G.

(I) Representative flow cytometry plots of IFNg production analysis in unstimulated and stimulated primary CD8+ T cells after treatment with LRAs. Cells were treated as indicated for 18 hours followed by PMA/Ionomycin stimulation for 7 hours in the presence of a protein transport inhibitor or remained unstimulated. IFN-g production was assessed by intracellular staining and analyzed by flow cytometry. Numbers in the plot show percentage of IL-2 producing cells.

(J) Percentage of INFg producing CD8+ T cells from 3 healthy donors as described in I

Figure 6

A

	DONOR 1	DONOR 2	DONOR 3	DONOR 4	DONOR 5
SEX	Male	Male	Female	Male	Male
AGE	66	67	42	75	47
TIME FROM HIV DIAGNOSIS TO LEUCAPHERESIS	52 months	86 months	55 months	108 months	85 months
TIME ON ART AT LEUKAPHERESIS	52 months	86 months	55 months	81 months	72 months
TIME WITH CONTINOUS HIV RNA <50 COPIES/ML	48 months	12 months	52 months	75 months	64 months

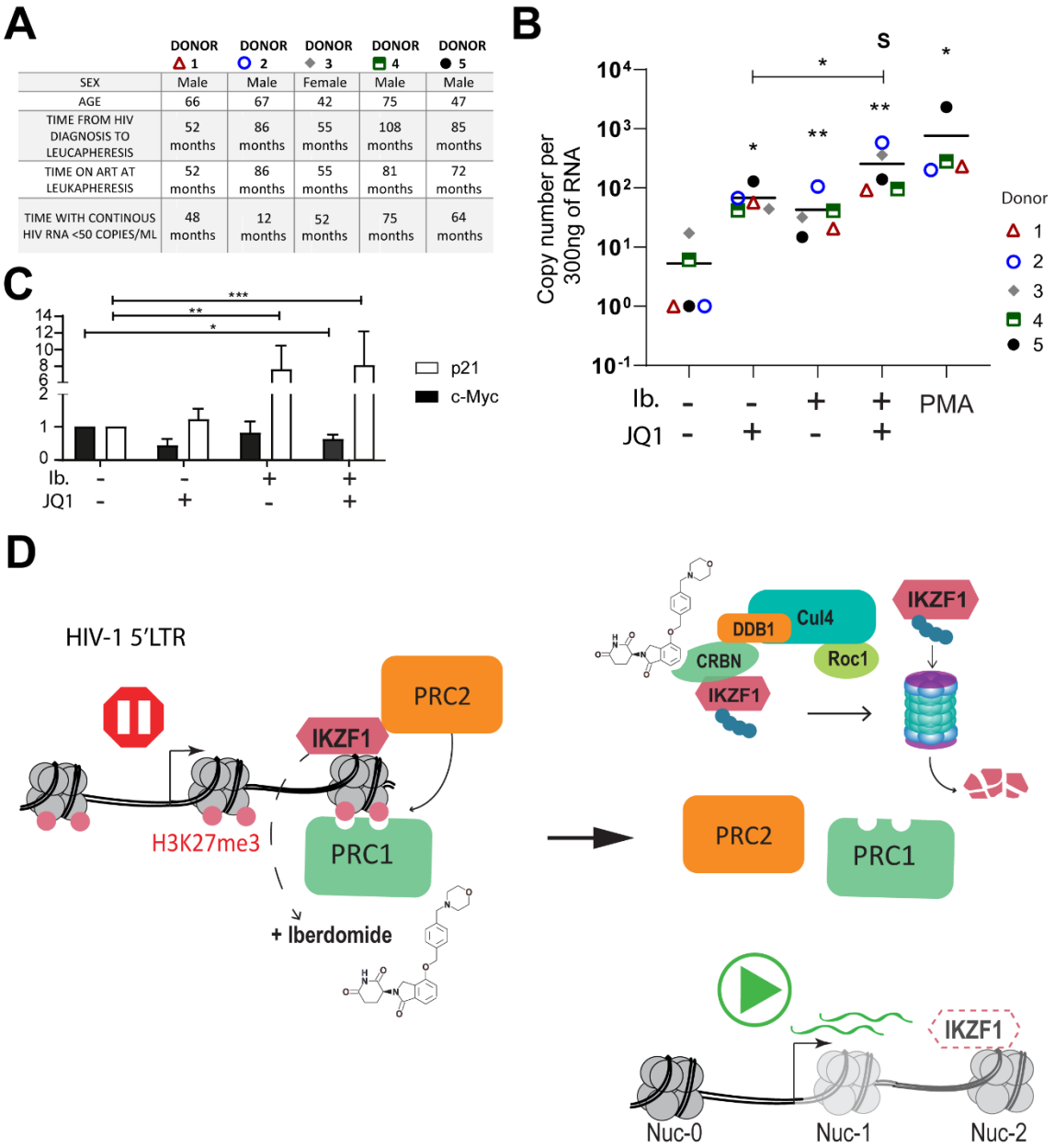


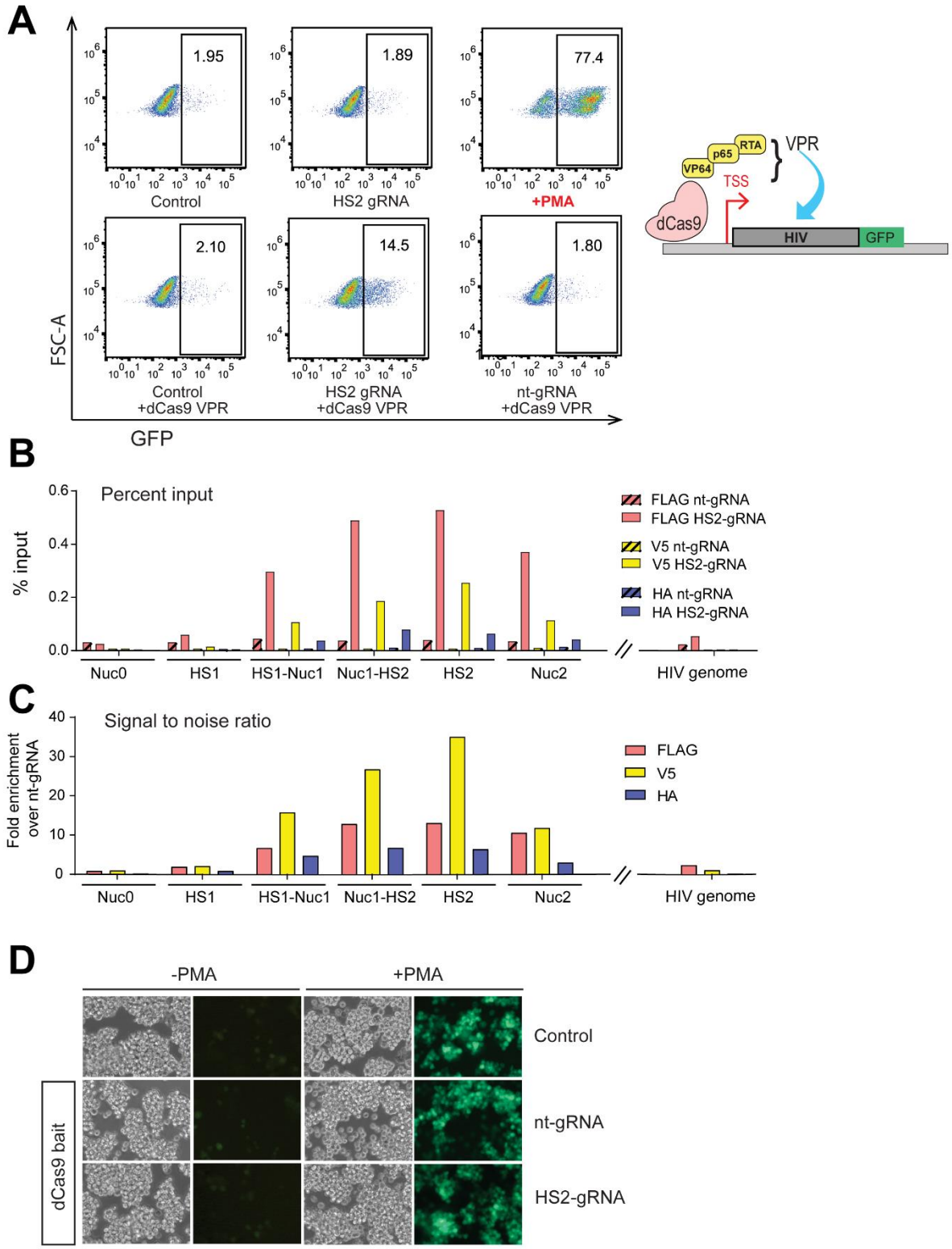
Figure 6 Targeting IKZF1 by iberdomide treatment reverses latency in primary CD4+ T cells obtained from cART suppressed HIV-1 infected patients.

(A) Clinical Information Table corresponding to the HIV-1 infected study participants.
(B) Iberdomide treatment causes significant increase in cell-associated HIV-1 RNA levels and synergistically enhances JQ1-mediated HIV-1 latency reversal in CD4+ T cells obtained from c-ART suppressed HIV-1 infected patient volunteers. Graph panel showing the average levels of cell-associated HIV-1 RNA in CD4+ T cells isolated from five infected, aviremic participants. Statistical significance was calculated using unpaired Mann-Whitney test, * – $p < 0,05$, ** – $p < 0,01$.

(C) qRT-PCR analysis of IKZF1 targets p21 and c-myc upon treatment with JQ1, iberdomide, or both compounds. RT-PCR was performed in primary CD4+ T cells isolated from 5 aviremic HIV-1 infected study partecipants. Data are represented as fold change mean (±SD) over untreated and are normalized with Cyclophilin A (CyPA). Statistical significance was calculated using ratio-paired t test * – $p < 0,05$, ** – $p < 0,01$, *** – $p < 0,001$.

(D) Proposed model for iberdomide mediated latency reversal. Iberdomide binds to CRBN, a subunit of the CRL^{CRBN} E3 ubiquitin ligase complex that acts as a substrate adaptor. Iberdomide biding induces recruitment of IKZF1 and its ubiquitination by the ligase. In latently infected cells, ubiquitination of IKZF1 and its degradation by the proteasome results in a decrease in IKZF1 levels, impaired LTR recruitment of PRC2 and PRC1, leading to chromatin derepression and latency reversal.

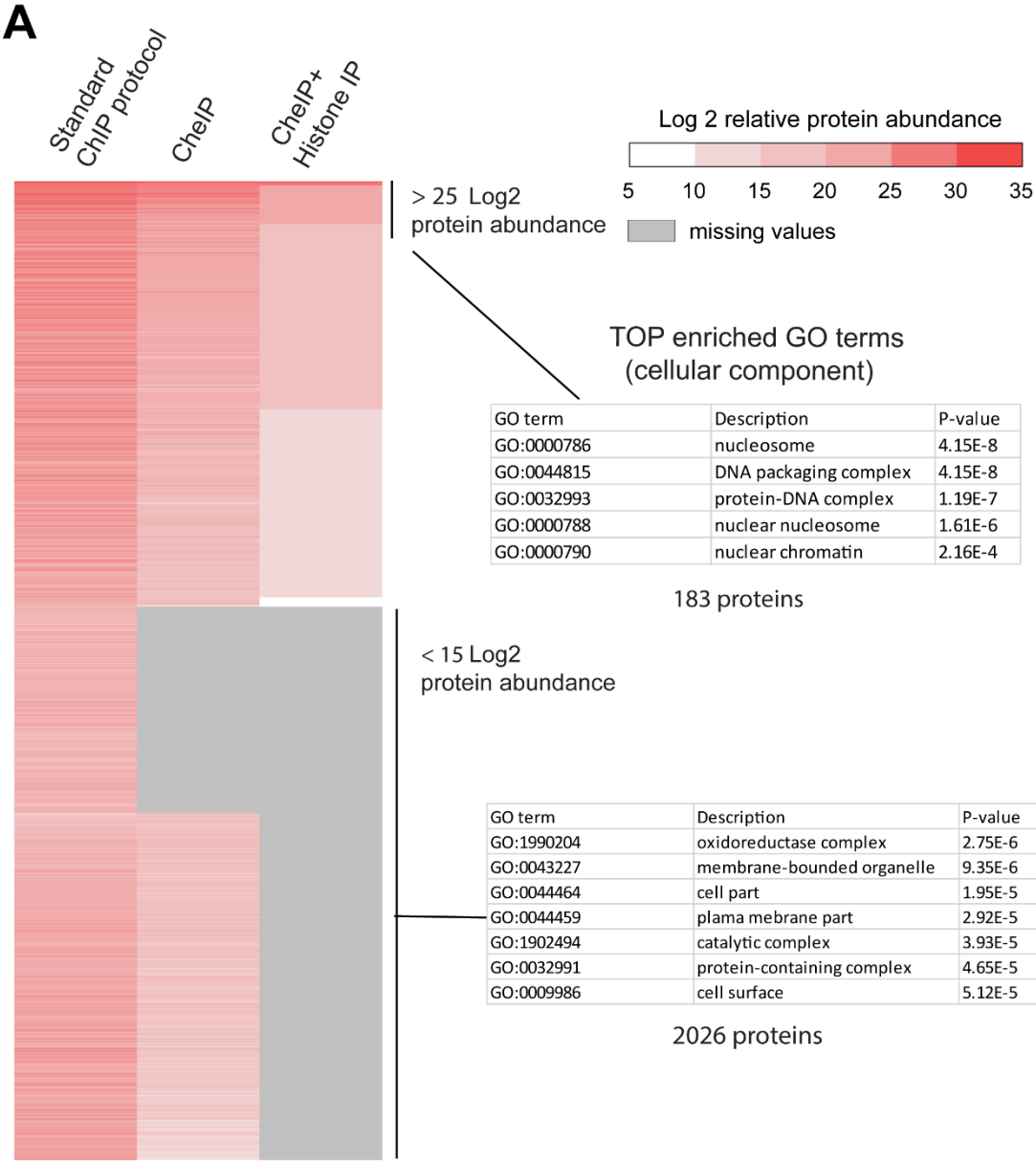
Figure S1


A
1.95
1.89
77.4
Control
HS2 gRNA
+PMA
2.10
14.5
1.80
Control +dCas9 VPR
HS2 gRNA +dCas9 VPR
nt-gRNA +dCas9 VPR
FSC-A
GFP
VP64
p65
RTA
VPR
TSS
dCas9
HIV
GFP
B
Percent input
% input
FLAG nt-gRNA
FLAG HS2-gRNA
V5 nt-gRNA
V5 HS2-gRNA
HA nt-gRNA
HA HS2-gRNA
Nuc0
HS1
HS1-Nuc1
Nuc1-HS2
HS2
Nuc2
HIV genome
C
Signal to noise ratio
Fold enrichment over nt-gRNA
FLAG
V5
HA
Nuc0
HS1
HS1-Nuc1
Nuc1-HS2
HS2
Nuc2
HIV genome
D
-PMA
+PMA
Control
nt-gRNA
HS2-gRNA
dCas9 bait

Supplementary Figure 1. Characterization of the experimental system.
(A) Functional validation of the gRNA designed against the HIV-1 5'LTR HS2 region (HS2-gRNA). The FACSs plots show J-Lat 11.1 cells examined by flow-cytometry at 72 hours after nucleofection with a dCas9 VPR construct to check for HIV-1 LTR dependent GFP expression, measured as % GFP positive cells. (B) ChIP-qPCR experiments performed using different antibodies, conjugated to affinity beads, against the different synthetic tags (FLAG, V5, HA) of the dCas9 construct. Cells expressing the HS2 gRNA and control cells expressing a non-targeting gRNA (nt-gRNA) are compared. HIV-1 5'LTR sequences recovery is calculated as a percentage of the input. (C) From the experiment shown in (B), the signal to noise ratio of the experiment is calculated by dividing the ChIP-qPCR signal obtained in the HS2 gRNA expressing pool with the signal obtained in the nt-gRNA expressing pool. Data are represented as fold enrichment over the non-targeting gRNA signal. (D) Microscopy pictures (bright field and fluorescence) of control J-lat 11.1 cells, cells expressing dCas9 bait and a non-targeting gRNA, cells expressing the bait and a gRNA targeting the HS2 region. Cells have been examined in unstimulated (-PMA) and in the presence of 20nM PMA (+PMA) to assess reactivation capacity.

Figure S2

A



TOP enriched GO terms (cellular component)

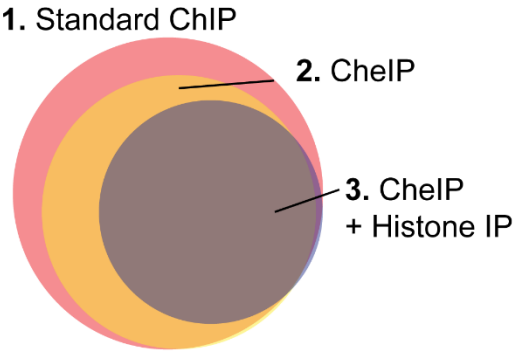
GO term	Description	P-value
GO:0000786	nucleosome	4.15E-8
GO:0044815	DNA packaging complex	4.15E-8
GO:0032993	protein-DNA complex	1.19E-7
GO:0000788	nuclear nucleosome	1.61E-6
GO:0000790	nuclear chromatin	2.16E-4

183 proteins

GO term	Description	P-value
GO:1990204	oxidoreductase complex	2.75E-6
GO:0043227	membrane-bounded organelle	9.35E-6
GO:0044464	cell part	1.95E-5
GO:0044459	plama mebrane part	2.92E-5
GO:1902494	catalytic complex	3.93E-5
GO:0032991	protein-containing complex	4.65E-5
GO:0009986	cell surface	5.12E-5

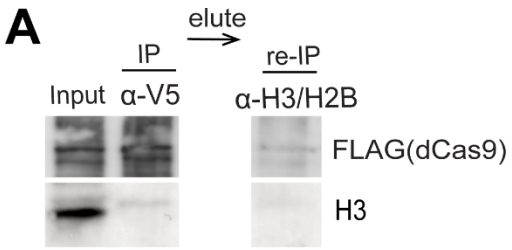
2026 proteins

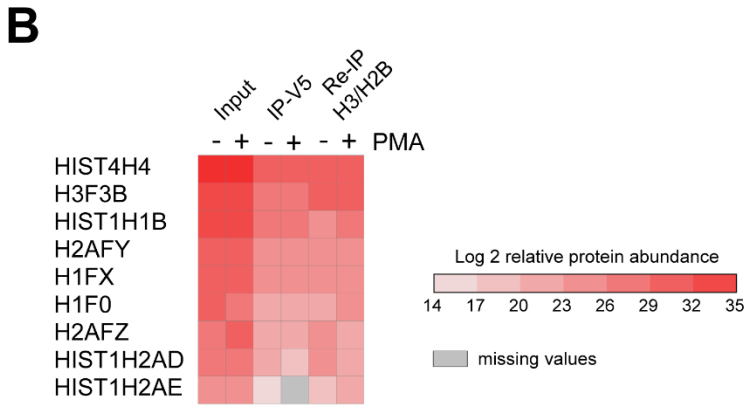
B

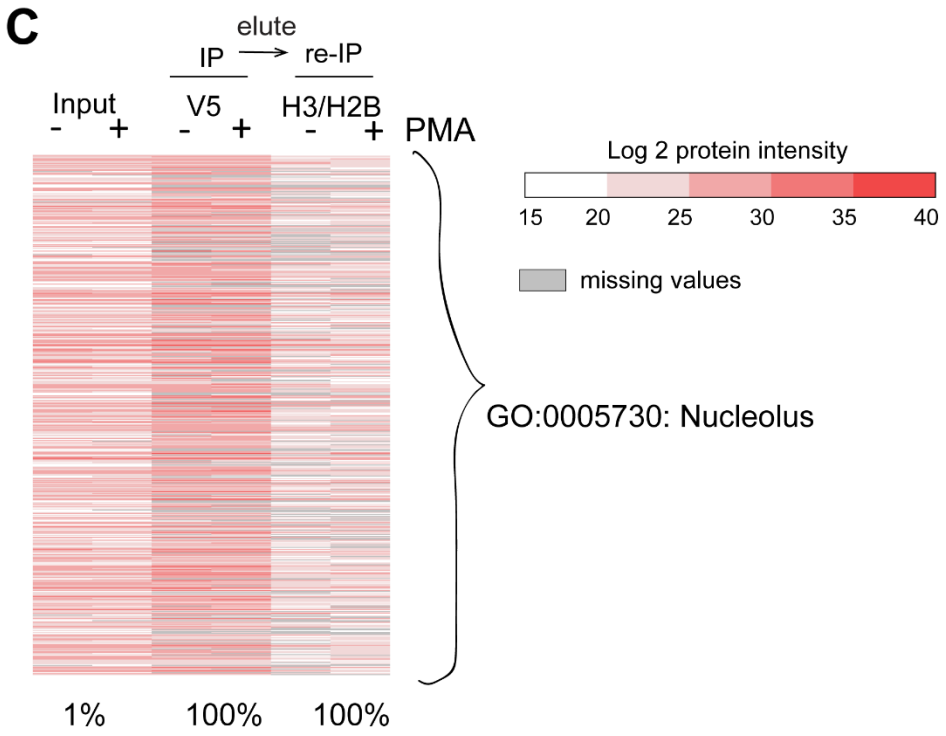


Supplementary Figure 2. Characterization of the chromatin preparation protocol by mass spectrometry. (A) The heatmap shows a comparison between the protein content of a standard ChIP protocol, CheIP, and CheIP followed by a histone enrichment step (CheIP + histone IP) with H2B and H3 conjugated affinity beads. The colors represent the Log2 transformation of the proteins relative abundance. The protein relative abundance was calculated based on protein/peptide spectral intensity values and normalized to the total protein content (B) Venn diagram showing the proportion of the number of hits identified in the different protocols

Figure S3


A
elute
IP
re-IP
Input
α-V5
α-H3/H2B
FLAG(dCas9)
H3

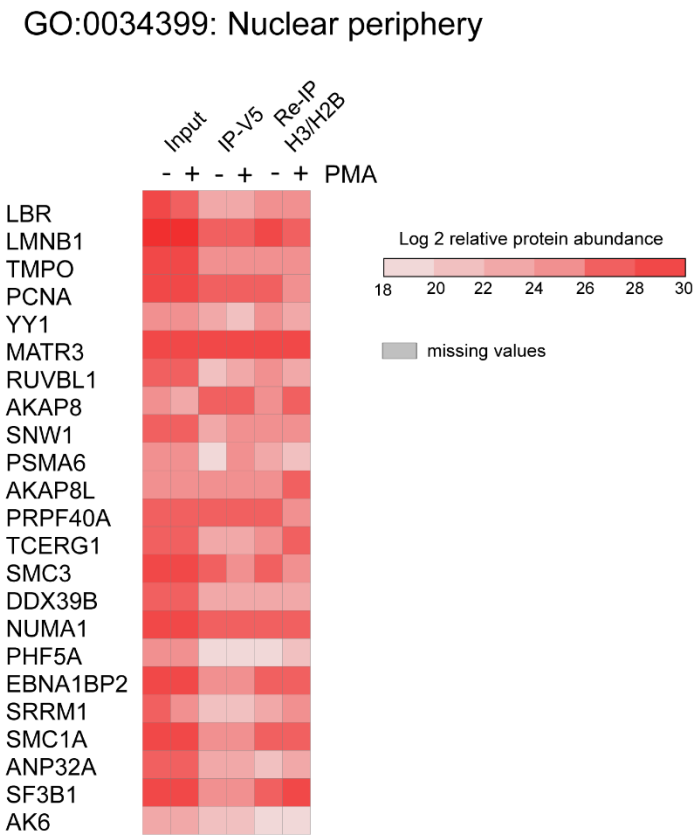

B
Input
IP-V5
Re-IP H3/H2B
- + - + - + PMA
HIST4H4
H3F3B
HIST1H1B
H2AFY
H1FX
H1F0
H2AFZ
HIST1H2AD
HIST1H2AE
Log 2 relative protein abundance
14 17 20 23 26 29 32 35
missing values


C
elute
IP
re-IP
Input
V5
H3/H2B
- + - + - + PMA
Log 2 protein intensity
15 20 25 30 35 40
missing values
GO:0005730: Nucleolus
1%
100%
100%

Supplementary Figure 3. Depletion of non localized bait and nucleolar proteins

(A) Western blotting with anti Cas9 and anti-V5 antibody indicates relative presence of HA-V5-FLAG-dCas9 bait in the fractions used in the sequential ChIP experiments in Figure 2D. (B) Detection of histones. The heatmap summarizes the relative abundance of histone proteins, detected by Catchet-MS. The colors represents the Log2 transformation of the proteins relative abundance. The protein relative abundance was calculated based on protein/peptide spectral intensity values and normalized to the total protein content. Missing values are represented by grey lines. (C) Heatmap displaying the content of nucleolar proteins (GO cellular compartment category GO:005730; nucleolus) at each step of the Catchet-MS purification pipeline used for isolation of the HIV-1 5'LTR. The colors range represents the represents the Log2 transformation of the proteins intensities scores. Values corresponding to the V5 based immunopurification are adjusted to the fraction of material analyzed by mass spectrometry, corresponding to 1:40 of the material used for the second, histone based (H2B/H3) immunopurification. Missing values are represented by grey lines.

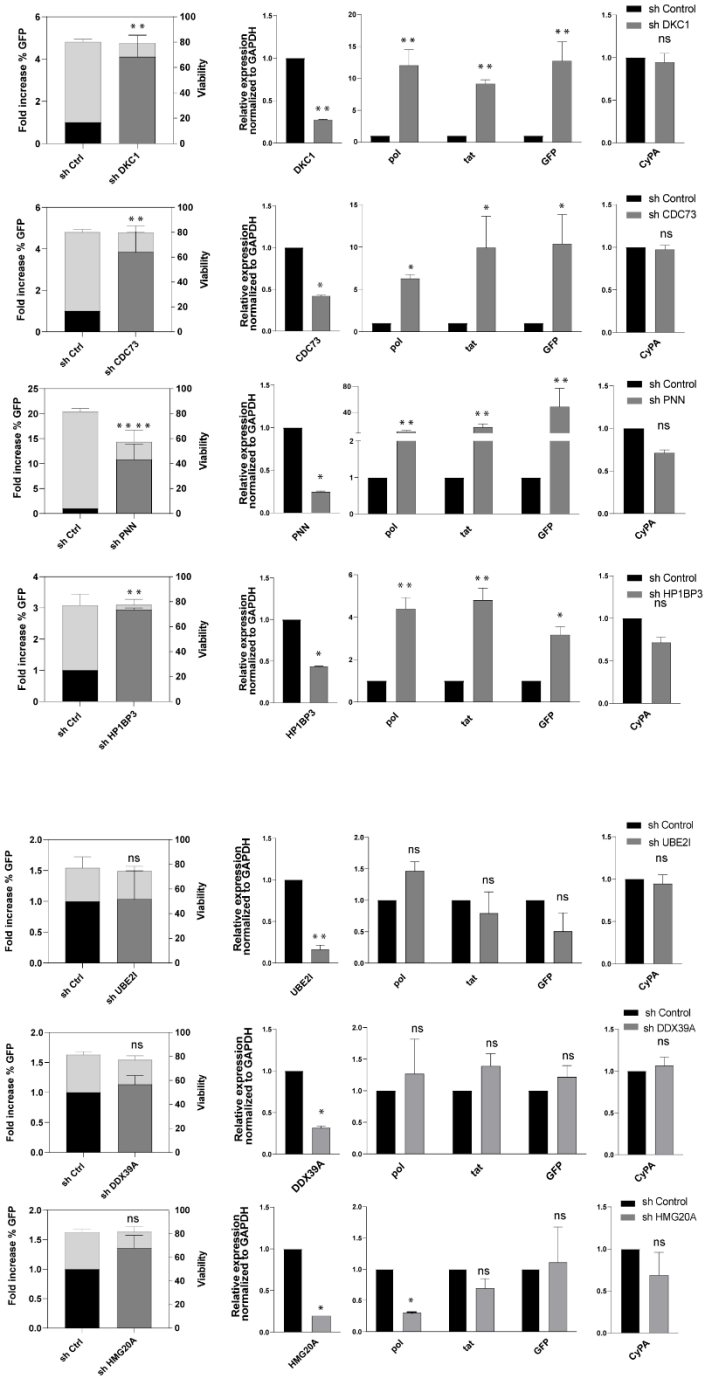
Figure S4



Supplementary Figure 4. Identification of nuclear body and nuclear periphery associated proteins with the latent HIV-1 LTR.

The heatmap summarizes the relative abundance of proteins associated with the nuclear periphery (GO:0034399), detected by Catchet-MS. The colors represents the Log2 transformation of the proteins relative abundance. The protein relative abundance was calculated based on protein/peptide spectral intensity values and normalized to the total protein content. Missing values are represented by grey lines.

Figure S5


Fold increase % GFP
Viability
sh Ctrl
sh DKC1
sh CDC73
sh PNN
sh HP1BP3
sh UBE2I
sh DDX39A
sh HMG20A
Relative expression normalized to GAPDH
DKC1
CDC73
PNN
HP1BP3
UBE2I
DDX39A
HMG20A
pol
tat
GFP
CyPA
sh Control
ns

Supplementary Figure 5. Functional validation of a selection of proteins bound downstream of the latent HIV-1 promoter.

Functional validation of the hits associated with the repressed HIV-1 LTR. shRNA mediated depletion followed by Flow cytometry and RT-PCR. Statistical significance was calculated using ratio-paired t-test and multiple comparison t-test on Log2 transforment fold changes * – $p < 0,05$, ** – $p < 0,01$, *** – $p < 0,001$, ****– $p < 0,0001$.

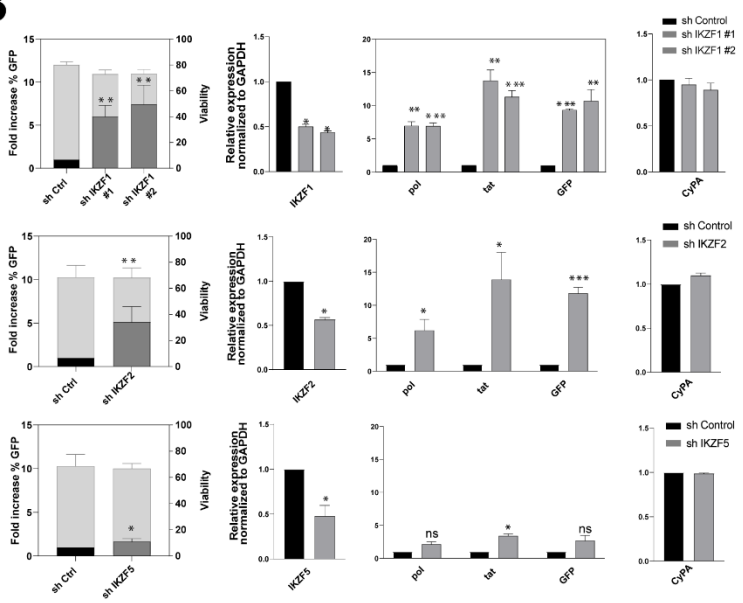
Figure S6

A

RNA seq, Jurkat cells

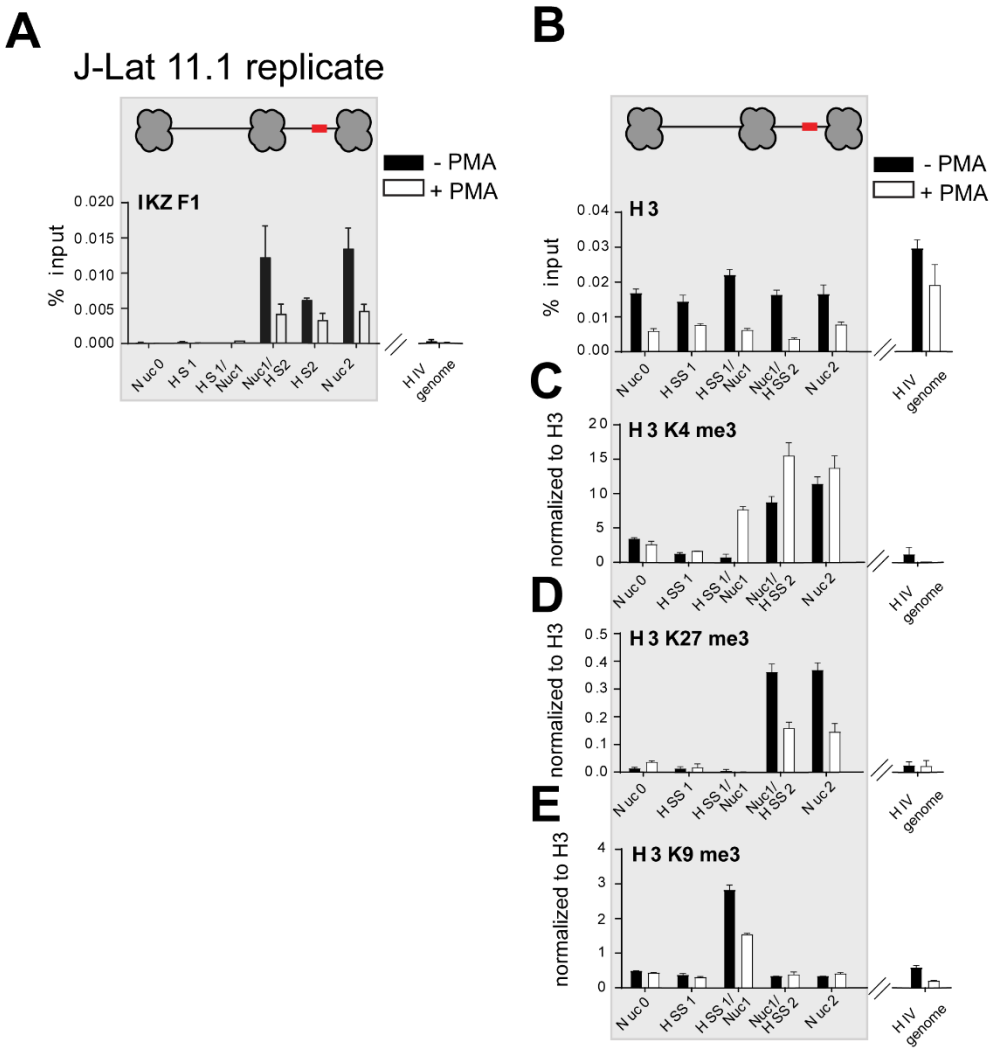
Target	Total reads
IKZF1	5202,5
IKZF2	3303
IKZF3	49
IKZF4	1
IKZF5	701,5

B



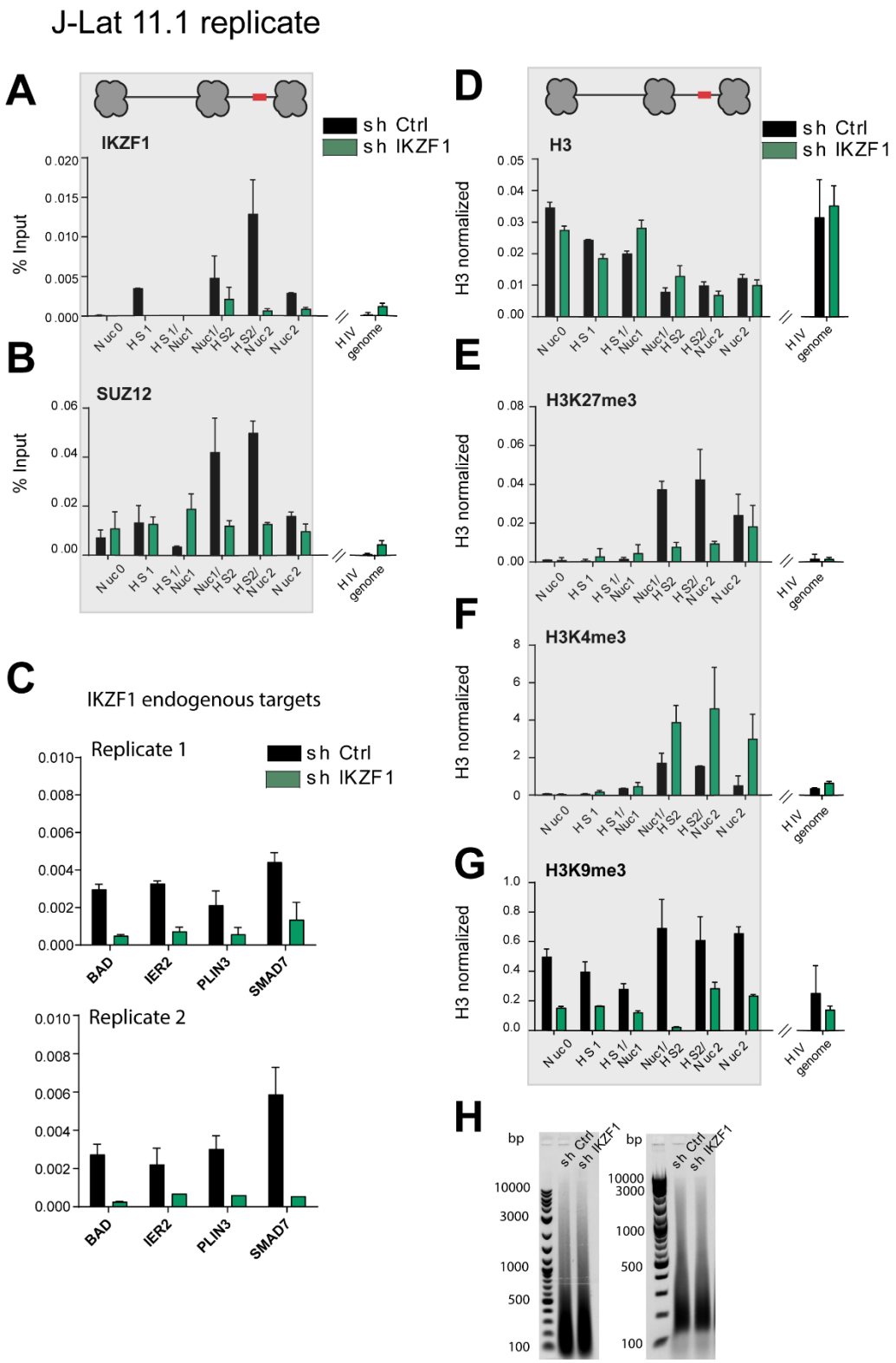
Supplementary Figure 6. IKZF family members expression and role in HIV-1 latency in J-Lat 11.1 cells. (A) Expression of IKZF family members in Jurkat cells. The panel shows the number of RNA seq reads in Jurkat cells. The data are published and available in Palstra et al., Science advances, 2018).
(B) IKZF2 and IKZF5, the two prominent Jurkat cells expressed IKZF members were depleted from J-Lat 11.1 cells following shRNA-mediated transduction and GFP expression was examined by Flow cytometry and RT-PCR. Statistical significance was calculated using ratio-paired t-test and multiple comparison t-test * – p < 0,05, ** – p < 0,01 , *** – p < 0,001, ****– p < 0,0001 .

Figure S7



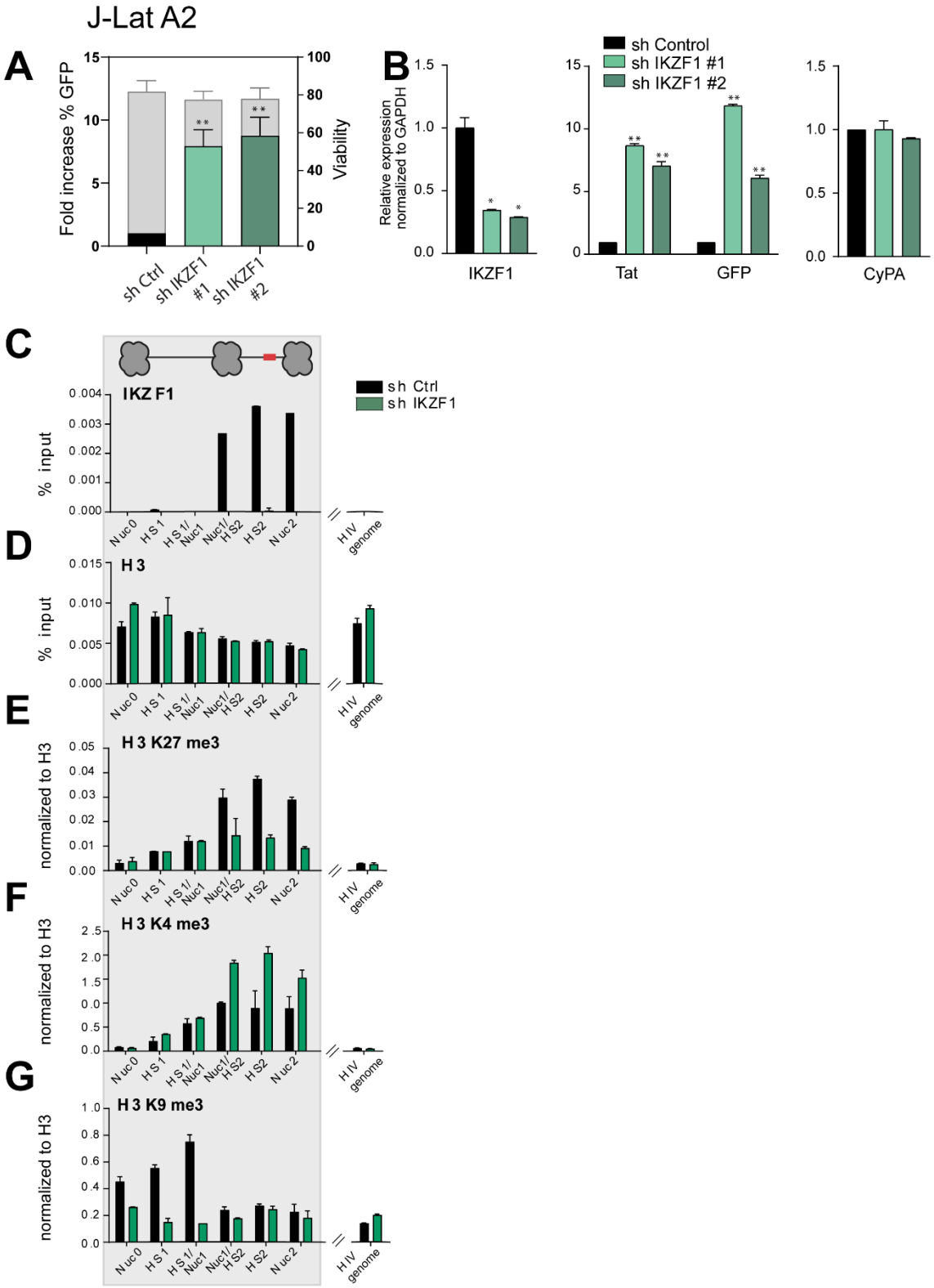
Supplementary Figure 7. Carachterization of IKZF1 binding and chromatin state at the 5'LTR following treatment with PMA in J-Lat 11.1 cells. (A) Replicate ChIP-qPCR analysis with IKZF1 antibody in latent and PMA stimulated J-Lat 11.1 cells as indicated. Data are presented as % input, error bars represent the standard deviation (SD) of two separate real-time PCR measurements. (B-C) ChIP-qPCR using antibodies specific for distinct histone marks in latent and PMA stimulated J-Lat 11.1 cells as indicated; total histone H3 (B), H3K4me3 (C), H3K27me3 (D), H3K9me3 (E). Total histone H3 data (G) are presented as % input mean (±SD), data corresponding to histone marks (C-E) are expressed as fold change over H3 signal. Error bars represent the standard deviation (SD) of two separate real-time PCR measurements.

Figure S8


J-Lat 11.1 replicate
A
IKZF1
% Input
sh Ctrl
sh IKZF1
Nuc0
HS1
HS1/Nuc1
Nuc1/HS2
HS2/Nuc2
Nuc2
HIV genome
B
SUZ12
% Input
C
IKZF1 endogenous targets
Replicate 1
Replicate 2
BAD
IER2
PLIN3
SMAD7
D
H3
H3 normalized
E
H3K27me3
F
H3K4me3
G
H3K9me3
H
bp
10000
3000
1000
500
100

Supplementary Figure 8. IKZF1, required for maintenance of HIV-1 latency in J-Lat 11.1 cells, binds downstream of the latent HIV-1 5'LTR to promote PRC2/PRC1 recruitement and the establishment of a repressive chromatin environment. (A) ChIP qPCR analysis using antibody against IKZF1 in J-Lat 11.1 cells transduced with scramble shRNA (shControl) and shIKZF1 probing binding to the HIV-1 5'LTR. (B) ChIP-qPCR analysis with SUZ12 in J-Lat 11.1 cells transduced with scramble shRNA (shControl) and shIKZF1 at the HIV-1 5'LTR. Data in (A) and (B) are presented as % input, error bars represent the standard deviation (SD) of two separate real-time PCR measurements. (C) ChIP qPCR analysis using antibody against IKZF1 in J-Lat 11.1 cells transduced with scramble shRNA (shControl) and shIKZF1 probing binding of IKZF1 to endogenous IKZF target genes BAD, IER2, PLIN3 and SMAD7 (Song et al., 2016). (D-G) ChIP-qPCR using antibodies specific for distinct histone marks in J-Lat 11.1 cells transduced with scramble shRNA (shControl) and shIKZF1; total histone H3 (D), H3K27me3 (E), H3K4me3 (F), H3K9me3 (G). Total histone H3 data (D) are represented as % input mean (±SD), histone marks data (E-G) are expressed as fold change over H3 signal. Error bars represent the standard deviation (SD) of two separate real-time PCR measurements. (H) Representative agarose gels demonstrating range of DNA size resulting from sonication of chromatin used in the ChIP experiments presented in the manuscript.

Figure S9


J-Lat A2
A
Fold increase % GFP
Viability
sh Ctrl
sh IKZF1 #1
sh IKZF1 #2
B
Relative expression normalized to GAPDH
sh Control
sh IKZF1 #1
sh IKZF1 #2
IKZF1
Tat
GFP
CyPA
C
IKZ F1
% input
sh Ctrl
sh IKZF1
D
H 3
E
H 3 K27 me3
normalized to H3
F
H 3 K4 me3
G
H 3 K9 me3
Nuc0
HS1
HS1/Nuc1
Nuc1/HS2
HS2
Nuc2
HIV genome

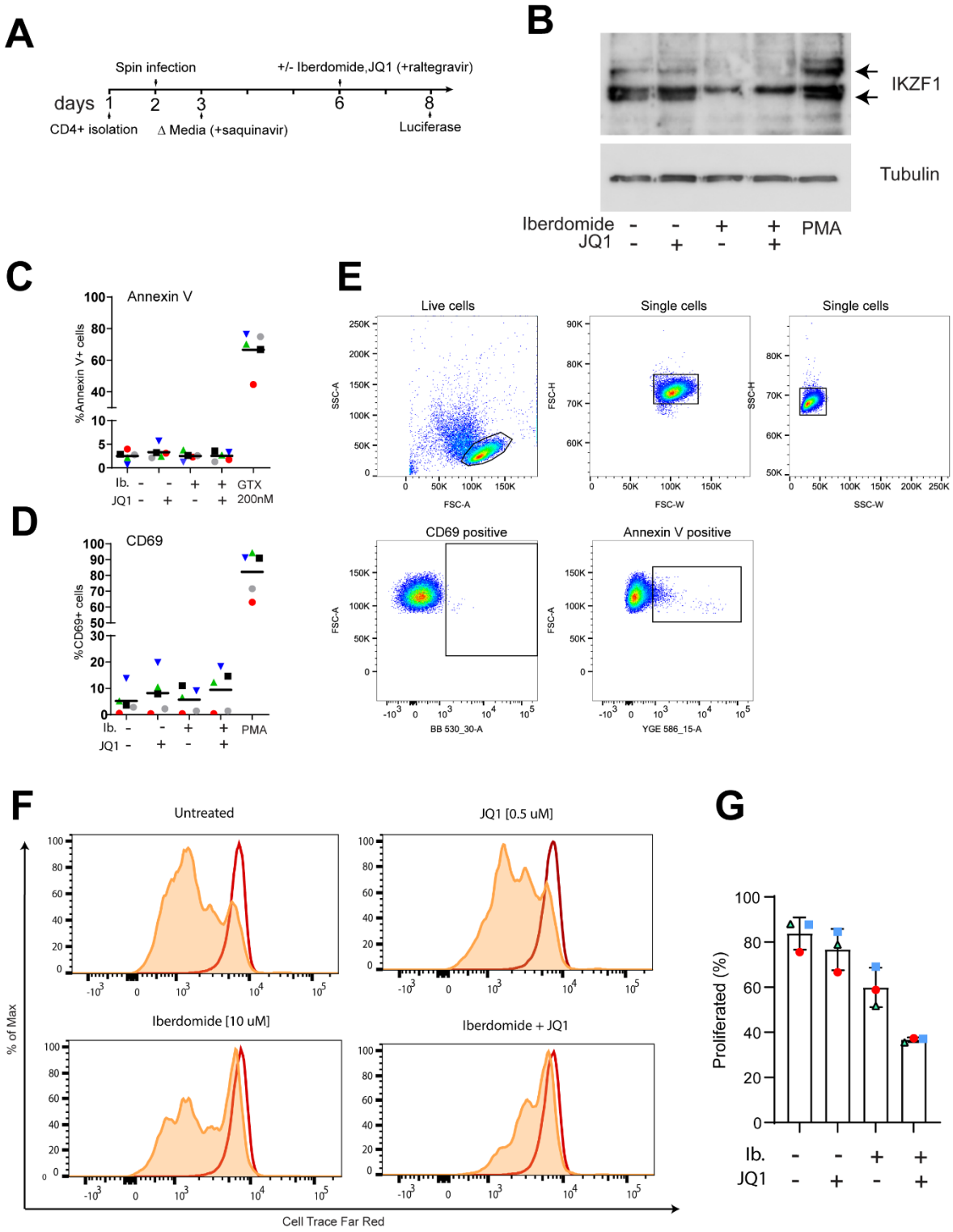
Supplementary figure 9. IKZF1, required for maintenance of HIV-1 latency in J-Lat A2 cells, binds downstream of the latent HIV-1 5'LTR to promote PRC2 recruitment

(A) Bar plot showing the fold increase in % GFP positive cells (left y-axes) measured by FACS analysis, following IKZF1 depletion in J-Lat A2 cells with two different shRNA constructs (#1 and #2). The right y-axis represents the percentage of live cells. Data are the mean of two independent experiments (±SD).

(B) qRT-PCR analysis measuring expression of HIV genes (pol, GFP, tat) in J-lat A2 cells transduced with scramble shRNA (sh Control) and sh IKZF1 #1 and #2. Data, normalized to GAPDH are represented as fold enrichment over sh Control and are the mean of three independent experiments (±SEM). Statistical significance was calculated using a ratio paired t test, * – p<0,05; **– p<0,01– ***p<0,001.

(C) ChIP qPCR analysis with IKZF1 antibody in J-Lat A2 cells transduced with scramble shRNA (shControl) and shIKZF1. Data are presented as % input, error bars represent the standard deviation (SD) of two separate real-time PCR measurements. (D-E) Histone marks ChIP qPCR analysis of the HIV-1 5' LTR in J-Lat A2 cells transduced with scramble shRNA (shControl) and shIKZF1; (D) Total histone H3 (E) H3K27me3 (F) H3K4me3 (G) H3K9me3. Total histone H3 data (D) are presented as % input mean (±SD), histone marks data (E-G) are expressed as fold change over H3 signal. Error bars represent the standard deviation (SD) of two separate real-time PCR measurements.

Figure S10


A
Spin infection
+/- Iberdomide,JQ1 (+raltegravir)
days 1 2 3 6 8
CD4+ isolation
Δ Media (+saquinavir)
Luciferase
B
IKZF1
Tubulin
Iberdomide - - + + PMA
JQ1 - + - +
C
Annexin V
%Annexin V+ cells
Ib. - - + + GTX
JQ1 - + - + 200nM
D
CD69
%CD69+ cells
Ib. - - + + PMA
JQ1 - + - +
E
Live cells
Single cells
Single cells
SSC-A
FSC-A
FSC-H
FSC-W
SSC-H
SSC-W
CD69 positive
Annexin V positive
BB 530_30-A
YGE 586_15-A
F
Untreated
JQ1 [0.5 uM]
Iberdomide [10 uM]
Iberdomide + JQ1
% of Max
Cell Trace Far Red
G
Proliferated (%)
Ib. - - + +
JQ1 - + - +

Supplementary figure 10 Iberdomide treatment of primary CD4+ T cells causes depletion of IKZF1 at the protein levels, reduces proliferation but does not cause cytotoxicity or induce T cell activation.

(A) schematic representation of the protocol for HIV-1 latency establishment in primary human CD4+ T cells.

(B) Western blot analysis using an antibody specific for IKZF1 indicates depletion of IKZF1 at the protein level in CD4+ T cells, upon treatment with iberdomide, JQ1 or a combination of both compounds as indicated. PMA is used as a control LRA. α-Tubulin is used as a loading control.

(C) Percentage of cells expressing the Annexin V marker of apoptosis in primary CD4+ T cells treated with iberdomide for 24 hours, JQ1 and the combination of both compounds. Treatment with a toxic concentration of Gliotoxin (GTX) 200nM was used as a positive control

(D) Percentage of cells expressing the CD69 marker of cell activation in primary CD4+ T cells treated with iberdomide for 24 hours, JQ1 and the combination of both compounds.

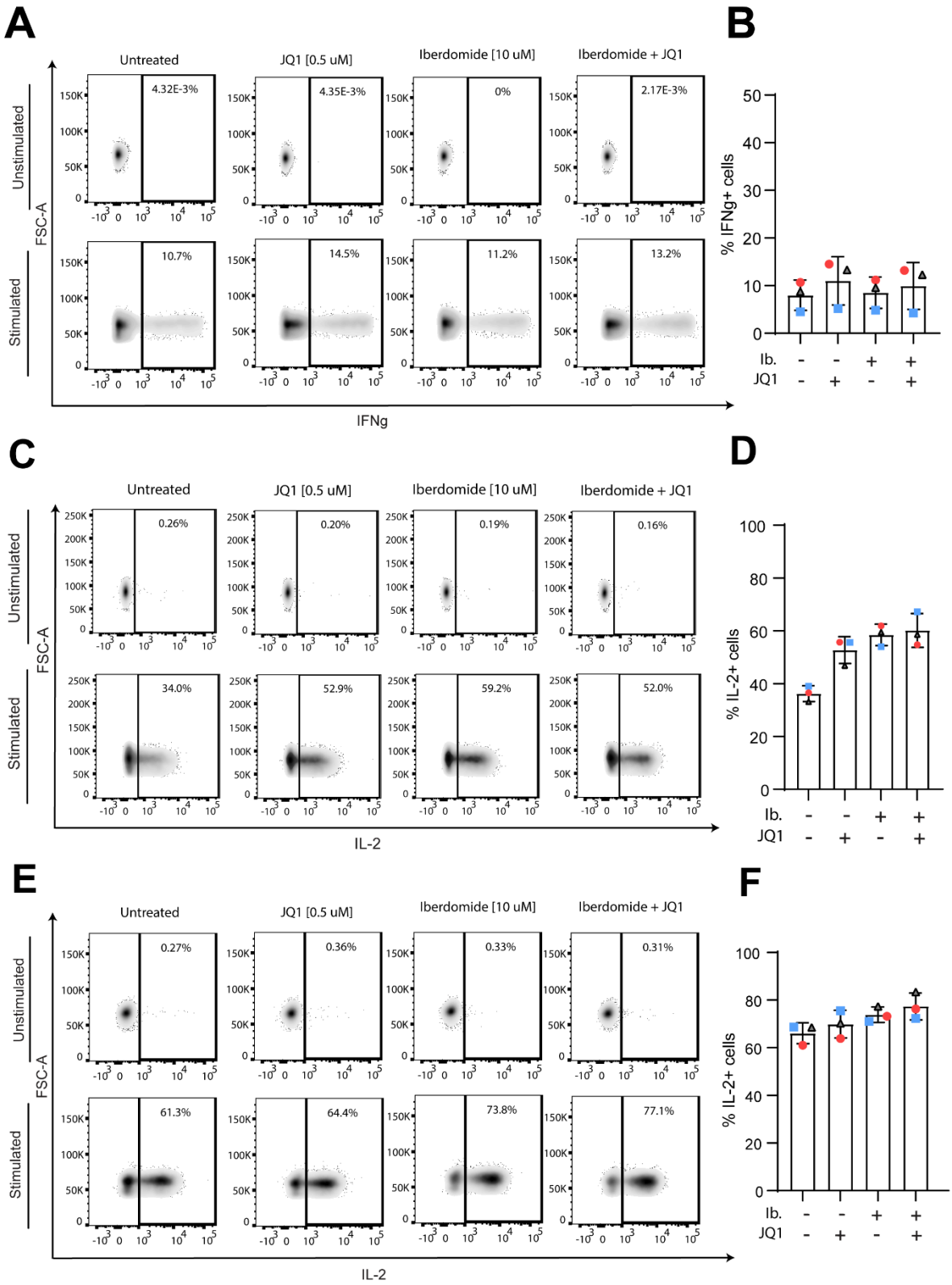
Experiments were performed in uninfected cells obtained from 5 healthy donors. Treatment with PMA was used as a positive control. Bars represent the average of experiments performed on samples deriving from two healthy donors.

(E) Representative flow cytometry plots and gating strategy for annexin V and CD69 staining.

(F) Representantive histogram of proliferative capacity of unstimulated or aCD3/CD28 stimulated CD8+ T cells in the presence or absence of LRAs. Cells were stained with a proliferation dye and analyzed 72 hours later by flow cytometry. Dividing cells show decreased intensity of proliferaiton dye as it becomes diluted upon cell division.

(G) Percentage of proliferated CD8+ T cells from 3 healthy donors as described in C.

Figure S11


A
Untreated
JQ1 [0.5 uM]
Iberdomide [10 uM]
Iberdomide + JQ1
Unstimulated
Stimulated
FSC-A
IFNg
4.32E-3%
4.35E-3%
0%
2.17E-3%
10.7%
14.5%
11.2%
13.2%
B
% IFNg+ cells
Ib.
JQ1
C
IL-2
0.26%
0.20%
0.19%
0.16%
34.0%
52.9%
59.2%
52.0%
D
% IL-2+ cells
E
0.27%
0.36%
0.33%
0.31%
61.3%
64.4%
73.8%
77.1%
F

Supplementary figure 11 Effect of iberdomide treatment alone and in combination with JQ1 on T cell toxicity, activation, proliferation capacity and effector function

(A) Representative flow cytometry plots (left panel) of INF-g production analysis in unstimulated and stimulated primary CD8+ T cells. . Cells were treated as indicated for 18 hours followed by PMA/Ionomycin stimulation for 7 hours in the presence of a protein transport inhibitor or remained unstimulated. IL-2 production was assessed by intracellular staining and analyzed by flow cytometry.
(B) Percentage of INF-g producing CD8+ T cells from 3 healthy donors as described in B.
(C) Representative flow cytometry plots (left panel) of IL2 production analysis in unstimulated and stimulated primary CD8+ T cells. Cells were treated as indicated for 18 hours followed by PMA/Ionomycin stimulation for 7 hours in the presence of a protein transport inhibitor or remained unstimulated. IL-2 production was assessed by intracellular staining and analyzed by flow cytometry.
(D) Percentage of IL-2 producing CD8+ T cells from 3 healthy donors as described in C
(E) Representative flow cytometry plots (left panel) of IL2 production analysis in unstimulated and stimulated primary CD8+ T cells. . Cells were treated as indicated for 18 hours followed by PMA/Ionomycin stimulation for 7 hours in the presence of a protein transport inhibitor or remained unstimulated. IL-2 production was assessed by intracellular staining and analyzed by flow cytometry.
(F) Percentage of IL-2 producing CD8+ T cells from 3 healthy donors as described in E

Supplementary Table 1 dCas9 V5/Histone ChIP sequencing peak calling summary

Table S1

Location coordinates	Reads V5/Histone	Reads_INPUT	avg_reads_V5/Histone	avg_reads_INPUT	avg_reads_V5/Histone	avg_reads_INPUT	fold_enrichment	fold_enrichment_norm	significance	fdr	qc_flag	pileup	saturation_treatment	saturation_control	MACS_fe
K03455.1:324-1353	163	0	15.8252	0.09709	15.8252	0	7.34873	7.34873	125.021	120.338	33	68	0.17952	0.0011	48.8854
K03455.1:7469-7950	58	1	12.0332	0.20747	12.0332	0.20747	5.85798	5.85798	59.4773	55.3671	33	44	0.16111	0.00278	24.244
chr1:163642157-163642744	52	0	8.84354	0.17007	8.84354	0	5.70044	5.70044	69.5395	65.3325	33	42	0.11159	0.00215	30.4648
chr13:106929710-106930148	35	0	7.97267	0	7.97267	0	5.12928	5.12928	40.4553	36.5264	33	27	0.11041	0	19.8376
chr18:69452480-69452979	32	0	6.4	0.2	6.4	0	5	5	42.3036	38.3524	33	28	0.08466	0.00265	20.546
chr10:10107793-10108316	30	0	5.72519	0	5.72519	0	4.90689	4.90689	36.4838	32.6128	33	25	0.07463	0	18.2645
K03455.1:4591-5133	29	0	5.3407	0	5.3407	0	4.85798	4.85798	31.4541	27.6399	33	22	0.06888	0	16.2951
chrX:79904028-79904440	25	0	6.05327	0	6.05327	0	4.64386	4.64386	33.2202	29.3807	33	23	0.08591	0	17.0036
chr10:81712946-81713497	49	0	8.87681	0.36232	8.87681	0	4.61471	5.61471	81.8569	77.5484	33	48	0.11395	0.00465	34.7157
chr10:42527461-42527960	47	1	9.4	0.4	9.4	0.2	4.55459	5.55459	65.5112	61.3377	33	40	0.12434	0.00529	29.0479
chr14:27677712-27678319	46	1	7.56579	0.32895	7.56579	0.16447	4.52356	5.52356	54.3442	50.2743	33	41	0.09465	0.00412	22.6278
chr10:67391032-67391407	21	0	5.58511	0.26596	5.58511	0	4.39232	4.39232	24.5825	20.8832	33	18	0.08268	0.00394	13.4612
chr8:83440464-83441020	53	0	9.51526	0.5386	9.51526	0	4.14296	5.72792	88.1367	83.7771	33	51	0.12184	0.0069	36.8412
chr5:45664104-45664588	35	0	7.21649	0.41237	7.21649	0	4.12928	5.12928	49.8391	45.809	33	32	0.09642	0.00551	23.38
chr11:33625473-33625961	35	0	7.15746	0.409	7.15746	0	4.12928	5.12928	42.3036	38.3524	33	28	0.09537	0.00545	20.546
chr2:42995018-42995354	17	1	5.04451	0.29674	5.04451	0.29674	4.08746	4.08746	21.2761	17.6472	33	16	0.07907	0.00465	12.0442
chr4:120251096-120251584	32	0	6.54397	0.409	6.54397	0	4	5	46.0438	42.0515	33	30	0.08719	0.00545	21.963
chr13:67564368-67564889	38	0	7.27969	0.57471	7.27969	0	3.66297	5.24793	53.4639	49.4081	33	35	0.095	0.0075	24.4341
chr5:164248717-164249084	67	1	18.2065	1.63043	18.2065	0.27174	3.48113	6.06609	101.348	96.8835	33	67	0.27236	0.02439	36.6354
chr2:117593970-117594523	44	0	7.94224	0.72202	7.94224	0	3.45943	5.45943	65.5112	61.3377	33	40	0.10185	0.00926	29.0479
chr2:167082840-167083383	31	1	5.69853	0.55147	5.69853	0.18382	3.36923	4.9542	31.4541	27.6399	33	22	0.07346	0.00711	16.2951
chr10:115327573-115327989	49	0	11.7506	1.19904	11.7506	0	3.29278	5.61471	83.9417	79.6163	33	49	0.1661	0.01695	35.4242
chr4:57790810-57791142	19	1	5.70571	0.6006	5.70571	0.3003	3.24793	4.24793	22.9176	19.2508	33	17	0.09005	0.00948	12.7527
chr7:57945316-57945811	37	0	7.45968	0.80645	7.45968	0	3.20945	5.20945	46.0438	42.0515	33	30	0.09893	0.0107	21.963

Table S2

			-PMA	+PMA	HIV interaction
DNA associated	Bait	dCas9			
	Chromatin remodeling	EED			Nguyen et al., 2017
		PRMT1			Hassa et al., 2008
		MTA2			Cismasiu et al., 2008
		RBBP4			Knoener et al., 2017
		UHRF1			Verdikt et al., 2018
		CHD4			Gastellegui et al., 2011
		HAT1			Galey et al., 2013
		ACTL6A			
	Chromosome organization	HIST1H1B			
		RPA2			Amacker et al., 1997
		DNMT1			Trejbalova et al., 2016
		RPA1			
		H1F0			
		HMGB1			
	DNA replication	MCM2			López-Huertas et al., 2013
		MCM4			
		MCM5			
		MCM6			Jarboui et al., 2012
		MCM7			Jarboui et al., 2012
		RPA3			Amacker et al., 1997
	Structural	BOLA2B			
		FHL1			
		HMGB2			
		IK			
		RAD52			
		PHF6			
		TOP2A			
		RFC4			
		SMC4			
		SAFB			Milev et al., 2012
		TOP1			Shoya et al., 2003
		XRCC6			Li et al., 2001
	Transcription complex	TCERG1			Coiras et al., 2013
		SSRP1			Huang et al., 2015
	Transcription factor	EDF1			Mariotti et al., 2008
		E4F1			
		RFXAP			
		BCCIP			
		ZC3H8			
		ZFR			Milev et al., 2012
		TFPT			
		CGGBP1			
		LEF1			Sharma et al., 2011
		FUBP3			
		FUBP1			
		TRIM28			Milev et al., 2012
		RB1			Prasad and Shanmgam 1993
	Ubiquitin proteasome system	CAND1			
		USP7			

		-PMA	+PMA	HIV interaction
RNA associated	DNAJC8			
	DDX42			
	ELAVL1			Milev et al., 2012
	HNRNPLL			
	PABPN1			Milev et al., 2012
	PTBP1			
	QKI			
	RBM15			Naji et al., 2011
	RBM39			Naji et al., 2011
	RBM7			
	RBMXL1			
	SNRPA			Simmons et al., 2011
	TARDBP			Ou et al., 1995
	TIA1			
	ZC3H11A			
	ADAR			Jarboui et al., 2012
	DDX39B			Milev et al., 2012
	LUC7L3			
	RBMX			Milev et al., 2012
	SNRNP200			Milev et al., 2012
	SNRPF			Naji et al., 2011
	SNW1			Brès et al., 2005
	EFTUD2			
	SARNP			Lin et al., 2014
	HNRNPDL			
	HNRNPUL2			
	API5			
	FAM98B			Naji et al., 2011
	MYEF2			
	RBM45			
	RNASEH2A			
	SSB			
	STRBP			
	PRPF8			Milev et al., 2012
	SF3A3			Naji et al., 2011
	DHX15			
	SF3B2			
	SRSF4			Milev et al., 2012
	TRA2B			Paz et al., 2014
	KHSRP			Barrero et al., 2013
	PRPF40A			
	PRPF6			Knoener et al., 2017
	DDX46			
	CPSF7			Milev et al., 2012
	KHDRBS3			Kim et al., 2013
	ZNF326			
	MBNL1			
	DAZAP1			Naji et al., 2012
	RALY			
	KHDRBS1			Marsh et al., 2008
	AKAP8L			Jager et al., 2011
	ALYREF			Jager et al., 2011
	TNPO1			Naji et al., 2012
	CPSF6			Sowd et al., 2016
	MATR3			Sarracino et al., 2018
	HNRNPA0			Milev et al., 2012
	SNRNP40			
	SNRPD3			König et al., 2008
	SNRPA1			Lin et al., 2014
	HNRNPUL1			Wagner et al., 2014
	HNRNPH3			Milev et al., 2012
	SNRPB			Fong et al., 2001
	SNRPD2			Fong et al., 2001
	DDX1			Milev et al., 2012
	CELF2			
	UCHL5			Rathore et al., 2020
	RBM12			
	NCOA5			
	U2AF1L5			
	SRSF1			Milev et al., 2012
	SRSF10			Milev et al., 2012
	SRSF6			Swanson et al., 2010

Log 2 relative protein abundance

18 20 22 24 26 28 30 32

missing values

Supplementary Table 2 List of factors enriched on both the latent and active state of the HIV-1 promoter. The table displays selected and functionally classified hits (n=122) identified in both experimental conditions with similar scores, including the HA-V5-FLAG-dCas9 bait and potential non-specific binding contaminants.

Table S3

			-PMA	+PMA	HIV interaction
DNA associated	Chromatin remodelling	CHD1L			
	Chromosome organization	H3F3B			Naji et al., 2012
		HMGB3			
		MAD2L1			Liu et al., 2014
		MTF2			
		PCNA			Zhoue and He, 2004
		WDR5			
		HMGA1			
	DNA metabolism	GINS3			
		MPG			Jager et al., 2011
		NUDT16L1			
	Structural	TOP2B			Lokeswara et al., 2013
	Transcription factor	BCLAF1			Zhou et al., 2008
		CBFB			Anderson and Harris 2015
		FLI1			
		GTF2B			Zhang et al., 2000
		GTF2E2			Zhou et al., 1996
		GTF2I			Taylor et al., 2011
		LDB1			
		TAF3			Zhou and Sharp 1995
		TBPL1			
		ZNF48			
		ZNF638			
		ZNF653			
		PRDM8			
	Transcription complex	MED20			Zhou et al., 2008
		PCBP3			Kim et al., 2013
		SSU72			Chen et al., 2014
		TAF15			Kashanchi et al., 1994
	Ubiquitin proteasome system	PSMA6			
		PSMD10			
		RNF114			
		UBA1			
		UBE2S			
		UBE3A			
		USP11			

		-PMA	+PMA	HIV interaction
RNA associated	CSTF3			Naji et al., 2012
	NUDT21			Barrero et al., 2013
	RNF40			
	SNRNP70			Naji et al., 2012
	CSTF1			
	CTNNBL1			
	DBR1			
	GPKOW			
	POLR2G			Agostini et al., 1996
	RBM10			
	RBM15B			Uranishi et al., 2009
	RBM27			
	SAFB2			
	SAP18			La Porte et al., 2016
	SLTM			
	SON			Le Sage et al., 2015
	ZRANB2			
	PRPF4			
	RAVER1			
	SNRNP40			Naji et al., 2012
	U2SURP			Jarboui et al., 2012
	NCBP1			
	SNRPE			Naji et al., 2012
	LUC7L			
	RBM5			
	C7orf55-LUC7L2			
	RBM6			Konig et al., 2008
	RNASEH2C			
	RTCA			
	DDX41			
	DHX40			
	GTF2F2			Cicala et al., 2002
	SRSF2			Jarboui et al., 2012
	U2AF2			Luznik et al., 1995
	SF3B1			Kyei et al., 2018
	RBM17			
	SF3B6			
	STAT3			Fan et al., 2015
	ANP32A			Naji et al., 2012
	ANP32E			Naji et al., 2012
	KPNA2			Naji et al., 2012
	CPSF1			Milev et al., 2012
	PARK7			Lopez-Huertas et al., 2013
	CPSF3			de la Vega et al., 2007
	MAGOH			
	MAGOHB			Jarboui et al., 2012
	C14orf166			
	FKBP4			

Log 2 relative protein abundance

18 20 22 24 26 28 30 32

missing values

Supplementary Table 3 List of putative factors enriched on the active (+PMA) HIV-1 promoter. The table displays selected and functionally classified hits (n=84) identified by Catchet-MS in the +PMA state

Table S4

Name	Arithmetic Mean	Median	SD	IQR	CV	Sum	Minimum	Maximum	Count
Proteins - # AAs	625.9015755	449	763.2457635	472.5	1.219434163	3694697	18	34350	5903
Proteins - # Peptides	7.049974589	3	9.941955035	8	1.410211471	41616	1	167	5903
Proteins - # Peptides (by Search Engine): A4 Mascot	7.049974589	3	9.941955035	8	1.410211471	41616	1	167	5903
Proteins - # Protein Groups	1.075385397	1	0.804874751	0	0.748452372	6348	0	11	5903
Proteins - # Protein Unique Peptides	5.691004574	2	9.409635432	6	1.653422574	33594	0	166	5903
Proteins - # PSMs	85.45383703	12	238.3124063	52	2.78878532	504434	1	3565	5903
Proteins - # PSMs (by Search Engine): A4 Mascot	85.45383703	12	238.3124063	52	2.78878532	504434	1	3565	5903
Proteins - # Unique Peptides	6.197187871	3	9.440269279	7	1.523315006	36582	0	166	5903
Proteins - Abundance Ratio (log2): (F27) / (F1)	-3.645077669	-4.022450988	3.256357008	5.050406102	-0.893357372	-18706.5386	-6.64385619	6.64385619	5132
Proteins - Abundance Ratio (log2): (F28) / (F2)	-3.080770486	-3.058825361	3.274785095	5.829354752	-1.062976002	-15810.51413	-6.64385619	6.64385619	5132
Proteins - Abundance Ratio (log2): (F29) / (F1)	-3.489604345	-4.214873516	3.490343712	5.662728971	-1.000211877	-17908.6495	-6.64385619	6.64385619	5132
Proteins - Abundance Ratio (log2): (F3) / (F1)	-3.7296844	-4.124734329	3.018716311	4.452048224	-0.809375804	-19140.74034	-6.64385619	6.64385619	5132
Proteins - Abundance Ratio (log2): (F30) / (F2)	-3.339627058	-3.54799393	3.405717493	5.870080586	-1.019789765	-17138.96606	-6.64385619	6.64385619	5132
Proteins - Abundance Ratio (log2): (F4) / (F2)	-3.658679244	-4.105673663	3.098987039	4.64434842	-0.847023429	-18776.34188	-6.64385619	6.64385619	5132
Proteins - Abundance Ratio: (F27) / (F1)	2.625520118	0.049843132	13.37986578	0.269872962	5.096081985	13474.16925	0	100	5132
Proteins - Abundance Ratio: (F28) / (F2)	2.531365802	0.100205681	12.73420222	0.472421481	5.030565796	12990.96929	0	100	5132
Proteins - Abundance Ratio: (F29) / (F1)	2.841371373	0.0364059	13.46780769	0.415348333	4.739897015	14581.91789	0	100	5132
Proteins - Abundance Ratio: (F3) / (F1)	2.387959743	0.050110282	13.29634644	0.183140747	5.568078139	12255.0094	0	100	5132
Proteins - Abundance Ratio: (F30) / (F2)	2.226408549	0.060088502	11.05184387	0.45543572	4.963978365	11425.92867	0	100	5132
Proteins - Abundance Ratio: (F4) / (F2)	2.611594288	0.051448102	13.91148538	0.205106275	5.326817202	13402.70189	0	100	5132
Proteins - Abundance: F1: Sample, control	59131448.73	1996852.621	569381322.9	13897669.05	9.629077845	3.03463E+11	0	35892627283	5132
Proteins - Abundance: F2: Sample, PMA	59208972.49	1739722.451	941523815.1	12793741.41	15.9017084	3.0386E+11	0	64933805348	5132
Proteins - Abundance: F27: Sample, control	1896039.707	5601.295166	34358041.65	92306.19226	18.12095048	9730475779	0	1684339964	5132
Proteins - Abundance: F28: Sample, PMA	3097478.294	16661.19287	51491849.94	178842.3175	16.62379686	15896258606	0	2854568181	5132
Proteins - Abundance: F29: Sample, control	864449.2629	2883.788208	14068899.68	60276.5614	16.27498603	4436353617	0	623270265.7	5132
Proteins - Abundance: F3: Sample, control	6563848.329	17273.24548	227245221.5	179896.2858	34.62073012	33685669625	0	15861526519	5132
Proteins - Abundance: F30: Sample, PMA	919751.1687	3881.058167	14206434.72	60436.61716	15.44595452	4720162998	0	740924934.1	5132
Proteins - Abundance: F4: Sample, PMA	7890541.895	19108.25586	273278590.4	203837.4481	34.63369107	40494261008	0	19132853673	5132
Proteins - Abundances (Grouped) Count: F1	0.96745908	1	0.17744899	0	0.183417566	4965	0	1	5132
Proteins - Abundances (Grouped) Count: F2	0.964925955	1	0.183984925	0	0.190672584	4952	0	1	5132
Proteins - Abundances (Grouped) Count: F27	0.585931411	1	0.492608442	1	0.840727144	3007	0	1	5132
Proteins - Abundances (Grouped) Count: F28	0.656664069	1	0.474868729	1	0.72315321	3370	0	1	5132
Proteins - Abundances (Grouped) Count: F29	0.538776306	1	0.498542705	1	0.925324109	2765	0	1	5132
Proteins - Abundances (Grouped) Count: F3	0.684333593	1	0.464826019	1	0.679238932	3512	0	1	5132
Proteins - Abundances (Grouped) Count: F30	0.55806703	1	0.496665165	1	0.889974032	2864	0	1	5132
Proteins - Abundances (Grouped) Count: F4	0.679462198	1	0.466728793	1	0.686909138	3487	0	1	5132
Proteins - Abundances (Grouped) CV [%]: F1	0	0	0	0		0	0	0	5132
Proteins - Abundances (Grouped) CV [%]: F2	0	0	0	0		0	0	0	5132
Proteins - Abundances (Grouped) CV [%]: F27	0	0	0	0		0	0	0	5132
Proteins - Abundances (Grouped) CV [%]: F28	0	0	0	0		0	0	0	5132
Proteins - Abundances (Grouped) CV [%]: F29	0	0	0	0		0	0	0	5132
Proteins - Abundances (Grouped) CV [%]: F3	0	0	0	0		0	0	0	5132
Proteins - Abundances (Grouped) CV [%]: F30	0	0	0	0		0	0	0	5132
Proteins - Abundances (Grouped) CV [%]: F4	0	0	0	0		0	0	0	5132
Proteins - Abundances (Grouped): F1	61325236.83	2070936.237	590505479.3	14413275.24	9.629077845	3.14721E+11	0	37224250649	5132
Proteins - Abundances (Grouped): F2	59208972.49	1739722.451	941523815.1	12793741.41	15.9017084	3.0386E+11	0	64933805348	5132
Proteins - Abundances (Grouped): F27	69605376.47	205628.7414	1261315580	3388645.943	18.12095048	3.57215E+11	0	61833682499	5132
Proteins - Abundances (Grouped): F28	65854165.71	354226.5198	1094746273	3802290.282	16.62379686	3.37964E+11	0	60689757327	5132
Proteins - Abundances (Grouped): F29	71814646.11	239572.4521	1168782362	5007511.848	16.27498603	3.68553E+11	0	51778554837	5132
Proteins - Abundances (Grouped): F3	69886905.22	183912.4869	2419535684	1915399.937	34.62073012	3.5866E+11	0	1.68882E+11	5132
Proteins - Abundances (Grouped): F30	67935838.7	286667.6885	1049333874	4464046.814	15.44595452	3.48647E+11	0	54727146342	5132
Proteins - Abundances (Grouped): F4	70339478.96	170338.7142	2436115784	1817089.381	34.63369107	3.60982E+11	0	1.70558E+11	5132
Proteins - Abundances (Normalized): F1: Sample, control	61325236.83	2070936.237	590505479.3	14413275.24	9.629077845	3.14721E+11	0	37224250649	5132
Proteins - Abundances (Normalized): F2: Sample, PMA	59208972.49	1739722.451	941523815.1	12793741.41	15.9017084	3.0386E+11	0	64933805348	5132
Proteins - Abundances (Normalized): F27: Sample, control	69605376.47	205628.7414	1261315580	3388645.943	18.12095048	3.57215E+11	0	61833682499	5132
Proteins - Abundances (Normalized): F28: Sample, PMA	65854165.71	354226.5198	1094746273	3802290.282	16.62379686	3.37964E+11	0	60689757327	5132
Proteins - Abundances (Normalized): F29: Sample, control	71814646.11	239572.4521	1168782362	5007511.848	16.27498603	3.68553E+11	0	51778554837	5132
Proteins - Abundances (Normalized): F3: Sample, control	69886905.22	183912.4869	2419535684	1915399.937	34.62073012	3.5866E+11	0	1.68882E+11	5132
Proteins - Abundances (Normalized): F30: Sample, PMA	67935838.7	286667.6885	1049333874	4464046.814	15.44595452	3.48647E+11	0	54727146342	5132
Proteins - Abundances (Normalized): F4: Sample, PMA	70339478.96	170338.7142	2436115784	1817089.381	34.63369107	3.60982E+11	0	1.70558E+11	5132
Proteins - Abundances Count: F1: Sample, control	6.950701481	3	10.45416281	7	1.504044281	35671	0	192	5132
Proteins - Abundances Count: F2: Sample, PMA	6.980904131	3	10.55383253	7	1.511814564	35826	0	193	5132
Proteins - Abundances Count: F27: Sample, control	2.914848012	1	6.213724014	3	2.13174889	14959	0	136	5132
Proteins - Abundances Count: F28: Sample, PMA	3.46745908	1	6.650409091	4	1.917948831	17795	0	143	5132
Proteins - Abundances Count: F29: Sample, control	2.273967264	1	4.790582305	2	2.106706803	11670	0	107	5132
Proteins - Abundances Count: F3: Sample, control	3.694466095	1	7.060157504	4	1.911008877	18960	0	137	5132
Proteins - Abundances Count: F30: Sample, PMA	2.677123928	1	5.843138828	3	2.182617983	13739	0	130	5132
Proteins - Abundances Count: F4: Sample, PMA	3.604247857	1	6.883036344	4	1.909701169	18497	0	135	5132
Proteins - calc. pI	7.21698289	7.01	1.629887877	2.65	0.225840618	42601.85	3.78	12.41	5903
Proteins - Coverage [%]	16.39207976	9.803921569	16.71646786	23.27724859	1.019789319	96762.44682	0.061135371	94.59854015	5903
Proteins - Coverage [%] (by Search Engine): A4 Mascot	16.39207976	9.803921569	16.71646786	23.27724859	1.019789319	96762.44682	0.061135371	94.59854015	5903
Proteins - MW [kDa]	69.923392	50.114	85.01040993	53.1425	1.215764961	412757.783	2.172	3813.652	5903
Proteins - Score Mascot: A4 Mascot	1992.34036	241.3936976	5841.714586	1166.60379	2.932086658	11760785.14	0	111969.5972	5903
Protein Groups - # Peptides	8.097394276	4	10.65309734	8	1.31562043	37912	1	167	4682
Protein Groups - # Proteins	1.355830842	1	1.94064817	0	1.431335024	6348	1	44	4682
Protein Groups - # PSMs	89.42994447	15	239.6488314	62	2.679738121	418711	1	3565	4682
Protein Groups - # Unique Peptides	7.451089278	4	10.14664217	8	1.361766285	34886	1	166	4682
Peptide Groups - # Missed Cleavages	0.211967266	0	0.456436943	0	2.15333693	8470	0	2	39959
Peptide Groups - # Protein Groups	1.046072224	1	0.314178563	0	0.300341177	41800	0	11	39959
Peptide Groups - # Proteins	1.147751445	1	0.963569809	0	0.839528291	45863	1	53	39959
Peptide Groups - # PSMs	9.256287695	4	18.45459183	9	1.993735765	369872	1	1016	39959
Peptide Groups - Abundance Ratio (log2): (F27) / (F1)	-4.427275803	-6.64385619	3.201323989	4.184645862	-0.723091158	-162768.7949	-6.64385619	6.64385619	36765
Peptide Groups - Abundance Ratio (log2): (F28) / (F2)	-3.96321723	-6.64385619	3.417120885	5.089851971	-0.862208828	-145707.6814	-6.64385619	6.64385619	36765
Peptide Groups - Abundance Ratio (log2): (F29) / (F1)	-4.616874934	-6.64385619	3.327400265	3.949680358	-0.720704007	-169739.4069	-6.64385619	6.64385619	36765
Peptide Groups - Abundance Ratio (log2): (F3) / (F1)	-4.380463896	-6.385266372	3.039468743	3.908941314	-0.693869146	-161047.7551	-6.64385619	6.64385619	36765
Peptide Groups - Abundance Ratio (log2): (F30) / (F2)	-4.250586005	-6.64385619	3.45884849	4.848126938	-0.813734503	-156272.7945	-6.64385619	6.64385619	36765
Peptide Groups - Abundance Ratio (log2): (F4) / (F2)	-4.383447925	-6.64385619	3.08928167	3.945770081	-0.704760664	-161157.463	-6.64385619	6.64385619	36765
Peptide Groups - Abundance Ratio: (F27) / (F1)	2.138980967	0.01	12.34151993	0.127078462	5.76981288	78639.63525	0	100	36765
Peptide Groups - Abundance Ratio: (F28) / (F2)	2.487092839	0.01	12.73635452	0.260723468	5.120980739	91437.96821	0	100	36765
Peptide Groups - Abundance Ratio: (F29) / (F1)	2.203438043	0.01	12.12087239	0.074909761	5.500890949	81009.39967	0	100	36765
Peptide Groups - Abundance Ratio: (F3) / (F1)	1.888389535	0.01	11.74717175	0.110001574	6.22073546	69426.64127	0	100	36765
Peptide Groups - Abundance Ratio: (F30) / (F2)	2.299673368	0.01	11.85249127	0.205131313	5.153989014	84547.49137	0	100	36765
Peptide Groups - Abundance Ratio: (F4) / (F2)	1.945044238	0.01	11.77744248	0.114117072	6.05510263	71509.55141	0	100	36765
Peptide Groups - Abundance: F1: Sample, control	9768077.383	1002071.805	133899705.8	3782878.344	13.70788749	3.59123E+11	0	14225552205	36765
Peptide Groups - Abundance: F2: Sample, PMA	10130474.93	860133.4023	191378462	3147426.527	18.8913613	3.72447E+11	0	17123692824	36765
Peptide Groups - Abundance: F27: Sample, control	275952.5729	0	7670051.194	12220.73145	27.79481674	10145396342	0	1009310882	36765
Peptide Groups - Abundance: F28: Sample, PMA	476491.1356	0	9768870.495	25158.16406	20.50168359	17518196601	0	979450692.5	36765

Peptide Groups - Abundance: F29: Sample, control	121942.8356	0	3009821.78	4152.288574	24.68223545	4483228352	0	419899557.7	36765
Peptide Groups - Abundance: F3: Sample, control	951464.38	1966.508301	84377961.21	28388.42505	88.68220711	34980587931	0	15861501896	36765
Peptide Groups - Abundance: F30: Sample, PMA	137151.7057	0	3343645.749	7889.714355	24.37917728	5042382461	0	462384751.1	36765
Peptide Groups - Abundance: F4: Sample, PMA	1136416.391	1308.020508	101585114	28943.14844	89.39075045	41780348630	0	19132837467	36765
Peptide Groups - Abundances (Grouped) Count: F1	0.96496668	1	0.183866543	0	0.190541857	35477	0	1	36765
Peptide Groups - Abundances (Grouped) Count: F2	0.969862641	1	0.170967522	0	0.17628014	35657	0	1	36765
Peptide Groups - Abundances (Grouped) Count: F27	0.408459132	0	0.491555533	1	1.203438713	15017	0	1	36765
Peptide Groups - Abundances (Grouped) Count: F28	0.484292126	0	0.499759999	1	1.031939138	17805	0	1	36765
Peptide Groups - Abundances (Grouped) Count: F29	0.317149463	0	0.465372509	1	1.46736023	11660	0	1	36765
Peptide Groups - Abundances (Grouped) Count: F3	0.515735074	1	0.499759143	1	0.969022989	18961	0	1	36765
Peptide Groups - Abundances (Grouped) Count: F30	0.375492996	0	0.484256528	1	1.289655287	13805	0	1	36765
Peptide Groups - Abundances (Grouped) Count: F4	0.503386373	1	0.499995332	1	0.993263543	18507	0	1	36765
Peptide Groups - Abundances (Grouped) CV [%]: F1	0	0	0	0		0	0	0	36765
Peptide Groups - Abundances (Grouped) CV [%]: F2	0	0	0	0		0	0	0	36765
Peptide Groups - Abundances (Grouped) CV [%]: F27	0	0	0	0		0	0	0	36765
Peptide Groups - Abundances (Grouped) CV [%]: F28	0	0	0	0		0	0	0	36765
Peptide Groups - Abundances (Grouped) CV [%]: F29	0	0	0	0		0	0	0	36765
Peptide Groups - Abundances (Grouped) CV [%]: F3	0	0	0	0		0	0	0	36765
Peptide Groups - Abundances (Grouped) CV [%]: F30	0	0	0	0		0	0	0	36765
Peptide Groups - Abundances (Grouped) CV [%]: F4	0	0	0	0		0	0	0	36765
Peptide Groups - Abundances (Grouped): F1	10130474.93	1039248.861	138867410.6	3923223.857	13.70788749	3.72447E+11	0	14753322925	36765
Peptide Groups - Abundances (Grouped): F2	10130474.93	860133.4023	191378462	3147426.527	18.8913613	3.72447E+11	0	17123692824	36765
Peptide Groups - Abundances (Grouped): F27	10130474.93	0	281574694.2	448634.3876	27.79481674	3.72447E+11	0	37052738731	36765
Peptide Groups - Abundances (Grouped): F28	10130474.93	0	207691791.6	534877.0025	20.50168359	3.72447E+11	0	20823683683	36765
Peptide Groups - Abundances (Grouped): F29	10130474.93	0	250042767.5	344953.8883	24.68223545	3.72447E+11	0	34883410091	36765
Peptide Groups - Abundances (Grouped): F3	10130474.93	20937.89685	898392875.9	302258.5336	88.68220711	3.72447E+11	0	1.68881E+11	36765
Peptide Groups - Abundances (Grouped): F30	10130474.93	0	246972644.3	582760.1856	24.37917728	3.72447E+11	0	34153254636	36765
Peptide Groups - Abundances (Grouped): F4	10130474.93	11660.22337	905570756.4	258010.9209	89.39075045	3.72447E+11	0	1.70558E+11	36765
Peptide Groups - Abundances (Normalized): F1: Sample, control	10130474.93	1039248.861	138867410.6	3923223.857	13.70788749	3.72447E+11	0	14753322925	36765
Peptide Groups - Abundances (Normalized): F2: Sample, PMA	10130474.93	860133.4023	191378462	3147426.527	18.8913613	3.72447E+11	0	17123692824	36765
Peptide Groups - Abundances (Normalized): F27: Sample, control	10130474.93	0	281574694.2	448634.3876	27.79481674	3.72447E+11	0	37052738731	36765
Peptide Groups - Abundances (Normalized): F28: Sample, PMA	10130474.93	0	207691791.6	534877.0025	20.50168359	3.72447E+11	0	20823683683	36765
Peptide Groups - Abundances (Normalized): F29: Sample, control	10130474.93	0	250042767.5	344953.8883	24.68223545	3.72447E+11	0	34883410091	36765
Peptide Groups - Abundances (Normalized): F3: Sample, control	10130474.93	20937.89685	898392875.9	302258.5336	88.68220711	3.72447E+11	0	1.68881E+11	36765
Peptide Groups - Abundances (Normalized): F30: Sample, PMA	10130474.93	0	246972644.3	582760.1856	24.37917728	3.72447E+11	0	34153254636	36765
Peptide Groups - Abundances (Normalized): F4: Sample, PMA	10130474.93	11660.22337	905570756.4	258010.9209	89.39075045	3.72447E+11	0	1.70558E+11	36765
Peptide Groups - Abundances Count: F1: Sample, control	4.479722562	2	5.67021581	5	1.265751557	164697	0	213	36765
Peptide Groups - Abundances Count: F2: Sample, PMA	4.473167415	2	5.75199837	5	1.285889357	164456	0	221	36765
Peptide Groups - Abundances Count: F27: Sample, control	1.062341901	0	2.420654201	1	2.27860183	39057	0	139	36765
Peptide Groups - Abundances Count: F28: Sample, PMA	1.401332789	0	3.020985854	2	2.155794738	51520	0	142	36765
Peptide Groups - Abundances Count: F29: Sample, control	0.753189174	0	2.029354528	1	2.694349038	27691	0	105	36765
Peptide Groups - Abundances Count: F3: Sample, control	1.63244934	1	3.639521693	2	2.22948523	60017	0	234	36765
Peptide Groups - Abundances Count: F30: Sample, PMA	1.020563035	0	2.502313446	1	2.451895041	37521	0	139	36765
Peptide Groups - Abundances Count: F4: Sample, PMA	1.579355365	1	3.661589668	2	2.318407718	58065	0	249	36765
Peptide Groups - Charge (by Search Engine): A4 Mascot	2.27760955	2	0.515622043	0	0.226387373	91011	2	6	39959
Peptide Groups - Concatenated Rank (by Search Engine): A4 Mascot	1.006306464	1	0.09549788	0	0.0948994	40211	1	6	39959
Peptide Groups - Ions Score (by Search Engine): A4 Mascot	48.26798994	44.03	19.5713523	25.11	0.405472702	1928740.61	13.06	189.06	39959
Peptide Groups - m/z [Da] (by Search Engine): A4 Mascot	615.8574228	592.2871094	160.9139442	235.5726624	0.261284411	24609046.76	374.204	1368.588257	39959
Peptide Groups - Rank (by Search Engine): A4 Mascot	1.006306464	1	0.09549788	0	0.0948994	40211	1	6	39959
Peptide Groups - RT [min] (by Search Engine): A4 Mascot	27.42003655	27.33202667	11.91351497	19.89909226	0.434482097	1095677.24	6.371940223	58.84110185	39959
Peptide Groups - Search Engine Rank (by Search Engine): A4 Mascot	1.006306464	1	0.09549788	0	0.0948994	40211	1	6	39959
Peptide Groups - Sequence Length	12.26459621	12	4.285822242	5	0.349446665	490081	6	49	39959
Peptide Groups - Theo. MH+ [Da]	1390.766984	1314.701204	444.3190626	557.3550343	0.319477718	55573657.92	747.4610811	4852.389837	39959
Peptide Groups - Top Apex RT [min]	34.91930421	34.84970565	11.98924577	19.88060326	0.343341485	1299696.503	7.332193851	65.35504679	37220
Peptide Groups - ΔCn (by Search Engine): A4 Mascot	7.4136E-05	0	0.001667698	0	22.49511887	2.9624	0	0.05	39959
Peptide Groups - ΔM [ppm] (by Search Engine): A4 Mascot	-0.335248381	-0.600007916	1.699005782	1.398146383	-5.067901533	-13396.19004	-10.00784576	9.95517133	39959
Peptide Groups - Δm/z [Da] (by Search Engine): A4 Mascot	-0.000179905	-0.000338092	0.001148927	0.000782437	-6.386301978	-7.188822863	-0.008193627	0.009014794	39959
Peptide Groups - ΔScore (by Search Engine): A4 Mascot	0.858403769	0.9231	0.194457775	0.1935	0.226534159	34300.9562	0	1	39959
PSMs - # Missed Cleavages	0.189973829	0	0.432670676	0	2.2775278	70266	0	2	369872
PSMs - # Protein Groups	1.132042977	1	0.563864963	0	0.498095015	418711	0	11	369872
PSMs - # Proteins	1.363806939	1	1.772538808	0	1.299699215	504434	1	53	369872
PSMs - Aligned m/z [Da]	610.9495306	592.848056	150.1197701	216.8768911	0.245715501	225973124.8	374.204	1368.588922	369872
PSMs - Apex RT [min]	34.82322869	34.91168594	11.78910601	19.04831871	0.338541441	11923438.68	7.063932419	68.82004149	342399
PSMs - Charge	2.302956158	2	0.531473438	1	0.230778791	851799	2	6	369872
PSMs - Concatenated Rank	1.01650571	1	0.145332399	0	0.142972536	375977	1	7	369872
PSMs - Expectation Value	0.008506321	0.001787946	0.013216423	0.010966158	1.553717836	3146.250091	1.36582E-19	0.062900933	369872
PSMs - First Scan	18628.62877	17656	11138.85506	17818	0.597942833	6890208179	37	47149	369872
PSMs - Homology Threshold	17.84616825	17	3.198909224	5	0.179249079	5240795	13	28	293665
PSMs - Identity Relaxed	23.63392201	25	3.248101016	4	0.137433855	8741526	13	29	369872
PSMs - Identity Strict	30.57488266	32	3.44780049	4	0.11276578	11308793	13	36	369872
PSMs - Intensity	411994.6241	161245.08	4049520.624	276167.2513	9.829061806	1.52385E+11	497.5424387	472309673.8	369872
PSMs - Ion Inject Time [ms]	50.05243485	50	46.32631629	23.6732502	0.925555698	18512994.18	0.01	405	369872
PSMs - Ions Score	42.3708514	38.63	16.0202789	19.55	0.378096696	15671791.55	13	189.06	369872
PSMs - Isolation Interference [%]	24.21034591	21.41119003	20.29280653	33.34288645	0.838187385	8952719.605	0	100	369789
PSMs - Last Scan	18628.62877	17656	11138.85506	17818	0.597942833	6890208179	37	47149	369872
PSMs - m/z [Da]	610.9493594	592.8479004	150.1197229	216.8767548	0.245715493	225973061.5	374.204	1368.588257	369872
PSMs - Matched Ions	8.940644331	8	3.653160233	4	0.40860145	3306894	3	47	369872
PSMs - MH+ [Da]	1387.947643	1332.702867	398.5069065	507.2485013	0.287119553	513362970.5	747.4601234	4852.38817	369872
PSMs - Original Precursor Charge	2.302956158	2	0.531473438	1	0.230778791	851799	2	6	369872
PSMs - Peptides Matched	315.2430219	332	151.7696314	247	0.481436926	116599567	1	926	369872
PSMs - Precursor Abundance	3770530.128	347015.5625	48778663.1	996942.2813	12.93681828	1.29103E+12	1319.942261	3558978048	342399
PSMs - Rank	1.003947312	1	0.076458628	0	0.076158008	371332	1	6	369872
PSMs - RT [min]	35.12487057	35.21066658	12.02516657	19.57748062	0.342354758	12991706.13	7.075704602	69.06229308	369872
PSMs - Search Engine Rank	1.01650571	1	0.145332399	0	0.142972536	375977	1	7	369872
PSMs - Theo. MH+ [Da]	1387.947993	1332.700536	398.5067833	507.247517	0.287119392	513363100.2	747.4610811	4852.389837	369872
PSMs - Total Ions	114.6922611	108	48.09247484	60	0.419317523	42421456	40	564	369872
PSMs - ΔCn	4.23541E-05	0	0.001265631	0	29.88212821	15.6656	0	0.05	369872
PSMs - ΔM [ppm]	-0.307557476	-0.584566922	1.679606174	1.499008222	-5.461113139	-113756.8988	-10.01231837	9.984084329	369872
PSMs - Δm/z [Da]	-0.000152428	-0.000325844	0.001143479	0.000842563	-7.501760949	-56.37888414	-0.009015198	0.010827309	369872
PSMs - ΔScore	0.864610257	0.9286	0.187494838	0.1905	0.216854747	315244.6811	0	1	364609
MS/MS Spectrum Info - # PSMs	0.192323832	0	0.402099916	0	2.090744099	369872	0	7	1923173
MS/MS Spectrum Info - First Scan	15231.46535	13295	10567.79198	15300	0.693813217	29292742918	3	47710	1923173
MS/MS Spectrum Info - Ion Inject Time [ms]	70.26806714	50	64.70516652	20.95400238	0.920833163	135137649.5	0.01	405	1923173
MS/MS Spectrum Info - Isolation Interference [%]	29.51677465	26.44338799	24.75437571	41.43119109	0.838654494	56369011.02	0	100	1909728
MS/MS Spectrum Info - Last Scan	15231.46535	13295	10567.79198	15300	0.693813217	29292742918	3	47710	1923173
MS/MS Spectrum Info - Original Mass	652.3559541	626.8167114	177.9248498	238.1844482	0.272741973	1254593357	370.1434	1399.947998	1923173
MS/MS Spectrum Info - Original Precursor Charge	2.376992085	2	0.600460824	1	0.252613725	4571367	2	6	1923173
MS/MS Spectrum Info - Precursor Charge	2.376992085	2	0.600460824	1	0.252613725	4571367	2	6	1923173

MS/MS Spectrum Info - Precursor Intensity	213557.8148	46452.89755	2334027.242	151371.0227	10.92925232	4.10709E+11	21.61298144	472309673.8	1923173
MS/MS Spectrum Info - Precursor m/z [Da]	652.3559541	626.8167114	177.9248498	238.1844482	0.272741973	1254593357	370.1434	1399.947998	1923173
MS/MS Spectrum Info - Precursor MH+ [Da]	1538.281566	1428.227848	551.2313513	581.1628149	0.358342298	2958381575	739.2795234	4999.598776	1923173
MS/MS Spectrum Info - RT [min]	33.45201969	33.14864308	12.60194188	19.59304474	0.376716922	64334021.07	6.012299866	69.09696422	1923173
Consensus Features - # Isotopes	2.828294601	2	1.073049029	1	0.379397899	6367696	1	11	2251426
Consensus Features - # PSMs	0.151898397	0	0.606765537	0	3.994548653	341988	0	28	2251426
Consensus Features - Abundances (Normalized): F1	314833.9195	1424.515476	8869282.148	65200.24473	28.1713043	7.08825E+11	0	2565607859	2251426
Consensus Features - Abundances (Normalized): F2	308451.6701	1634.985413	12021096.99	57789.875	38.97238418	6.94456E+11	0	3558978048	2251426
Consensus Features - Abundances (Normalized): F27	1592416.602	0	99125250.65	0	62.24831525	3.58521E+12	0	31240017314	2251426
Consensus Features - Abundances (Normalized): F28	1170373.197	0	62065790.36	61398.46206	53.03076873	2.63501E+12	0	21653060950	2251426
Consensus Features - Abundances (Normalized): F29	2857699.299	0	210421347.8	0	73.63313134	6.4339E+12	0	92229373061	2251426
Consensus Features - Abundances (Normalized): F3	1142345.686	0	62942361.17	49136.51464	55.09922432	2.57191E+12	0	22181353259	2251426
Consensus Features - Abundances (Normalized): F30	2926949.378	0	200681879.1	97307.62522	68.56349502	6.58981E+12	0	70348920112	2251426
Consensus Features - Abundances (Normalized): F4	1050457.155	0	56942597.8	41397.52739	54.20744437	2.36503E+12	0	18410173385	2251426
Consensus Features - Abundances (per File): F1	303571.3636	1373.556274	8552001.259	62867.83594	28.1713043	6.83468E+11	0	2473828352	2251426
Consensus Features - Abundances (per File): F2	308451.6701	1634.985413	12021096.99	57789.875	38.97238418	6.94456E+11	0	3558978048	2251426
Consensus Features - Abundances (per File): F27	43377.18236	0	2700156.522	0	62.24831525	97660516171	0	850973248	2251426
Consensus Features - Abundances (per File): F28	55048.99403	0	2919290.471	2887.902405	53.03076873	1.23939E+11	0	1018460800	2251426
Consensus Features - Abundances (per File): F29	34398.77777	0	2532889.722	0	73.63313134	77446302640	0	1110185984	2251426
Consensus Features - Abundances (per File): F3	107290.2541	0	5911609.778	4614.950806	55.09922432	2.41556E+11	0	2083294976	2251426
Consensus Features - Abundances (per File): F30	39626.58241	0	2716936.986	1317.401886	68.56349502	89216317936	0	952420736	2251426
Consensus Features - Abundances (per File): F4	117838.1801	0	6387706.594	4643.891724	54.20744437	2.65304E+11	0	2065216384	2251426
Consensus Features - Abundances Count: F1	0.512485865	1	0.51957881	1	1.013840275	1153824	0	7	2251426
Consensus Features - Abundances Count: F2	0.513884978	1	0.517792059	1	1.007603027	1156974	0	7	2251426
Consensus Features - Abundances Count: F27	0.248284421	0	0.449655857	0	1.811051439	558994	0	6	2251426
Consensus Features - Abundances Count: F28	0.287191762	0	0.469100232	1	1.633404207	646591	0	6	2251426
Consensus Features - Abundances Count: F29	0.226378748	0	0.430265863	0	1.900645999	509675	0	7	2251426
Consensus Features - Abundances Count: F3	0.351991138	0	0.49598871	1	1.409094312	792482	0	5	2251426
Consensus Features - Abundances Count: F30	0.266895292	0	0.465703721	1	1.744892976	600895	0	6	2251426
Consensus Features - Abundances Count: F4	0.33983351	0	0.491769307	1	1.44708892	765110	0	5	2251426
Consensus Features - Avg. Apex RT [min]	34.15686315	32.67115726	14.79300682	22.84614764	0.433090321	76901649.77	6.042416096	69.54751187	2251426
Consensus Features - Avg. m/z [Da]	588.5644275	552.9387608	162.0658118	223.1567236	0.27535781	1325109255	375.0216788	1399.732844	2251426
Consensus Features - Charge	2.37724269	2	1.45475301	1	0.611949725	5352186	0	12	2251426
Consensus Features - Fraction No.	6.61378433	7	3.358662346	6	0.507827619	14890446	1	12	2251426
Consensus Features - Left RT [min]	33.75953003	32.20213839	14.80070439	22.86659648	0.438415593	76007083.66	6.000916958	69.49367123	2251426
Consensus Features - Max. Apex S/N	71.5143955	12.54756069	527.2665243	21.86564887	7.372872562	161009369.4	5	45418.12109	2251426
Consensus Features - Right RT [min]	34.58304349	33.15800289	14.79615169	22.86233732	0.427844117	77861163.26	6.086333752	69.59352475	2251426
LCMS Features - # Isotopes	2.129449163	2	1.107326312	2	0.520005986	15177634	1	11	7127493
LCMS Features - Aligned m/z [Da]	585.6580588	552.9435201	156.3686768	220.63798	0.266996543	4174273714	375.0213422	1399.732844	7127493
LCMS Features - Apex RT [min]	33.87854927	32.43117868	14.57580316	21.84757997	0.430236934	241469122.8	6.025818825	69.54971676	7127493
LCMS Features - Apex S/N	39.10979803	6.544229984	352.3536005	11.05107164	9.009343395	278754811.7	1.000031114	45418.12109	7127493
LCMS Features - Area	3749613.925	117245.5938	174610615.8	463043.6094	46.56762518	2.67253E+13	172.289093	93029072896	7127493
LCMS Features - Charge	2.472280646	2	1.631848322	1	0.66005788	17621163	0	12	7127493
LCMS Features - Intensity	330111.2665	14159.41797	10547387.41	54186.26904	31.95100708	2.35287E+12	59.92860031	3558978048	7127493
LCMS Features - Intensity (aligned)	1620190.163	61683.59919	74967646.15	199591.2197	46.27089329	1.15479E+13	0.785402608	32437220452	7127493
LCMS Features - Left RT [min]	33.72749859	32.27890015	14.56616046	21.82295609	0.431877876	240392510.1	6.000916958	69.49367123	7127493
LCMS Features - m/z [Da]	585.6575936	552.9435493	156.3684233	220.6374607	0.266996322	4174270399	375.02105	1399.729259	7127493
LCMS Features - Right RT [min]	34.04627201	32.61019488	14.58636863	21.87874245	0.428427777	242664565.4	6.048110485	69.59352475	7127493

Supplementary Table 4. Mass spectrometry statistics

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell lines

Latent HIV-1 infected Jurkat cells (clone J-Lat 11.1 and clone J-Lat A2) were cultured in RPMI-1640 media supplemented with 7% FBS and 100 µg/ml penicillin-streptomycin at 37°C in a humidified 95% air, 5% CO_2 incubator. Primary CD4+ T cells were cultured in RPMI-1640 media supplemented with 7% FBS and 100 µg/ml penicillin-streptomycin at 37°C in a humidified 95%-air-5%CO2 incubator. HEK 293T cells were cultured in DMEM supplemented with 10% FBS and 100 µg/ml penicillin-streptomycin at 37 ºC in a humidified 95% air 5% CO2 incubator and used exclusively for the production of lentiviral shRNA vectors.

Establishment of J-lat 11.1 cells expressing the HA-V5-FLAG-dCas9 bait

In order to generate J-Lat 11.1 cells expressing the triple tagged (HA, V5, 3xFLAG) dCas9 construct, cells were then nucleofected with a modified version of the PX462 plasmid (AddGene) and selected for resistance to puromycin. pSpCas9n(BB)-2A-Puro (PX462) V2.0 was a gift from Feng Zhang (Addgene plasmid # 62987 ; http://n2t.net/addgene:62987 ; RRID:Addgene_62987). The original plasmid encodes for a Cas9 D10A nickase mutant, we modified the original sequence by introducing an H841A mutation that renders the Cas9 fully catalytically inactive (dCas9). Additionally, we introduced a triple FLAG epitope sequence (3xFLAG), a V5 epitope sequence and an HA epitope sequence at the N-terminal region of the dCas9. The plasmid contains a chicken β-actin promoter that drives the expression of the dCas9 bait and a U6 promoter that drives the expression of the gRNA scaffold. The sequence of the gRNA is 5'-GAAGCGCGCACGGCAAG-3' and is located between nucleotides 708 and 724 on the HIV-1 5' LTR, downstream the transcriptional start site (TSS) and in-between Nuc-1 and Nuc-2 (Figure 1A). Within the plasmid, the dCas9 open reading frame (ORF) sequence additionally contains an IRES sequence that is followed by a puromycin resistance cassette, allowing for puromycin selection of dCas9 expressing cells. The plasmid sequence information is available upon request. 48 hours after nucleofection the cells were selected for a week with 1µg/ml puromycin and expanded in a polyclonal population. Unstimulated J-Lat 11.1 cells were used to model the latent, repressed, promoter while cells stimulated with 10 µm of phorbol 12 myristate 13-acetate (PMA) (Sigma), a potent PKC agonist, were used to model the transcriptionally active promoter.

Plasmids nucleofection

The expression plasmids for the HA-V5-FLAG-dCas9 bait and the dCas9-VPR were delivered to J-Lat 11.1 cells by nucleofection using Amaxa Nucleofector (Lonza) as previously described (Rafati et al., 2011) and the Cell Line Nucleofector Kit R (Lonza). Briefly, cells were split to 4×10^5 cells/ml one day before nucleofection, 8 million cells were centrifuged at 600rcf for 5 min at room temperature, resuspended in 100 µl of solution R, and nucleofected with 2 µgs of plasmid, using program O28 and the Cell line nucleofector Kit R. Nucleofected cells were then resuspended in 500 µl of pre-warmed, serum-free antibiotic-free RPMI at 37 °C for 15 min and then plated in 5 ml of pre-warmed complete media. Seventy-two hours post-nucleofection cells nucleofected with the dCas9 VPR plasmid were analyzed with flow cytometry while cells

nucleofected with the dCas9 construct for purification (HA, V5, FLAG) were selected for 10 days in puromycin 0.5µg/ml and expanded as a polyclonal population.

Bacterial strains

DH5α and Stable3 bacterial cells were used to propagate the plasmids used in the study.

METHOD DETAILS

Western Blotting

$10x10^6$ cells were lysed with 100 µl cold lysis buffer containing 150 mM NaCl, 30 mM Tris (pH 7.5), 1 mM EDTA, 1% Triton X-100, 10% Glycerol, 0.5 mM DTT, protease inhibitor cocktail tablets (EDTA-free) (Roche) at 4°C per 30 minutes. Cell lysates were clarified by centrifugation (14,000 rpm for 30 min at 4°C), mixed with 4x sodium dodecyl sulfate (SDS)-loading buffer with 0.1M DTT and boiled at 95°C per 5 min. Samples were run in a 10% SDS-polyacrylamide gel at 100V. The proteins were then transferred to polyvinylidene difluoride (PVDF) membranes (Immobilion) and the membranes probed with primary antibodies (anti-V5 epitope (Invitrogen), anti-HA epitope (Abcam), anti-FLAG epitope (Sigma) and anti-tubulin (Abcam) overnight at 4 °C. Secondary antibodies were used accordingly to the species reactivity and incubated for 1 hour at room temperature. Proteins were detected by chemical luminescence using the SuperSignal West Pico or SuperSignal West Pico Femto Chemiluminescent Substrate (Thermo Scientific).
For visualization of dCas9 protein enrichment during the ChIP-MS protocol, the protein complexes bound to the beads were decrosslinked and eluted and at 95°C for 30min in 4x sodium dodecyl sulfate (SDS)-loading buffer containing 0.1M DTT. The proteins were then resolved in a 10% polyacrylamide SDS-PAGE gel and western blotting was performed using a mouse anti-FLAG antibody, a secondary anti-mouse HRP antibody, and imaged as above.

Cytoplasmic and nuclear fractionation

Approximately $2x10^7$ J-lat 11.1 cells expressing the HA-V5-FLAG-dCas9 bait were subjected to cellular fractionation to separate the cytoplasmic and nuclear fraction. We adapted the cellular fractionation protocol described in (Ten Have et al., 2012) to cells in suspension. Briefly, the cells were collected by centrifugation at 800rcf for 5 minutes at 4C and washed twice with PBS. The cell pellet was then transferred and resuspended into a pre-chilled 7ml dounce homogenizer, where the cell membrane was broken using 10 strokes of a tight pestle. The dounced cells were then centrifuged for 5 minutes at 4C, 800rcf and the supernatant was retained as cytoplasmic fraction.
The nuclear pellets were then resuspended in 3ml of a solution (S1) containing 0.25M Sucrose and 10mM MgCl2 and layered over a 3ml cushion of a 0.35M Sucrose and 0.5mM MgCl2 solution (S2) by slowly pipetting S1 on top S2. The cell pellets were then centrifuged again for 10 minutes at 4C at 2500rcf and the remaining pellets retained as the nuclear fraction.

Chromatin immunoprecipitation and HIV-1 5'LTR purification using Catchet-MS:

In order to more specifically enrich for chromatin and reduce background from cellular contaminants and from the non-chromatin associated bait, we improved the stringency of our in-house ChIP protocol (Rafati et al., 2011) by incorporating a few steps from the Chromatin enrichment for proteomics (CheP) (Kustatscher et al., 2014) during the preparation of our input material. We have then followed a traditional ChIP protocol in which we have enriched for the dCas9 bait and its associated complexes using an Anti-V5 Agarose Affinity Gel (Sigma). To further eliminate the background originating from the purification of non-chromatin associated bait complexes we have additionally added a histone enrichment step, downstream of the bait enrichment step, using histone H2B (Abcam) and H3 antibodies (Abcam) conjugated to Protein G SepharoseTM 4 Fast flow (GE Healthcare) and Protein A SepharoseTM 4 Fast flow (GE Healthcare) beads by dimethyl pimelimidate cross-linking as described in (Chalkley and Verrijzer, 2004).

Chromatin input preparation using CheIP

Approximately 3 billion cells per condition were collected, cross-linked and processed in batches of 400million cells. After collection, by centrifugation at 800rcf for 5 minutes, each batch was washed two times in 40 ml PBS supplemented with 1mM MgCl2 and 1mM CaCl2, and cross-linked in 40 ml PBS with 1% formaldehyde (Formaldehyde methanol free 16%, Polysciences inc.) at room temperature for 30 min with vertical rotation at 15rpm. The reaction was then quenched with the addition of 1 M glycine to a final concentration of 125 mM. Cross-linked cells pellet were then washed in 20 ml cold buffer B (0.25% Triton x-100, 1mM EDTA, 0.5mM EGTA, 20mM HEPES pH 7.6) and 20 ml cold buffer C (0.15M NaCl, 1mM EDTA, 0.5mM EGTA, 20mM HEPES pH 7.6), resuspended in 2 ml ChIP incubation buffer (1% Triton x-100, 150mM NaCl, 1mM EDTA, ph 8.0, 0.5 mM EGTA, 20mM HEPES pH 7.6, protease inhibitor cocktail tablets (EDTA-free, Roche) and then transferred in 15ml polystyrene Falcon tubes compatible with sonication. In order to isolate and expose the cross-linked nuclei to the subsequent denaturation steps the samples were sonicated with 15 ml probes at medium intensity for 10 minutes (30' ON/30' OFF intervals; Diogenode Bioruptor) and the cell nuclei were spun at 800rcf for 10 minutes. The pellets were then re-suspended in 2 ml ChIP incubation buffer containing 0.15% SDS and sonicated again for 10 minutes at medium intensity (30' ON/30' OFF intervals; Diogenode Bioruptor). The sonicated nuclei in suspension were then aliquoted (500µls per tube) in 2ml Eppendorf tubes and denatured for 5 minutes at room temperature with 350µls 10% SDS and 1.2 ml 8M Urea buffer (10mM Tris pH 7.4, 1mM EDTA and 8 M urea). The insoluble fraction of the nuclei, enriched for chromatin, was then precipitated by centrifugation at 16200rcf at room temperature for 20minutes. The precipitated pellets were then resuspended and combined (4 into 1) in a single 2ml Falcon tube and denatured again for 5 minutes at room temperature with 500µls of 4% SDS buffer (50mM Tris ph 7.4, 10 mM EDTA, 4% SDS) and 8M Urea buffer and then precipitated again by centrifugation at 16200rcf at room temperature for 20minutes.

The pellets were then resuspended and washed by centrifugation at 16200rcf at room temperature for 20minutes in 2ml 4% SDS buffer and then resuspended in 2.5 ml ChIP incubation buffer containing 1% SDS. The content of each tube was then transferred in a 15ml polystyrene tube and sonicated (3 tubes at the time) at a high intensity (30' ON/30' OFF intervals; Diogenode

Bioruptor), using metal probes, to obtain chromatin fragments in a 200-500bp size range. The different chromatin batches were then pulled together concentrated through 100kDa Centricon filters, to buffer exchange with 0.1% SDS ChIP incubation buffer and remove smaller proteins and fragments. The sonicated chromatin (a total of 50 ml per condition) was then spun at max speed (20817 rcf) at 4°C for 30 minutes to remove cellular debris and diluted to 50 ml per condition. The chromatin was then precleared overnight at 4°C with vertical rotation at 15rpm (a total of 5 reactions per condition in 15ml tubes) with 500µls of a mix of Protein G and A Sepharose™ 4 Fast flow (GE Healthcare) beads and then used for immunoprecipitation.

HA-V5-FLAG-dCas9 bait enrichment

Per each condition, dCas9 containing complexes were isolated from the sonicated chromatin using 500µls (a total of 5 reaction where 100µls of affinity beads were used to purify 10ml of chromatin) of Anti-V5 Agarose Affinity Gel (Sigma) and then washed two times per each buffer (5minutes, 15rpm vertical rotation washes) with buffer 1 (0.1% SDS, 0.1% DOC, 1% Triton x-100, 150mM NaCl, 1 mM EDTA pH 8.0, 0.5 mM EGTA, 20mM HEPES pH8.0.), buffer 2 (500mM NaCl: 0.1% SDS, 0.1% DOC, 1% Triton x-100, 500mM NaCl, 1 mM EDTA pH 8.0, 0.5mM EGTA, 20mM HEPES pH8.0), buffer 3 (0.25M LiCL, 0.5%DOC, 0.5% NP-40, 1mM EDTA pH 8.0, 0.5mM EGTA pH 8.0, 20mM HEPES pH 8.0), buffer 4 (1mM EDTA pH8.0, 0.5mM EGTA pH 8.0, 20mM HEPES pH 8.0) to remove unspecific binding and eluted in a total volume of 1 ml elution buffer 1% SDS, 0.1M NaHCO3 per each condition. At this stage a 50µl aliquot of the eluate was taken to be analyzed by qPCR as quality control for dCas9 enrichment over the chromatin input at the desired chromatin locus.

Locus enrichment

The eluate was then diluted 10 times in ChIP incubation buffer without SDS (1% Triton x-100, 150mM NaCl, 1mM EDTA, ph 8.0, 0.5 mM EGTA, 20mM HEPES pH 7.6, protease inhibitor cocktail tablets (EDTA-free) (Roche) and immunoprecipitated again in a single reaction with 100µls of a mix of Protein G and A SepharoseTM 4 Fast flow (GE Healthcare) pre-conjugated with histone H3 and histone H2B antibodies in order to specifically enrich for chromatin-associated dCas9 complexes and eliminate from the purification unspecific non-chromatin associated dCas9 complexes. The immunoprecipitated material, bound to the beads, was then washed again 2 times per each buffer with buffer 1, buffer 2, buffer 3 and buffer 4 to remove unspecific binding and the protein complexes bound to the beads were finally eluted and decrosslinked by boiling the sample for 30min in 100µls of 4x sodium dodecyl sulfate (SDS)-loading buffer, containing 0.1M DTT.

Mass spectrometry

Proteins were resolved in a 15% polyacrylamide SDS-PAGE gel, visualized by Colloidal Blue Staining Kit (ThermoFisher) and prepared for nanoflow LC-MS/MS analysis. SDS-PAGE gel lanes were cut into 1-mm slices and protein were subjected to in-gel reduction with dithiothreitol, alkylation with iodoacetamide and digested with trypsin (TPCK Trypsin, ThermoScientific, Rockford, IL, USA), as described previously (Schwertman et al., 2012). Peptides were extracted

from the gel blocks in 30 % acetonitrile (Biosolve) and 0.5 % formic acid (Biosolve) and dried in a SpeedVac Vacuum Concentrator (ThermoFisher). Samples were analyzed by LC-MS/MS on a 20 cm x 75 µm C18 column (BEH C18, 130 Å, 3.5 µm, Waters) after trapping on a nanoAcquity UPLC Symmetry C18 trapping column (Waters, 100 Å, 5 µm, 180 µm x 20 mm) on an EASY-nLC 1000 coupled to a Fusion Tribrid Orbitrap mass spectrometer (ThermoFisher Scientific), essentially as described in (Sap et al., 2017). Data analysis was performed with the Mascot software suite (Daemon version 2.3.2 and Distiller version 2.3.02, Matrix Science) within the Proteome Discoverer (version 2.2, ThermoFisher Scientific) framework. Spectra were searched against a Uniprot database (version April 2017, taxonomy Homo sapiens) concatenated with a fasta database containing amino acid sequences of the triple tagged dCas9 (HA, V5, 3xFLAG) construct. Protein Mascot scores and peptide numbers were taken directly from the Mascot output and reported. The same procedure was used to analyze input samples for the protocol characterization (Supplementary Figure 3). Heat maps were generated using MORPHEUS (http://software.broadintistute.org/morpheus/index.html).

Chromatin quality control

100µl aliquots of chromatin were taken per each batch of the purification protocol, resuspended in 500µls of ChIP elution buffer and, after addition of 24µl 5M NaCl (200mM final concentration), decrosslinked overnight at 65°C on a heat block. DNA isolation was then performed using phenol-chloroform isoamyl isolation method and ethanol precipitation in the presence of glycogen as a carrier. The DNA pellet was resuspended in 100µls of nuclease-free water. 5µls of DNA was then mixed with sample 4x sample buffer and run on a 1% agarose DNA electrophoreses gel to check for fragmentation between 200 and 500bps.

ChIP- qPCR

Real-time qPCR was used to detect the specific enrichment of the 5'-LTR locus at a DNA level after the ChIP procedure. Reactions of 10 µl were performed with the GoTaq qPCR Master Mix kit (Promega) in a CFX Connect Real-Time PCR thermocycler (Biorad) Relative quantitation was calculated with the 2-ΔCt method where relative enrichment over the adjusted DNA input was used for data representation.
To evaluate IKZF1, CBX8 and SUZ12 binding to the HIV-1 promoter and histone marks profile a smaller scale ChIP experiment (50 million cells) with identical preparation of the chromatin input as for the mass spectrometry experiment. For the immunoprecipitation, protein G-beads or A-beads were used in combination with the anti-IKZF1 antibody (Cell signal technology), anti-SUZ12 (abcam), anti CBX8 (homemade) and histone antibodies: Histone H3 (abcam), histone H3K27me3 (Millipore), H3K4me3 (Millipore), H3K9me3 (Abcam). DNA was PCI extracted and used to evaluate the enrichment of ChIP-Ikaros at the HIV-1 5'LTR by the same primer sets summarized in the key resources table. Antibodies used are summarized in the Key resource table.

ChIP-Sequencing data analysis

In order to align the short reads produced by the ChIP-Seq procedure and also capture reads mapped to the HIV genome, we constructed a custom version of the human genome (UCSC version hg38) where we attached the HIV genome as an additional chromosome. We used the HIV strain K03455.1 (https://www.ncbi.nlm.nih.gov/nuccore/K03455). Subsequently, the resulting genome was indexed and reads were aligned with bwa 0.7.17 (1) and the bwa mem algorithm. In order to compensate for differences in library sizes between ChIP and Input DNA sequencing samples, the total number of reads of each sample was equalized by uniformly downsampling reads relatively to the sample with the lower number of reads. ChIP peaks were then called with MACS2 (2) using default parameters. The peaks returned by MACS2 were further filtered by i) imposing an additional MACS FDR threshold of 0.1%, ii) excluding peaks demonstrating a fold enrichment less than 3 in log2 scale (where fold enrichment is the ratio of normalized reads under a peak area in the ChIP sample to the respective number of normalized reads in the Input DNA sample). Finally, in order to visualize the ChIP-Seq signals, we constructed a custom UCSC track hub based on the combined hg38-HIV genomes using instructions and tools provided by UCSC (Li and Durbin, 2010; Zhang et al., 2008).

Production of shRNA lentiviral vectors

Lentiviral constructs containing the desired shRNA sequence were amplified from bacterial glycerol stocks obtained in house from the Erasmus Center for Biomics and part of the MISSION® shRNA library (sigma). Briefly, 5.0 x 10^6 HEK293T cells were plated in a 10 cm dish and transfected with 12,5 µg of plasmids mix. 4,5 µg of pCMVΔR8.9 (envelope), 2 µg of pCMV-VSV-G (packaging) and 6ug of shRNA vector were mixed in 500 µL serum-free DMEM and combined with 500 µL DMEM containing 125 µL of 10 mM polyethyleneimine (PEI, Sigma). The resulting 1 mL mixture was applied to HEK293T cells after 15 min incubation at room temperature. The transfection mix was removed after 12 hr and replaced with fresh RPMI medium. Virus-containing medium was harvested twice and replaced with fresh medium at 36 hours 48 hr, 60, 72 hr post-transfection. After each harvest, the collected medium was filtered through a cellulose acetate membrane (0.45 µm pore) and used directly for shRNA infections or stored at -80°C for subsequent use. Lentivaral vectors used in the study are summarized in the summarized the key resources table.

Flow Cytometry of J-lat 11.1

GFP expression in the J-Lat cell lines was analyzed by Flow Cytometry. The live-cell population was defined by using the forward scatter area versus side scatter area profiles (FSC-A, SSC-A). Single cells were then gated by using forward scatter height (FSC-H) versus forward scatter width (FSC-W) and side scatter height (SSC-H) versus side scatter width (SSC-W). GFP intensity to differentiate between GFP-positive and negative cells. Between 2 - 4x105 events were collected per sample on an LSR Fortessa (BD Biosciences) and analyzed using FlowJo software (version 9.7.4, Tree Star).

Fluorescence microscopy

Fluorescence microscopy and bright field pictures of J-Lat 11.1 cells in the absence and presence of PMA were acquired using an Olympus IX70 Fluorescence Phase Contrast inverted Microscope.

Total RNA Isolation and quantitative real-time PCR (RT-qPCR)

Total RNA was isolated from the cells using Trizol (Thermo Fisher) and the Total RNA Zol out kit (A&A Biotechnology) and residual genomic DNA digested with DNAseI (Life technologies). cDNA synthesis reactions were performed with Superscript II Reverse Transcriptase (Life Technologies) kit following with random priming and manufacturer's instructions. RT-qPCR reactions were conducted CFX Connect Real-Time PCR Detection System thermocycler (BioRad) using GoTaq qPCR Master Mix (Promega) following manufacturer protocol. Amplification was performed on the using following thermal program starting with 3 min at 95 °C, followed by 40 cycles of 95 °C for 10 s and 60 °C for 30 s. The specificity of the RT-qPCR products was assessed by melting curve analysis. Primers used for real-time PCR are listed in Table 1. Expression data was calculated using 2-ΔΔCt method by Livak Schmittgen (Schmittgen and Livak, 2008). Cyclophilin A (CyPA), GAPDH and β2 microglobulin (B2M) were used as housekeeping genes for the analysis.

Isolation and ex vivo infection of primary CD4+ T cells

Peripheral blood mononuclear cells (PBMCs) were isolated by Ficoll density gradient sedimentation of buffy coats from healthy donors (GE Healthcare). Total CD4+ T cells were separated by negative selection using EasySep Human CD4+ T Cell Enrichment Cocktail (StemCell Technologies). Primary CD4+ T cells were plated at a concentration of $1{,}5 \times 10^{6}$ cells/mL left overnight for recovery. HIV-1 latency *ex vivo* model was generated by spinoculation according to Lassen and Greene method as described elsewhere (Lassen et al., 2012). Pseudotyped viral particles were obtained from co-transfecting HEK 293T cells with HXB2 Env and pNL4.3.Luc.R-E- plasmids using PEI. Supernatants were collected at 48 and 72h post-transfection, filtered with a 0.45 µM filter and stored at -80ºC. Expression vectors HXB2 Env and pNL4.3.Luc.R-E- were donated by Dr. Nathaniel Landau and Drs. Kathleen Page and Dan Littman, respectively. Antiretroviral drugs Saquinavir Mesylate and Raltegravir were kindly provided by the Centre for AIDS reagents, NIBSC.

Flow cytometry for T cells activation and toxicity assay

Primary CD4+ T cells isolated from buffy coats of healthy volunteers were treated with 10uM of iberdomide, 500nM of JQ1, or both compounds, with PMA/ionomycin used as a positive control. The cells were examined by flow-cytometry at 24 and 48 hours, live cells were gated using forward scatter area versus side scatter area profiles (FSC-A, SSC-A). Cells were then stained for expression of Annexin V to examine the percentage of cells undergoing apoptosis and with the surface receptor CD69 to measure T cells activation. For Annexin V and CD69 staining, 10^{6} cells

were washed with PBS supplemented with 3% FBS and 2.5 mM CaCl2 followed by staining with Annexin V-PE (Becton and Dickinson), CD69-FITC (eBIOSCIENCE) for 20 minutes at 4C in the presence of 2.5 mM CaCl2. Cells were then washed with PBS/FBS/CaCl2 and analyzed by flow cytometry. Between 2 - $4x10^5$ events were collected per sample within 3 hours after staining on an LSRFortessa (BD Biosciences, 4 lasers, 18 parameters) and analyzed using FlowJo software (version 9.7.4, Tree Star).

T cell proliferation and functionality assay

To analyze the effect of the LRA on CD8+ and CD4+ T cells, proliferation and cytokine expression were analyzed by flow cytometry. Primary CD8+ and CD4+ T cells were isolated from buffy coats form healthy donors by Ficoll gradient (GE Healthcare) followed by negative selection with RosetteSep Human CD8+ T Cell Enrichment Cocktail or RosetteSep Human CD4+ T Cell Enrichment Cocktail (StemCell Technologies) respectively. Isolated cells were left overnight for recovery. To analyze T cell proliferation capacity, 1 million CD8+ or CD4+ T cells were stained with 0,07 uM CellTrace Far Red Cell Proliferation dye (ThermoFisher Scientific) following manufacturer's instructions. Cells were then cultivated for 72 hours with either unstimulated or stimulated conditions in the presence of the LRA, and analyzed by flow cytometry. Stimulation of T cells was performed using Anti-CD3/CD28 coated Dynabeads (ThermoFisher Scientific) following manufacturer's protocol. Proliferation capacity was determined by a decrease in proliferation dye intensity in daughter cells upon cell division. To analyze T cell functionality by means of cytokine expression 1 million CD8+ or CD4+ T cells were left untreated or treated with the LRA for 18 hours. Cells were then left unstimulated or stimulated with 50 ng/mL PMA and 1uM Ionomycin for 7 hours in the presence of a protein transport inhibitor (BD GolgiPlug™, BD Biosciences). To stain for intracellular cytokines cells were washed with PBS supplemented with 3% FBS followed by a fixation and permeabilization step with FIX & PERM Kit (Invitrogen) following manufacturer's protocol and incubated with 1:25 BV510 anti-IL2 (563265, BD Biosciences) and PE-Cy7 anti-IFNg (27-7319-41, eBioscience) in permeabilization buffer for 45 minutes at 4C. Stained cells were washed with PBS supplemented with 3% FBS and analyzed by flow cytometry.

HIV-1 latency reversal in primary CD4+ T cells isolated from aviremic patients

Primary CD4+ T cells from 5 aviremic patients (maintained viremia below 50 copies/mL for at least two years) were isolated as described previously (Stoszko et al., 2016). Two million CD4+ T cells were plated at the cell density of 10^6/ml and treated as indicated. After 16 hours cells were lysed in TRI reagent and RNA was isolated with Total RNA Zol-out kit (A&A Biotechnology), cDNA synthesis was performed as described above. Absolute quantification of cell-associated HIV-1 gag RNA was performed in a nested PCR as described previously (Pasternak et al., 2008). Briefly, first round of amplification was performed in a final volume of 25 µl using 4 µl of cDNA, 2.5 µl of 10× PCR buffer (Life Technologies), 1 µl of 50mM MgCl2 (Life Technologies), 1 µl of 10 mM dNTPs (Life Technologies), 0.075 µl of 100 µM Gag Forward primer, 0.075 of 100 µM SK437 Reverse primer, and 0.2 µl Platinum Taq polymerase (Life Technologies). The second round of amplification was performed in a final volume of 25 µl using 2 µl of pre-amplified cDNA, 2.5 µl of

10× PCR buffer (Life Technologies), 1 µl of 50mM MgCl2 (Life Technologies), 1 µl of 10 mM dNTPs (Life Technologies), 0.05 µl of 100 µM Gag Forward primer, 0.05 µl of 100 µM Gag Reverse primer, 0.075 µl of 50 µM Gag Probe and 0.2 µl Platinum Taq polymerase. The absolute number of gag copies in the PCR was calculated using a standard curve ranging from 8 to 512 copies of a plasmid containing the full-length HIV-1 genome (pNL4.3.Luc.R-E-). The amount of HIV-1 cellular associated RNA was expressed as number of copies/µg of input RNA in reverse transcription.

The study was conducted according to the ethical principles of the Declaration of Helsinki. All patients involved in the study gave their written informed consent. The study protocol and any amendment were approved by The Netherlands Medical Ethics Committee (MEC-2012-583).

STATISTICAL ANALYSIS

Statistical analysis was conducted as indicated in the figure legends using Prism version 8.3.0 (GraphPad software).

KEY RESOURCES TABLE

Reagent	Source	Identifier
Antibodies		
Cas9 antibody (mAb)	Active Motif	61577
Monoclonal ANTI-FLAG_ M2 antibody produced in mouse	Sigma	F3165-.2MG
Mouse monoclonal V5 Tag	Invitrogen	46-0705
HA tag antibody	Abcam	ab18181
Anti-Flag M2 affinity gel	Sigma	A-2220
Anti-V5 Agarose Affinity Gel antibody produced in mouse	Sigma	A7345-1ml
Anti-Tubulin antibody [YL1/2]	Abcam	ab6160
Anti-HA Agarose	Pierce	26181
Ikaros (D10E5) Rabbit mAb	Cell Signaling Techn	#CST081119Q
Anti-Ikaros antibody - N-terminal	Abcam	ab229275
anti mouse HRP antibody	Promega	W4021
anti Rabbit HRP antibody	Sigma	A0545
Anti-Histone H2B antibody - ChIP Grade	Abcam	ab1790
Histone H3 antibody - ChIP Grade	Abcam	ab1791-100
Anti-timethyl-Histone H3 (Lys4) Antibody	Millipore	07-030
Anti-Histone H3 (tri methyl K9) antibody - ChIP Grade	Abcam	ab8898
Anti-trimethyl-Histone H3 (Lys27) Antibody	Millipore	07-449

Antibody: donkey anti goat IgG HRP	Promega	V8051
Annexin V-PE	Becton and Dickinson	559763
CD69-FITC	eBIOSCIENCE	11-0699-42
Chemicals, peptides and recombinant proteins		
anhydrous DMSO	Sigma	276855-100ml
Iberdomide	MedChemExpress	HY-101291
JQ1	Cayman Chemical	11187-1
Phorbol 12 myristate 13 acetate (PMA)	Sigma	P8139
Random Primers	Invitrogen	48190-011
DNaseI amplification grade	Invitrogen	18068-015
Superscript II Reverse Transcriptase	Invitrogen	18064-014
Oligo(dT)12-18 Primer	Invitrogen	18418-012
100 mM dNTP Set	Invitrogen	10297-018
Formaldehyde methanol free 16%	Polyscience Inc.	18814
Gliotoxin	ApexBio	A4443
Commercial Assays		
GoTaq qPCR Mastermix	Promega	A6002
Platinum Taq DNA Polymerase	Invitrogen	10966-050
Easy sep human CD4 T-cell enrichment kit	Stem Cell Technologies	19052
Ficoll Paque Plus	GE Healthcare	17-1440-02
Total RNA Zol-Out	A en A	030-100
GenElute plasmid Mini prep kit	Sigma	PLN350
Plasmid Midi AX	A en A	A en A
Pure Link Maxiprep kit	Invitrogen	K210017
Nano-Glo® Luciferase Assay System	Promega	N1120
Supersignal West Pico	Thermo Scientific	345788
Supersignal West Femto	Thermo Scientific	34095
Colloidal Blue Staining Kit	Thermo Scientific	LC6025
Cell Line Nucleofector Kit R	Lonza	VVCA-1001
Deposited Data		
Mass spectrometry	this paper	(link)
Western blot original scans	this paper	(link)
Experimental models: cell lines		
J-Lat A2	(Jordan et al., 2003)	NIH AIDS Reagent Program
J-Lat 11.1	(Jordan et al., 2003)	NIH AIDS Reagent Program
HeK 293T	ATCC	CRL-3216
Experimental Models: Organisms/Strains		

DH5a	Thermo	18265017
Recombinant DNA		
sh IKZF1 #1	Sigma	TRCN0000107874
sh IKZF1 #2	Sigma	TRCN0000107871
sh CDC73	Sigma	TRCN0000008728
sh Control	Sigma	SHC002
sh CTR9	Sigma	TRCN0000008739
sh DKC1	Sigma	TRCN0000352996
sh NFRKB	Sigma	TRCN0000014868
sh HP1BP3	Sigma	TRCN0000180320
sh CHD1	Sigma	TRCN0000021309
sh DDX3Y	Sigma	TRCN0000000020
sh EED	Sigma	TRCN0000021204
sh SIN3A	Sigma	TRCN0000021774
sh SUZ12	Sigma	TRCN0000038724
UBE2I	Sigma	TRCN0000007205
FAM19A2	Sigma	TRCN0000138849
CBX8	Sigma	TRCN0000363282
HMG20A	Sigma	TRCN0000015578
UCHL5	Sigma	TRCN0000234906
RNGTT	Sigma	TRCN0000284836
PRPF6	Sigma	TRCN0000293597
PNN	Sigma	TRCN0000072278
DDX39	Sigma	TRCN0000050668
Lentiviral packaging construct pCMVDR8.91	(Naldini et., 1996)	N/A
VSV-G glycoprotein-expressing vector	(Naldini et., 1996)	N/A
Plasmid: SpCas9n-2A-Puro V2.0	(Ran et al., 2013)	62987
Oligonucleotides		
PCR primers	Integrated DNA Technologies IDT	
Softwares and Algorithms		
Prism version 8.3.0	GraphPad	https://www.graphpad.com/scientific-software/prism/
Morpheus heatmapper	Broad institute	https://software.broadinstitute.org/morpheus/
Proteome Discoverer	ThermoFisher Scientific	https://www.thermofisher.com/order/catalog/product/OPTON-30812
Mascot Daemon	Matrix Science	http://www.matrixscience.com/daemon_install.html

Table of Oligonucleotides

Primer	Application	Forward (5' to 3')	Reverse (5' to 3')
Nuc-0	ChIP-qPCR	CCACACACAAGGCTACTTCC	AACTGGTACTAGCTTGTAGCAC
HS1	ChIP-qPCR	TGTGAGCCTGCATGGGATGG	GAAAGTCCCCAGCGGAAAGT
HS1/Nuc1	ChIP-qPCR	AGTGGCGAGCCCTCAGATG	AGCTTTATTGAGGCTTAAGC
Nuc1/HS2	ChIP-qPCR	CGTCTGTTGTGTGACTCTGGT	TCGAGAGAGCTCCTCTGGTT
Nuc1	ChIP-qPCR	TCTCTGGCTAACTAGGGAACC	AAAGGGTCTGAGGGATCTCTAG
HS2/Nuc2	ChIP-qPCR	GCCCGAACAGGGACTTGAAA	TTGGCGTACTCACCAGTCG
Nuc2/Gag	ChIP-qPCR	GGTGCGAGAGCGTCAGTAT	AGCTCCCTGCTTGCCCATA
HIV-1 genome (Vif)	ChIP-qPCR	GGTCTGCATACAGGAGAAAGAG	TTAGTTGGTCTGCTAGGTCAGG
Cyclophilin A (PPIA)	RT-qPCR	TCATCTGCACTGCCAAGACTG	CATGCCTTCTTTCACTTTGCC
GAPDH	RT-qPCR	CAAGAAGGTGGTGAAGCAG	GCCAAATTCGTTGTCATACC
β2-microglobulin	RT-qPCR	ATGAGTATGCCTGCCGTGTG	CCAAATGCGGCATCTTCAAAC
HIV-1 Pol	RT-qPCR	GGTTTATTACAGGGACAGCAGAGA	ACCTGCCATCTGTTTTCCATA
HIV-1 Tat	RT-qPCR	CAAAAGCCTTAGGCATCTCCT	CCACCTTCTTCTTCGATTCCT
GFP	RT-qPCR	GAAGCAGCACGACTTCTTCAA	GCTTGTCGGCCATGATATAGA
IKZF1	RT-qPCR	GGGTCAAGACATGTCCCAAGT	ACATTACTGGCCACGACTCTG
CDC73	RT-qPCR	GTTTATGTCCGACGTGCAGC	TACTTGCCGATGTTGACGCT
DKC1	RT-qPCR	CACTTACCCTCGGAAGTGGG	TTTTGGCAGACTCACTGTAGTCAA
SUZ12	RT-qPCR	GCCTTTGAGAAGCCAACACAG	AGCTGCAAATGAGCTGACAAG
UBE2I	RT-qPCR	GAAAGGGACTCCGTGGGAAG	GCTTGAGCTGGGTCTTGGAT
PNN	RT-qPCR	AAGAGCGCACACGTAGAGAC	CCAAAAGCCGCAGTTCTGTC
HP1BP3	RT-qPCR	GCCGCCGCCATTACG	ATCTTCTACCTTCTCACCTAACTTG
DDX39A	RT-qPCR	CGGCGGAAAACCGAAGTTGG	GCTGTGGATGGAAACGTAGGA
HMG20A	RT-qPCR	AGCTACACATCACTTGACACCA	ATCTGCAAAAAGGGGCGGTA
IKZF2	RT-qPCR	TTCCTTCCTCCTCCCCTTGA	TGAAAGCTCATTGTCACACGTT
IKZF5	RT-qPCR	GAAGCAGAGGCTCTTCAGGG	TGAGGTTTTTCACCTGTGTGG
p21	RT-qPCR	AGCAGAGGAAGACCATGTGGAC	TTTCGACCCTGAGAGTCTCCAG
c-Myc	RT-qPCR	TCTCCACACATCAGCACAACTACGC	CGCCTCTTGACATTCTCCTCGGTG
Gag1 Fw/ SK437 Rv	Nested PCR	TCAGCCCAGAAGTAATACCCATGT	TGCTATGTCAGTTCCCCTTGGTTCTCT
Gag1 Fw/ Gag2 Rv	Nested PCR	TCAGCCCAGAAGTAATACCCATGT	CACTGTGTTTAGCATGGTGTTT
Gag Taqman probe	[6FAM]ATTATCAGAAGGAGCCACCCCACAAGA[BHQ1]		
U6 promoter	Sequencing.	ACTATCATATGCTTACCGTAAC	

References

Battivelli, E., Dahabieh, M.S., Abdel-Mohsen, M., Svensson, J.P., Tojal Da Silva, I., Cohn, L.B., Gramatica, A., Deeks, S., Greene, W.C., Pillai, S.K*., et al.* (2018). Distinct chromatin functional states correlate with HIV latency reactivation in infected primary CD4(+) T cells. Elife *7*.

Bjorklund, C.C., Kang, J., Amatangelo, M., Polonskaia, A., Katz, M., Chiu, H., Couto, S., Wang, M., Ren, Y., Ortiz, M*., et al.* (2020). Iberdomide (CC-220) is a potent cereblon E3 ligase modulator with antitumor and immunostimulatory activities in lenalidomide- and pomalidomide-resistant multiple myeloma cells with dysregulated CRBN. Leukemia *34*, 1197-1201.

Bottardi, S., Mavoungou, L., and Milot, E. (2015). IKAROS: a multifunctional regulator of the polymerase II transcription cycle. Trends Genet *31*, 500-508.

Boutboul, D., Kuehn, H.S., Van de Wyngaert, Z., Niemela, J.E., Callebaut, I., Stoddard, J., Lenoir, C., Barlogis, V., Farnarier, C., Vely, F*., et al.* (2018). Dominant-negative IKZF1 mutations cause a T, B, and myeloid cell combined immunodeficiency. J Clin Invest *128*, 3071-3087.

Bracken, A.P., Dietrich, N., Pasini, D., Hansen, K.H., and Helin, K. (2006). Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. Genes Dev *20*, 1123-1136.

Byrum, S.D., Raman, A., Taverna, S.D., and Tackett, A.J. (2012). ChAP-MS: a method for identification of proteins and histone posttranslational modifications at a single genomic locus. Cell Rep *2*, 198-205.

Chalkley, G.E., and Verrijzer, C.P. (2004). Immuno-depletion and purification strategies to study chromatin-remodeling factors in vitro. Methods Enzymol *377*, 421-442.

Chen, B., Gilbert, L.A., Cimini, B.A., Schnitzbauer, J., Zhang, W., Li, G.W., Park, J., Blackburn, E.H., Weissman, J.S., Qi, L.S*., et al.* (2013). Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. Cell *155*, 1479-1491.

Chun, T.W., Justement, J.S., Lempicki, R.A., Yang, J., Dennis, G., Jr., Hallahan, C.W., Sanford, C., Pandya, P., Liu, S., McLaughlin, M*., et al.* (2003). Gene expression and viral prodution in latently infected, resting CD4+ T cells in viremic versus aviremic HIV-infected individuals. Proc Natl Acad Sci U S A *100*, 1908-1913.

Cismasiu, V.B., Paskaleva, E., Suman Daya, S., Canki, M., Duus, K., and Avram, D. (2008). BCL11B is a general transcriptional repressor of the HIV-1 long terminal repeat in T lymphocytes through recruitment of the NuRD complex. Virology *380*, 173-181.

Coull, J.J., Romerio, F., Sun, J.M., Volker, J.L., Galvin, K.M., Davie, J.R., Shi, Y., Hansen, U., and Margolis, D.M. (2000). The human factors YY1 and LSF repress the human immunodeficiency virus type 1 long terminal repeat via recruitment of histone deacetylase 1. J Virol *74*, 6790-6799.

Davis, K.L. (2011). Ikaros: master of hematopoiesis, agent of leukemia. Ther Adv Hematol *2*, 359-368.

Dejardin, J., and Kingston, R.E. (2009). Purification of proteins associated with specific genomic Loci. Cell *136*, 175-186.

Diaz, T., Rodriguez, V., Lozano, E., Mena, M.P., Calderon, M., Rosinol, L., Martinez, A., Tovar, N., Perez-Galan, P., Blade, J*., et al.* (2017). The BET bromodomain inhibitor CPI203 improves lenalidomide and dexamethasone activity in in vitro and in vivo models of multiple myeloma by blockade of Ikaros and MYC signaling. Haematologica *102*, 1776-1784.

Dieudonne, M., Maiuri, P., Biancotto, C., Knezevich, A., Kula, A., Lusic, M., and Marcello, A. (2009). Transcriptional competence of the integrated HIV-1 provirus at the nuclear periphery. EMBO J *28*, 2231-2243.

Fecteau, J.F., Corral, L.G., Ghia, E.M., Gaidarova, S., Futalan, D., Bharati, I.S., Cathers, B., Schwaederle, M., Cui, B., Lopez-Girona, A*., et al.* (2014). Lenalidomide inhibits the proliferation of CLL cells via a cereblon/p21(WAF1/Cip1)-dependent mechanism independent of functional p53. Blood *124*, 1637-1644.

Fuchs, O. (2019). Treatment of Lymphoid and Myeloid Malignancies by Immunomodulatory Drugs. Cardiovasc Hematol Disord Drug Targets *19*, 51-78.

Gandhi, A.K., Kang, J., Havens, C.G., Conklin, T., Ning, Y., Wu, L., Ito, T., Ando, H., Waldman, M.F., Thakurta, A.*, et al.* (2014). Immunomodulatory agents lenalidomide and pomalidomide co-stimulate T cells by inducing degradation of T cell repressors Ikaros and Aiolos via modulation of the E3 ubiquitin ligase complex CRL4(CRBN.). Br J Haematol *164*, 811-821.
Gauchier, M., van Mierlo, G., Vermeulen, M., and Dejardin, J. (2020). Purification and enrichment of specific chromatin loci. Nat Methods *17*, 380-389.
Geimer Le Lay, A.S., Oravecz, A., Mastio, J., Jung, C., Marchal, P., Ebel, C., Dembele, D., Jost, B., Le Gras, S., Thibault, C.*, et al.* (2014). The tumor suppressor Ikaros shapes the repertoire of notch target genes in T cells. Sci Signal *7*, ra28.
Georgopoulos, K. (2017). The making of a lymphocyte: the choice among disparate cell fates and the IKAROS enigma. Genes Dev *31*, 439-450.
Goffin, V., Demonte, D., Vanhulle, C., de Walque, S., de Launoit, Y., Burny, A., Collette, Y., and Van Lint, C. (2005). Transcription factor binding sites in the pol gene intragenic regulatory region of HIV-1 are important for virus infectivity. Nucleic Acids Res *33*, 4285-4310.
Grau-Exposito, J., Luque-Ballesteros, L., Navarro, J., Curran, A., Burgos, J., Ribera, E., Torrella, A., Planas, B., Badia, R., Martin-Castillo, M.*, et al.* (2019). Latency reversal agents affect differently the latent reservoir present in distinct CD4+ T subpopulations. PLoS Pathog *15*, e1007991.
Grau-Exposito, J., Serra-Peinado, C., Miguel, L., Navarro, J., Curran, A., Burgos, J., Ocana, I., Ribera, E., Torrella, A., Planas, B.*, et al.* (2017). A Novel Single-Cell FISH-Flow Assay Identifies Effector Memory CD4(+) T cells as a Major Niche for HIV-1 Transcription in HIV-Infected Patients. mBio *8*.
Hamperl, S., Brown, C.R., Perez-Fernandez, J., Huber, K., Wittner, M., Babl, V., Stockl, U., Boeger, H., Tschochner, H., Milkereit, P.*, et al.* (2014). Purification of specific chromatin domains from single-copy gene loci in Saccharomyces cerevisiae. Methods Mol Biol *1094*, 329-341.
Harr, J.C., Luperchio, T.R., Wong, X., Cohen, E., Wheelan, S.J., and Reddy, K.L. (2015). Directed targeting of chromatin to the nuclear lamina is mediated by chromatin state and A-type lamins. J Cell Biol *208*, 33-52.
Huang, H., Santoso, N., Power, D., Simpson, S., Dieringer, M., Miao, H., Gurova, K., Giam, C.Z., Elledge, S.J., and Zhu, J. (2015). FACT Proteins, SUPT16H and SSRP1, Are Transcriptional Suppressors of HIV-1 and HTLV-1 That Facilitate Viral Latency. J Biol Chem *290*, 27297-27310.
Ide, S., and Dejardin, J. (2015). End-targeting proteomics of isolated chromatin segments of a mammalian ribosomal RNA gene promoter. Nat Commun *6*, 6674.
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.
Karn, J., and Stoltzfus, C.M. (2012). Transcriptional and posttranscriptional regulation of HIV-1 gene expression. Cold Spring Harb Perspect Med *2*, a006916.
Khan, S., Iqbal, M., Tariq, M., Baig, S.M., and Abbas, W. (2018). Epigenetic regulation of HIV-1 latency: focus on polycomb group (PcG) proteins. Clin Epigenetics *10*, 14.
Kim, Y., Anderson, J.L., and Lewin, S.R. (2018). Getting the "Kill" into "Shock and Kill": Strategies to Eliminate Latent HIV. Cell Host Microbe *23*, 14-26.
Kronke, J., Udeshi, N.D., Narla, A., Grauman, P., Hurst, S.N., McConkey, M., Svinkina, T., Heckl, D., Comer, E., Li, X.*, et al.* (2014). Lenalidomide causes selective degradation of IKZF1 and IKZF3 in multiple myeloma cells. Science *343*, 301-305.
Kustatscher, G., Wills, K.L., Furlan, C., and Rappsilber, J. (2014). Chromatin enrichment for proteomics. Nat Protoc *9*, 2090-2099.
Lassen, K.G., Hebbeler, A.M., Bhattacharyya, D., Lobritz, M.A., and Greene, W.C. (2012). A flexible model of HIV-1 latency permitting evaluation of many primary CD4 T-cell reservoirs. PLoS One *7*, e30176.
Lassen, K.G., Ramyar, K.X., Bailey, J.R., Zhou, Y., and Siliciano, R.F. (2006). Nuclear retention of multiply spliced HIV-1 RNA in resting CD4+ T cells. PLoS Pathog *2*, e68.

Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics *26*, 589-595.
Li, Z., Perez-Casellas, L.A., Savic, A., Song, C., and Dovat, S. (2011). Ikaros isoforms: The saga continues. World J Biol Chem *2*, 140-145.
Liu, X., Zhang, Y., Chen, Y., Li, M., Zhou, F., Li, K., Cao, H., Ni, M., Liu, Y., Gu, Z*., et al.* (2017). In Situ Capture of Chromatin Interactions by Biotinylated dCas9. Cell *170*, 1028-1043 e1019.
Lu, G., Middleton, R.E., Sun, H., Naniong, M., Ott, C.J., Mitsiades, C.S., Wong, K.K., Bradner, J.E., and Kaelin, W.G., Jr. (2014). The myeloma drug lenalidomide promotes the cereblon-dependent destruction of Ikaros proteins. Science *343*, 305-309.
Lusic, M., and Giacca, M. (2015). Regulation of HIV-1 latency by chromatin structure and nuclear architecture. J Mol Biol *427*, 688-694.
Ma, H., Tu, L.C., Naseri, A., Huisman, M., Zhang, S., Grunwald, D., and Pederson, T. (2016). CRISPR-Cas9 nuclear dynamics and target recognition in living cells. J Cell Biol *214*, 529-537.
Malik, B., and Hemenway, C.S. (2013). CBX8, a component of the Polycomb PRC1 complex, modulates DOT1L-mediated gene expression through AF9/MLLT3. FEBS Lett *587*, 3038-3044.
Marke, R., van Leeuwen, F.N., and Scheijen, B. (2018). The many faces of IKZF1 in B-cell precursor acute lymphoblastic leukemia. Haematologica *103*, 565-574.
Moros, A., Rodriguez, V., Saborit-Villarroya, I., Montraveta, A., Balsas, P., Sandy, P., Martinez, A., Wiestner, A., Normant, E., Campo, E*., et al.* (2014). Synergistic antitumor activity of lenalidomide with the BET bromodomain inhibitor CPI203 in bortezomib-resistant mantle cell lymphoma. Leukemia *28*, 2049-2059.
Mousseau, G., and Valente, S.T. (2017). Role of Host Factors on the Regulation of Tat-Mediated HIV-1 Transcription. Curr Pharm Des *23*, 4079-4090.
Mullighan, C.G., Su, X., Zhang, J., Radtke, I., Phillips, L.A., Miller, C.B., Ma, J., Liu, W., Cheng, C., Schulman, B.A*., et al.* (2009). Deletion of IKZF1 and prognosis in acute lymphoblastic leukemia. N Engl J Med *360*, 470-480.
Myers, S.A., Wright, J., Peckner, R., Kalish, B.T., Zhang, F., and Carr, S.A. (2018). Discovery of proteins associated with a predefined genomic locus via dCas9-APEX-mediated proximity labeling. Nat Methods *15*, 437-439.
Ne, E., Palstra, R.J., and Mahmoudi, T. (2018). Transcription: Insights From the HIV-1 Promoter. Int Rev Cell Mol Biol *335*, 191-243.
Ng, S.Y., Yoshida, T., Zhang, J., and Georgopoulos, K. (2009). Genome-wide lineage-specific transcriptional networks underscore Ikaros-dependent lymphoid priming in hematopoietic stem cells. Immunity *30*, 493-507.
Nguyen, K., Das, B., Dobrowolski, C., and Karn, J. (2017). Multiple Histone Lysine Methyltransferases Are Required for the Establishment and Maintenance of HIV-1 Latency. mBio *8*.
Oravecz, A., Apostolov, A., Polak, K., Jost, B., Le Gras, S., Chan, S., and Kastner, P. (2015). Ikaros mediates gene silencing in T cells through Polycomb repressive complex 2. Nat Commun *6*, 8823.
Ott, M., Geyer, M., and Zhou, Q. (2011). The control of HIV transcription: keeping RNA polymerase II on track. Cell Host Microbe *10*, 426-435.
Palstra, R.J., de Crignis, E., Roling, M.D., van Staveren, T., Kan, T.W., van Ijcken, W., Mueller, Y.M., Katsikis, P.D., and Mahmoudi, T. (2018). Allele-specific long-distance regulation dictates IL-32 isoform switching and mediates susceptibility to HIV-1. Sci Adv *4*, e1701729.
Pasternak, A.O., Adema, K.W., Bakker, M., Jurriaans, S., Berkhout, B., Cornelissen, M., and Lukashov, V.V. (2008). Highly sensitive methods based on seminested real-time reverse transcription-PCR for quantitation of human immunodeficiency virus type 1 unspliced and multiply spliced RNA and proviral DNA. J Clin Microbiol *46*, 2206-2211.
Pereira, L.A., Bentley, K., Peeters, A., Churchill, M.J., and Deacon, N.J. (2000). A compilation of cellular transcription factor interactions with the HIV-1 LTR promoter. Nucleic Acids Res *28*, 663-668.

Pourfarzad, F., Aghajanirefah, A., de Boer, E., Ten Have, S., Bryn van Dijk, T., Kheradmandkia, S., Stadhouders, R., Thongjuea, S., Soler, E., Gillemans, N.*, et al.* (2013). Locus-specific proteomics by TChP: targeted chromatin purification. Cell Rep *4*, 589-600.
Powell, M.D., Read, K.A., Sreekumar, B.K., and Oestreich, K.J. (2019). Ikaros Zinc Finger Transcription Factors: Regulators of Cytokine Signaling Pathways and CD4(+) T Helper Cell Differentiation. Front Immunol *10*, 1299.
Rafati, H., Parra, M., Hakre, S., Moshkin, Y., Verdin, E., and Mahmoudi, T. (2011). Repressive LTR nucleosome positioning by the BAF complex is required for HIV latency. PLoS Biol *9*, e1001206.
Rathore, A., Iketani, S., Wang, P., Jia, M., Sahi, V., and Ho, D.D. (2020). CRISPR-based gene knockout screens reveal deubiquitinases involved in HIV-1 latency in two Jurkat cell models. Sci Rep *10*, 5350.
Rojas-Araya, B., Ohlmann, T., and Soto-Rifo, R. (2015). Translational Control of the HIV Unspliced Genomic RNA. Viruses *7*, 4326-4351.
Roux, K.J., Kim, D.I., Raida, M., and Burke, B. (2012). A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. J Cell Biol *196*, 801-810.
Sap, K.A., Bezstarosti, K., Dekkers, D.H.W., Voets, O., and Demmers, J.A.A. (2017). Quantitative Proteomics Reveals Extensive Changes in the Ubiquitinome after Perturbation of the Proteasome by Targeted dsRNA-Mediated Subunit Knockdown in Drosophila. J Proteome Res *16*, 2848-2862.
Schafer, P.H., Ye, Y., Wu, L., Kosek, J., Ringheim, G., Yang, Z., Liu, L., Thomas, M., Palmisano, M., and Chopra, R. (2018). Cereblon modulator iberdomide induces degradation of the transcription factors Ikaros and Aiolos: immunomodulation in healthy volunteers and relevance to systemic lupus erythematosus. Ann Rheum Dis *77*, 1516-1523.
Schmidtmann, E., Anton, T., Rombaut, P., Herzog, F., and Leonhardt, H. (2016). Determination of local chromatin composition by CasID. Nucleus *7*, 476-484.
Schmittgen, T.D., and Livak, K.J. (2008). Analyzing real-time PCR data by the comparative C(T) method. Nat Protoc *3*, 1101-1108.
Schwertman, P., Lagarou, A., Dekkers, D.H., Raams, A., van der Hoek, A.C., Laffeber, C., Hoeijmakers, J.H., Demmers, J.A., Fousteri, M., Vermeulen, W.*, et al.* (2012). UV-sensitive syndrome protein UVSSA recruits USP7 to regulate transcription-coupled repair. Nat Genet *44*, 598-602.
Schwickert, T.A., Tagoh, H., Gultekin, S., Dakic, A., Axelsson, E., Minnich, M., Ebert, A., Werner, B., Roth, M., Cimmino, L.*, et al.* (2014). Stage-specific control of early B cell development by the transcription factor Ikaros. Nat Immunol *15*, 283-293.
Siliciano, J.D., Kajdas, J., Finzi, D., Quinn, T.C., Chadwick, K., Margolick, J.B., Kovacs, C., Gange, S.J., and Siliciano, R.F. (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.
Siliciano, J.M., and Siliciano, R.F. (2015). The Remarkable Stability of the Latent Reservoir for HIV-1 in Resting Memory CD4+ T Cells. J Infect Dis *212*, 1345-1347.
Song, C., Pan, X., Ge, Z., Gowda, C., Ding, Y., Li, H., Li, Z., Yochum, G., Muschen, M., Li, Q.*, et al.* (2016). Epigenetic regulation of gene expression by Ikaros, HDAC1 and Casein Kinase II in leukemia. Leukemia *30*, 1436-1440.
Stoszko, M., De Crignis, E., Rokx, C., Khalid, M.M., Lungu, C., Palstra, R.J., Kan, T.W., Boucher, C., Verbon, A., Dykhuizen, E.C.*, et al.* (2016). Small Molecule Inhibitors of BAF; A Promising Family of Compounds in HIV-1 Latency Reversal. EBioMedicine *3*, 108-121.
Stoszko, M., Ne, E., Abner, E., and Mahmoudi, T. (2019). A broad drug arsenal to attack a strenuous latent HIV reservoir. Curr Opin Virol *38*, 37-53.
Sun, L., Liu, A., and Georgopoulos, K. (1996). Zinc finger-mediated protein interactions modulate Ikaros activity, a molecular control of lymphocyte development. EMBO J *15*, 5358-5369.

Tsui, C., Inouye, C., Levy, M., Lu, A., Florens, L., Washburn, M.P., and Tjian, R. (2018). dCas9-targeted locus-specific protein isolation method identifies histone gene regulators. Proc Natl Acad Sci U S A *115*, E2734-E2741.
Turner, A.W., Dronamraju, R., Potjewyd, F., James, K.S., Winecoff, D.K., Kirchherr, J.L., Archin, N.M., Browne, E.P., Strahl, B.D., Margolis, D.M.*, et al.* (2020). Evaluation of EED Inhibitors as a Class of PRC2-Targeted Small Molecules for HIV Latency Reversal. ACS Infect Dis *6*, 1719-1733.
Vermeulen, M., and Dejardin, J. (2020). Locus-specific chromatin isolation. Nat Rev Mol Cell Biol *21*, 249-250.
Vidal, M., and Starowicz, K. (2017). Polycomb complexes PRC1 and their function in hematopoiesis. Exp Hematol *48*, 12-31.
Waldrip, Z.J., Byrum, S.D., Storey, A.J., Gao, J., Byrd, A.K., Mackintosh, S.G., Wahls, W.P., Taverna, S.D., Raney, K.D., and Tackett, A.J. (2014). A CRISPR-based approach for proteomic analysis of a single genomic locus. Epigenetics *9*, 1207-1211.
Yukl, S.A., Kaiser, P., Kim, P., Telwatte, S., Joshi, S.K., Vu, M., Lampiris, H., and Wong, J.K. (2018). HIV latency in isolated patient CD4(+) T cells may be due to blocks in HIV transcriptional elongation, completion, and splicing. Sci Transl Med *10*.
Zhang, J., Jackson, A.F., Naito, T., Dose, M., Seavitt, J., Liu, F., Heller, E.J., Kashiwagi, M., Yoshida, T., Gounari, F.*, et al.* (2011). Harnessing of the nucleosome-remodeling-deacetylase complex controls lymphocyte development and prevents leukemogenesis. Nat Immunol *13*, 86-94.
Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W.*, et al.* (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol *9*, R137.
Zhao, Y., Karijolich, J., Glaunsinger, B., and Zhou, Q. (2016). Pseudouridylation of 7SK snRNA promotes 7SK snRNP formation to suppress HIV-1 transcription and escape from latency. EMBO Rep *17*, 1441-1451.